



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) EP 0 786 519 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
30.07.1997 Bulletin 1997/31

(51) Int Cl.⁶: C12N 15/00

(21) Application number: 97100117.7

(22) Date of filing: 07.01.1997

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE
Designated Extension States:
AL LT LV RO SI

(30) Priority: 05.01.1996 US 9861

(71) Applicant: HUMAN GENOME SCIENCES, INC.
Rockville, MD 20850-3338 (US)

(72) Inventors:
• Kunsch, Charles A.
Gaithersburg, Maryland 20882 (US)

- Choi, Gil H.
Rockville, Maryland 20850 (US)
- Barash, Steven C.
Rockville, Maryland 20850 (US)
- Dillon, Patrick J.
Gaithersburg, Maryland 20879 (US)
- Fannon, Michael R.
Silver Spring, Maryland 20906 (US)
- Rosen, Craig A.
Laytonsville, Maryland 20882 (US)

(74) Representative: VOSSIUS & PARTNER
Postfach 86 07 67
81634 München (DE)

(54) **Staphylococcus aureus polynucleotides and sequences**

(57) The present invention provides polynucleotide sequences of the genome of *Staphylococcus aureus*, polypeptide sequences encoded by the polynucleotide sequences, corresponding polynucleotides and polypeptides, vectors and hosts comprising the polynu-

cleotides, and assays and other uses thereof. The present invention further provides polynucleotide and polypeptide sequence information stored on computer readable media, and computer-based systems and methods which facilitate its use.

EP 0 786 519 A2

Description

The present invention relates to the field of molecular biology. In particular, it relates to, among other things, nucleotide sequences of *Staphylococcus aureus*, contigs, ORFs, fragments, probes, primers and related polynucleotides thereof, peptides and polypeptides encoded by the sequences, and uses of the polynucleotides and sequences thereof, such as in fermentation, polypeptide production, assays and pharmaceutical development, among others.

The genus *Staphylococcus* includes at least 20 distinct species. (For a review see Novick, R. P., *The Staphylococcus as a Molecular Genetic System*, Chapter 1, pgs. 1-37 in *MOLECULAR BIOLOGY OF THE STAPHYLOCOCCI*, R. Novick, Ed., VCH Publishers, New York (1990)). Species differ from one another by 80% or more, by hybridization kinetics, whereas strains within a species are at least 90% identical by the same measure.

The species *Staphylococcus aureus*, a gram-positive, facultatively aerobic, clump-forming cocci, is among the most important etiological agents of bacterial infection in humans, as discussed briefly below.

Human Health and *S. Aureus*

Staphylococcus aureus is a ubiquitous pathogen. (See, for instance, Mims *et al.*, *MEDICAL MICROBIOLOGY*, Mosby-Year Book Europe Limited, London, UK (1993)). It is an etiological agent of a variety of conditions, ranging in severity from mild to fatal. A few of the more common conditions caused by *S. aureus* infection are burns, cellulitis, eyelid infections, food poisoning, joint infections, neonatal conjunctivitis, osteomyelitis, skin infections, surgical wound infection, scalded skin syndrome and toxic shock syndrome, some of which are described further below.

Burns

Burn wounds generally are sterile initially. However, they generally compromise physical and immune barriers to infection, cause loss of fluid and electrolytes and result in local or general physiological dysfunction. After cooling, contact with viable bacteria results in mixed colonization at the injury site. Infection may be restricted to the non-viable debris on the burn surface ("eschar"), it may progress into full skin infection and invade viable tissue below the eschar and it may reach below the skin, enter the lymphatic and blood circulation and develop into septicaemia. *S. aureus* is among the most important pathogens typically found in burn wound infections. It can destroy granulation tissue and produce severe septicaemia.

Cellulitis

Cellulitis, an acute infection of the skin that expands from a typically superficial origin to spread below the cutaneous layer, most commonly is caused by *S. aureus* in conjunction with *S. pyogenes*. Cellulitis can lead to systemic infection. In fact, cellulitis can be one aspect of synergistic bacterial gangrene. This condition typically is caused by a mixture of *S. aureus* and microaerophilic *streptococci*. It causes necrosis and treatment is limited to excision of the necrotic tissue. The condition often is fatal.

Eyelid infections

S. aureus is the cause of styes and of "sticky eye" in neonates, among other eye infections. Typically such infections are limited to the surface of the eye, and may occasionally penetrate the surface with more severe consequences.

Food poisoning

Some strains of *S. aureus* produce one or more of five serologically distinct, heat and acid stable enterotoxins that are not destroyed by digestive process of the stomach and small intestine (enterotoxins A-E). Ingestion of the toxin, in sufficient quantities, typically results in severe vomiting, but not diarrhoea. The effect does not require viable bacteria. Although the toxins are known, their mechanism of action is not understood.

Joint infections

S. aureus infects bone joints causing diseases such as osteomyelitis.

Osteomyelitis

S. aureus is the most common causative agent of haematogenous osteomyelitis. The disease tends to occur in

children and adolescents more than adults and it is associated with non-penetrating injuries to bones. Infection typically occurs in the long end of growing bone, hence its occurrence in physically immature populations. Most often, infection is localized in the vicinity of sprouting capillary loops adjacent to epiphysal growth plates in the end of long, growing bones.

Skin infections

S. aureus is the most common pathogen of such minor skin infections as abscesses and boils. Such infections often are resolved by normal host response mechanisms, but they also can develop into severe internal infections. Recurrent infections of the nasal passages plague nasal carriers of *S. aureus*.

Surgical Wound Infections

Surgical wounds often penetrate far into the body. Infection of such wound thus poses a grave risk to the patient. *S. aureus* is the most important causative agent of infections in surgical wounds. *S. aureus* is unusually adept at invading surgical wounds; sutured wounds can be infected by far fewer *S. aureus* cells than are necessary to cause infection in normal skin. Invasion of surgical wound can lead to severe *S. aureus* septicaemia. Invasion of the blood stream by *S. aureus* can lead to seeding and infection of internal organs, particularly heart valves and bone, causing systemic diseases, such as endocarditis and osteomyelitis.

Scalded Skin Syndrome

S. aureus is responsible for "scalded skin syndrome" (also called toxic epidermal necrosis, Ritter's disease and Lyell's disease). This disease occurs in older children, typically in outbreaks caused by flowering of *S. aureus* strains produce exfoliation (also called scalded skin syndrome toxin). Although the bacteria initially may infect only a minor lesion, the toxin destroys intercellular connections, spreads epidermal layers and allows the infection to penetrate the outer layer of the skin, producing the desquamation that typifies the disease. Shedding of the outer layer of skin generally reveals normal skin below, but fluid lost in the process can produce severe injury in young children if it is not treated properly.

Toxic Shock Syndrome

Toxic shock syndrome is caused by strains of *S. aureus* that produce the so-called toxic shock syndrome toxin. The disease can be caused by *S. aureus* infection at any site, but it is too often erroneously viewed exclusively as a disease solely of women who use tampons. The disease involves toxemia and septicaemia, and can be fatal.

Nocosomal Infections

In the 1984 National Nosocomial Infection Surveillance Study ("NNIS") *S. aureus* was the most prevalent agent of surgical wound infections in many hospital services, including medicine, surgery, obstetrics, pediatrics and newborns.

Resistance to drugs of S. aureus strains

Prior to the introduction of penicillin the prognosis for patients seriously infected with *S. aureus* was unfavorable. Following the introduction of penicillin in the early 1940s even the worst *S. aureus* infections generally could be treated successfully. The emergence of penicillin-resistant strains of *S. aureus* did not take long, however. Most strains of *S. aureus* encountered in hospital infections today do not respond to penicillin; although, fortunately, this is not the case for *S. aureus* encountered in community infections.

It is well known now that penicillin-resistant strains of *S. aureus* produce a lactamase which converts penicillin to penicilloic acid, and thereby destroys antibiotic activity. Furthermore, the lactamase gene often is propagated episomally, typically on a plasmid, and often is only one of several genes on an episomal element that, together, confer multidrug resistance.

Methicillins, introduced in the 1960s, largely overcame the problem of penicillin resistance in *S. aureus*. These compounds conserve the portions of penicillin responsible for antibiotic activity and modify or alter other portions that make penicillin a good substrate for inactivating lactamases. However, methicillin resistance has emerged in *S. aureus*, along with resistance to many other antibiotics effective against this organism, including aminoglycosides, tetracycline, chloramphenicol, macrolides and lincosamides. In fact, methicillin-resistant strains of *S. aureus* generally are multiply drug resistant.

The molecular genetics of most types of drug resistance in *S. aureus* has been elucidated (See Lyon *et al.*, *Microbiology Reviews* 51: 88-134 (1987)). Generally, resistance is mediated by plasmids, as noted above regarding penicillin resistance; however, several stable forms of drug resistance have been observed that apparently involve integration of a resistance element into the *S. aureus* genome itself.

Thus far each new antibiotic gives rise to resistance strains, strains emerge that are resistance to multiple drugs and increasingly persistent forms of resistance begin to emerge. Drug resistance of *S. aureus* infections already poses significant treatment difficulties, which are likely to get much worse unless new therapeutic agents are developed.

Molecular Genetics of *Staphylococcus Aureus*

Despite its importance in, among other things, human disease, relatively little is known about the genome of this organism.

Most genetic studies of *S. aureus* have been carried out using the strain NCTC8325, which contains prophages psi11 psi12 and psi13, and the UV-cured derivative of this strain, 8325-4 (also referred to as RN450), which is free of the prophages.

These studies revealed that the *S. aureus* genome, like that of other *staphylococci*, consists of one circular, covalently closed, double-stranded DNA and a collection of so-called variable accessory genetic elements, such as prophages, plasmids, transposons and the like.

Physical characterization of the genome has not been carried out in any detail. Pattee *et al.* published a low resolution and incomplete genetic and physical map of the chromosome of *S. aureus* strain NCTC 8325. (Pattee *et al.* Genetic and Physical Mapping of Chromosome of *Staphylococcus aureus* NCTC 8325, Chapter 11, pgs. 163-169 in. MOLECULAR BIOLOGY OF THE STAPHYLOCOCCI, R.P. Novick, Ed., VCH Publishers, New York, (1990) The genetic map largely was produced by mapping insertions of Tn551 and Tn4001, which, respectively, confer erythromycin and gentamicin resistance, and by analysis of SmaI-digested DNA by Pulsed Field Gel Electrophoresis ("PFGE").

The map was of low resolution; even estimating the physical size of the genome was difficult, according to the investigators. The size of the largest SmaI chromosome fragment, for instance, was too large for accurate sizing by PFGE. To estimate its size, additional restriction sites had to be introduced into the chromosome using a transposon containing a SmaI recognition sequence.

In sum, most physical characteristics and almost all of the genes of *Staphylococcus aureus* are unknown. Among the few genes that have been identified, most have not been physically mapped or characterized in detail. Only a very few genes of this organism have been sequenced. (See, for instance Thornsberry, J., *Antimicrobial Chemotherapy* 21 Suppl C: 9-16 (1988), current versions of GENBANK and other nucleic acid databases, and references that relate to the genome of *S. aureus* such as those set out elsewhere herein.)

It is clear that the etiology of diseases mediated or exacerbated by *S. aureus* infection involves the programmed expression of *S. aureus* genes, and that characterizing the genes and their patterns of expression would add dramatically to our understanding of the organism and its host interactions. Knowledge of *S. aureus* genes and genomic organization would dramatically improve understanding of disease etiology and lead to improved and new ways of preventing, ameliorating, arresting and reversing diseases. Moreover, characterized genes and genomic fragments of *S. aureus* would provide reagents for, among other things, detecting, characterizing and controlling *S. aureus* infections. There is a need therefore to characterize the genome of *S. aureus* and for polynucleotides and sequences of this organism.

The present invention is based on the sequencing of fragments of the *Staphylococcus aureus* genome. The primary nucleotide sequences which were generated are provided in SEQ ID NOS: 1-5,191.

The present invention provides the nucleotide sequence of several thousand contigs of the *Staphylococcus aureus* genome, which are listed in tables below and set out in the Sequence Listing submitted herewith, and representative fragments thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan. In one embodiment, the present invention is provided as contiguous strings of primary sequence information corresponding to the nucleotide sequences depicted in SEQ ID NOS:1-5,191.

The present invention further provides nucleotide sequences which are at least 95%, preferably 99% and most preferably 99.9%, identical to the nucleotide sequences of SEQ ID NOS:1-5,191.

The nucleotide sequence of SEQ ID NOS:1-5,191, a representative fragment thereof, or a nucleotide sequence which is at least 95%, preferably 99% and most preferably 99.9%, identical to the nucleotide sequence of SEQ ID NOS:1-5,191 may be provided in a variety of mediums to facilitate its use. In one application of this embodiment, the sequences of the present invention are recorded on computer readable media. Such media includes, but is not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media.

The present invention further provides systems, particularly computer-based systems which contain the sequence

information herein described stored in a data storage means. Such systems are designed to identify commercially important fragments of the *Staphylococcus aureus* genome.

Another embodiment of the present invention is directed to fragments, preferably isolated fragments, of the *Staphylococcus aureus* genome having particular structural or functional attributes. Such fragments of the *Staphylococcus aureus* genome of the present invention include, but are not limited to, fragments which encode peptides, hereinafter referred to as open reading frames or ORFs," fragments which modulate the expression of an operably linked ORF, hereinafter referred to as expression modulating fragments or EMFs," and fragments which can be used to diagnose the presence of *Staphylococcus aureus* in a sample, hereinafter referred to as diagnostic fragments or "DFs."

Each of the ORFs in fragments of the *Staphylococcus aureus* genome disclosed in Tables 1-3, and the EMFs found 5' to the ORFs, can be used in numerous ways as polynucleotide reagents. For instance, the sequences can be used as diagnostic probes or amplification primers for detecting or determining the presence of a specific microbe in a sample, to selectively control gene expression in a host and in the production of polypeptides, such as polypeptides encoded by ORFs of the present invention, particular those polypeptides that have a pharmacological activity.

The present invention further includes recombinant constructs comprising one or more fragments of the *Staphylococcus aureus* genome of the present invention. The recombinant constructs of the present invention comprise vectors, such as a plasmid or viral vector, into which a fragment of the *Staphylococcus aureus* has been inserted.

The present invention further provides host cells containing any of the isolated fragments of the *Staphylococcus aureus* genome of the present invention. The host cells can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic cell, such as a yeast cell, or a procaryotic cell such as a bacterial cell.

The present invention is further directed to polypeptides and proteins, preferably isolated polypeptides and proteins, encoded by ORFs of the present invention. A variety of methods, well known to those of skill in the art, routinely may be utilized to obtain any of the polypeptides and proteins of the present invention. For instance, polypeptides and proteins of the present invention having relatively short, simple amino acid sequences readily can be synthesized using commercially available automated peptide synthesizers. Polypeptides and proteins of the present invention also may be purified from bacterial cells which naturally produce the protein. Yet another alternative is to purify polypeptide and proteins of the present invention can from cells which have been altered to express them.

The invention further provides polypeptides, preferably isolated polypeptides, comprising *Staphylococcus aureus* epitopes and vaccine compositions comprising such polypeptides. Also provided are methods for vaccinating an individual against *Staphylococcus aureus* infection.

The invention further provides methods of obtaining homologs of the fragments of the *Staphylococcus aureus* genome of the present invention and homologs of the proteins encoded by the ORFs of the present invention. Specifically, by using the nucleotide and amino acid sequences disclosed herein as a probe or as primers, and techniques such as PCR cloning and colony/plaque hybridization, one skilled in the art can obtain homologs.

The invention further provides antibodies which selectively bind polypeptides and proteins of the present invention. Such antibodies include both monoclonal and polyclonal antibodies.

The invention further provides hybridomas which produce the above-described antibodies. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

The present invention further provides methods of identifying test samples derived from cells which express one of the ORFs of the present invention, or a homolog thereof. Such methods comprise incubating a test sample with one or more of the antibodies of the present invention, or one or more of the DFs or antigens of the present invention, under conditions which allow a skilled artisan to determine if the sample contains the ORF or product produced therefrom.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the above-described assays.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the antibodies, antigens, or one of the DFs of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of bound antibodies, antigens or hybridized DFs.

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents capable of binding to a polypeptide or protein encoded by one of the ORFs of the present invention. Specifically, such agents include, as further described below, antibodies, peptides, carbohydrates, pharmaceutical agents and the like. Such methods comprise steps of: (a) contacting an agent with an isolated protein encoded by one of the ORFs of the present invention; and (b) determining whether the agent binds to said protein.

The present genomic sequences of *Staphylococcus aureus* will be of great value to all laboratories working with this organism and for a variety of commercial purposes. Many fragments of the *Staphylococcus aureus* genome will be immediately identified by similarity searches against GenBank or protein databases and will be of immediate value to *Staphylococcus aureus* researchers and for immediate commercial value for the production of proteins or to control gene expression.

The methodology and technology for elucidating extensive genomic sequences of bacterial and other genomes

has and will greatly enhance the ability to analyze and understand chromosomal organization. In particular, sequenced contigs and genomes will provide the models for developing tools for the analysis of chromosome structure and function, including the ability to identify genes within large segments of genomic DNA, the structure, position, and spacing of regulatory elements, the identification of genes with potential industrial applications, and the ability to do comparative genomic and molecular phylogeny.

FIGURE 1 is a block diagram of a computer system (102) that can be used to implement computer-based systems of present invention.

FIGURE 2 is a schematic diagram depicting the data flow and computer programs used to collect, assemble, edit and annotate the contigs of the *Staphylococcus aureus* genome of the present invention. Both Macintosh and Unix platforms are used to handle the AB 373 and 377 sequence data files, largely as described in Kerlavage *et al.*, *Proceedings of the Twenty-Sixth Annual Hawaii International Conference on System Sciences*, 585, IEEE Computer Society Press, Washington D.C. (1993). Factura (AB) is a Macintosh program designed for automatic vector sequence removal and end-trimming of sequence files. The program Loadis runs on a Macintosh platform and parses the feature data extracted from the sequence files by Factura to the Unix based *Staphylococcus aureus* relational database. Assembly of contigs (and whole genome sequences) is accomplished by retrieving a specific set of sequence files and their associated features using extrseq, a Unix utility for retrieving sequences from an SQL database. The resulting sequence file is processed by seq_filter to trim portions of the sequences with more than 2% ambiguous nucleotides. The sequence files were assembled using TIGR Assembler, an assembly engine designed at The Institute for Genomic Research (TIGR) for rapid and accurate assembly of thousands of sequence fragments. The collection of contigs generated by the assembly step is loaded into the database with the lassie program. Identification of open reading frames (ORFs) is accomplished by processing contigs with zorf. The ORFs are searched against *S. aureus* sequences from Genbank and against all protein sequences using the BLASTN and BLASTP programs, described in Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990). Results of the ORF determination and similarity searching steps were loaded into the database. As described below, some results of the determination and the searches are set out in Tables 1-3.

The present invention is based on the sequencing of fragments of the *Staphylococcus aureus* genome and analysis of the sequences. The primary nucleotide sequences generated by sequencing the fragments are provided in SEQ ID NOS:1-5,191. (As used herein, the "primary sequence" refers to the nucleotide sequence represented by the IUPAC nomenclature system.)

In addition to the aforementioned *Staphylococcus aureus* polynucleotide and polynucleotide sequences, the present invention provides the nucleotide sequences of SEQ ID NOS:1-5,191, or representative fragments thereof, in a form which can be readily used, analyzed, and interpreted by a skilled artisan.

As used herein, a "representative fragment of the nucleotide sequence depicted in SEQ ID NOS:1-5,191" refers to any portion of the SEQ ID NOS:1-5,191 which is not presently represented within a publicly available database. Preferred representative fragments of the present invention are *Staphylococcus aureus* open reading frames (ORFs), expression modulating fragment (EMFs) and fragments which can be used to diagnose the presence of *Staphylococcus aureus* in sample ("DFs"). A non-limiting identification of preferred representative fragments is provided in Tables 1-3.

As discussed in detail below, the information provided in SEQ ID NOS:1-5,191 and in Tables 1-3 together with routine cloning, synthesis, sequencing and assay methods will enable those skilled in the art to clone and sequence all "representative fragments" of interest, including open reading frames encoding a large variety of *Staphylococcus aureus* proteins.

While the presently disclosed sequences of SEQ ID NOS:1-5,191 are highly accurate, sequencing techniques are not perfect and, in relatively rare instances, further investigation of a fragment or sequence of the invention may reveal a nucleotide sequence error present in a nucleotide sequence disclosed in SEQ ID NOS:1-5,191. However, once the present invention is made available (*i.e.*, once the information in SEQ ID NOS:1-5,191 and Tables 1-3 has been made available), resolving a rare sequencing error in SEQ ID NOS:1-5,191 will be well within the skill of the art. The present disclosure makes available sufficient sequence information to allow any of the described contigs or portions thereof to be obtained readily by straightforward application of routine techniques. Further sequencing of such polynucleotide may proceed in like manner using manual and automated sequencing methods which are employed ubiquitous in the art. Nucleotide sequence editing software is publicly available. For example, Applied Biosystem's (AB) AutoAssembler can be used as an aid during visual inspection of nucleotide sequences. By employing such routine techniques potential errors readily may be identified and the correct sequence then may be ascertained by targeting further sequencing effort, also of a routine nature, to the region containing the potential error.

Even if all of the very rare sequencing errors in SEQ ID NOS:1-5,191 were corrected, the resulting nucleotide sequences would still be at least 95% identical, nearly all would be at least 99% identical, and the great majority would be at least 99.9% identical to the nucleotide sequences of SEQ ID NOS:1-5,191.

As discussed elsewhere herein, polynucleotides of the present invention readily may be obtained by routine application of well known and standard procedures for cloning and sequencing DNA. Detailed methods for obtaining

libraries and for sequencing are provided below, for instance. A wide variety of *Staphylococcus aureus* strains that can be used to prepare *S. aureus* genomic DNA for cloning and for obtaining polynucleotides of the present invention are available to the public from recognized depository institutions, such as the American Type Culture Collection (ATCC).

The nucleotide sequences of the genomes from different strains of *Staphylococcus aureus* differ somewhat. However, the nucleotide sequences of the genomes of all *Staphylococcus aureus* strains will be at least 95% identical, in corresponding part, to the nucleotide sequences provided in SEQ ID NOS:1-5,191. Nearly all will be at least 99% identical and the great majority will be 99.9% identical.

Thus, the present invention further provides nucleotide sequences which are at least 95%, preferably 99% and most preferably 99.9% identical to the nucleotide sequences of SEQ ID NOS:1-5,191, in a form which can be readily used, analyzed and interpreted by the skilled artisan.

Methods for determining whether a nucleotide sequence is at least 95%, at least 99% or at least 99.9% identical to the nucleotide sequences of SEQ ID NOS:1-5,191 are routine and readily available to the skilled artisan. For example, the well known fasta algorithm described in Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* **85**: 2444 (1988) can be used to generate the percent identity of nucleotide sequences. The BLASTN program also can be used to generate an identity score of polynucleotides compared to one another.

COMPUTER RELATED EMBODIMENTS

The nucleotide sequences provided in SEQ ID NOS:1-5,191, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 99% and most preferably at least 99.9% identical to a polynucleotide sequence of SEQ ID NOS:1-5,191 may be "provided" in a variety of mediums to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid molecule, which contains a nucleotide sequence of the present invention; i.e., a nucleotide sequence provided in SEQ ID NOS:1-5,191, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 99% and most preferably at least 99.9% identical to a polynucleotide of SEQ ID NOS:1-5,191. Such a manufacture provides a large portion of the *Staphylococcus aureus* genome and parts thereof (e.g., a *Staphylococcus aureus* open reading frame (ORF)) in a form which allows a skilled artisan to examine the manufacture using means not directly applicable to examining the *Staphylococcus aureus* genome or a subset thereof as it exists in nature or in purified form.

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories, such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. Likewise, it will be clear to those of skill how additional computer readable media that may be developed also can be used to create analogous manufactures having recorded thereon a nucleotide sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data-processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. Thus, by providing in computer readable form the nucleotide sequences of SEQ ID NOS:1-5,191, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably at least 99% and most preferably at least 99.9% identical to a sequence of SEQ ID NOS:1-5,191 the present invention enables the skilled artisan routinely to access the provided sequence information for a wide variety of purposes.

The examples which follow demonstrate how software which implements the BLAST (Altschul *et al.*, *J. Mol. Biol.* **215**:403-410 (1990)) and BLAZE (Brutlag *et al.*, *Comp. Chem.* **17**:203-207 (1993)) search algorithms on a Sybase system was used to identify open reading frames (ORFs) within the *Staphylococcus aureus* genome which contain homology to ORFs or proteins from both *Staphylococcus aureus* and from other organisms. Among the ORFs discussed

herein are protein encoding fragments of the *Staphylococcus aureus* genome useful in producing commercially important proteins, such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify, among other things, commercially important fragments of the *Staphylococcus aureus* genome.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention.

As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means.

As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the present genomic sequences which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBI). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymic active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the *Staphylococcus aureus* genomic sequences possessing varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments of the *Staphylococcus aureus* genome. In the present examples, implementing software which implement the BLAST and BLAZE algorithms, described in Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990), was used to identify open reading frames within the *Staphylococcus aureus* genome. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention. Of course, suitable proprietary systems that may be known to those of skill also may be employed in this regard.

Figure 1 provides a block diagram of a computer system illustrative of embodiments of this aspect of present invention. The computer system 102 includes a processor 106 connected to a bus 104. Also connected to the bus 104 are a main memory 108 (preferably implemented as random access memory, RAM) and a variety of secondary storage devices 110, such as a hard drive 112 and a removable medium storage device 114. The removable medium storage device 114 may represent, for example, a floppy disk drive, a CD-ROM drive, a magnetic tape drive, *etc.* A removable storage medium 116 (such as a floppy disk, a compact disk, a magnetic tape, *etc.*) containing control logic and/or data recorded therein may be inserted into the removable medium storage device 114. The computer system 102 includes appropriate software for reading the control logic and/or the data from the removable medium storage device 114, once it is inserted into the removable medium storage device 114.

A nucleotide sequence of the present invention may be stored in a well known manner in the main memory 108, any of the secondary storage devices 110, and/or a removable storage medium 116. During execution, software for

accessing and processing the genomic sequence (such as search tools, comparing tools, etc.) reside in main memory 108, in accordance with the requirements and operating parameters of the operating system, the hardware system and the software program or programs.

5 BIOCHEMICAL EMBODIMENTS

Other embodiments of the present invention are directed to fragments of the *Staphylococcus aureus* genome, preferably to isolated fragments. The fragments of the *Staphylococcus aureus* genome of the present invention include, but are not limited to fragments which encode peptides, hereinafter open reading frames (ORFs), fragments which
10 modulate the expression of an operably linked ORF, hereinafter expression modulating fragments (EMFs) and fragments which can be used to diagnose the presence of *Staphylococcus aureus* in a sample, hereinafter diagnostic fragments (DFs).

As used herein, an "isolated nucleic acid molecule" or an "isolated fragment of the *Staphylococcus aureus* genome" refers to a nucleic acid molecule possessing a specific nucleotide sequence which has been subjected to purification means to reduce, from the composition, the number of compounds which are normally associated with the composition.
15 Particularly, the term refers to the nucleic acid molecules having the sequences set out in SEQ ID NOS:1-5,191, to representative fragments thereof as described above, to polynucleotides at least 95%, preferably at least 99% and especially preferably at least 99.9% identical in sequence thereto, also as set out above.

A variety of purification means can be used to generate the isolated fragments of the present invention. These
20 include, but are not limited to methods which separate constituents of a solution based on charge, solubility, or size.

In one embodiment, *Staphylococcus aureus* DNA can be mechanically sheared to produce fragments of 15-20 kb in length. These fragments can then be used to generate an *Staphylococcus aureus* library by inserting them into lambda clones as described in the Examples below. Primers flanking, for example, an ORF, such as those enumerated in Tables 1-3 can then be generated using nucleotide sequence information provided in SEQ ID NOS: 1-5,191. Well
25 known and routine techniques of PCR cloning then can be used to isolate the ORF from the lambda DNA library of *Staphylococcus aureus* genomic DNA. Thus, given the availability of SEQ ID NOS:1-5,191, the information in Tables 1, 2 and 3, and the information that may be obtained readily by analysis of the sequences of SEQ ID NOS:1-5,191 using methods set out above, those of skill will be enabled by the present disclosure to isolate any ORF-containing or other nucleic acid fragment of the present invention.

The isolated nucleic acid molecules of the present invention include, but are not limited to single stranded and
30 double stranded DNA, and single stranded RNA.

As used herein, an "open reading frame," ORF, means a series of triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

Tables 1, 2 and 3 list ORFs in the *Staphylococcus aureus* genomic contigs of the present invention that were
35 identified as putative coding regions by the GeneMark software using organism-specific second-order Markov probability transition matrices. It will be appreciated that other criteria can be used, in accordance with well known analytical methods, such as those discussed herein, to generate more inclusive, more restrictive or more selective lists.

Table 1 sets out ORFs in the *Staphylococcus aureus* contigs of the present invention that are at least 80 amino acids long and over a continuous region of at least 50 bases which are 95% or more identical (by BLAST analysis) to
40 an *S. aureus* nucleotide sequence available through Genbank in November 1996.

Table 2 sets out ORFs in the *Staphylococcus aureus* contigs of the present invention that are not in Table 1 and match, with a BLASTP probability score of 0.01 or less, a polypeptide sequence available through Genbank by September 1996.

Table 3 sets out ORFs in the *Staphylococcus aureus* contigs of the present invention that do not match significantly,
45 by BLASTP analysis, a polypeptide sequence available through Genbank by September 1996.

In each table, the first and second columns identify the ORF by, respectively, contig number and ORF number within the contig; the third column indicates the reading frame, taking the first 5' nucleotide of the contig as the start of the +1 frame; the fourth column indicates the first nucleotide of the ORF, counting from the 5' end of the contig strand; and the fifth column indicates the length of each ORF in nucleotides.

In Tables 1 and 2, column six, lists the "Reference" for the closest matching sequence available through Genbank.
50 These reference numbers are the databases entry numbers commonly used by those of skill in the art, who will be familiar with their denominators. Descriptions of the nomenclature are available from the National Center for Biotechnology Information. Column seven in Tables 1 and 2 provides the "gene name" of the matching sequence; column eight provides the BLAST identity score from the comparison of the ORF and the homologous gene; and column nine
55 indicates the length in nucleotides of the highest scoring segment pair identified by the BLAST identity analysis.

In Table 3, the last column, column six, indicates the length of each ORF in amino acid residues.

The concepts of percent identity and percent similarity of two polypeptide sequences is well understood in the art. For example, two polypeptides 10 amino acids in length which differ at three amino acid positions (e.g., at positions

1, 3 and 5) are said to have a percent identity of 70%. However, the same two polypeptides would be deemed to have a percent similarity of 80% if, for example at position 5, the amino acids moieties, although not identical, were "similar" (*i.e.*, possessed similar biochemical characteristics). Many programs for analysis of nucleotide or amino acid sequence similarity, such as fasta and BLAST specifically list per cent identity of a matching region as an output parameter. Thus, for instance, Tables 1 and 2 herein enumerate the per cent identity" of the highest scoring segment pair" in each ORF and its listed relative. Further details concerning the algorithms and criteria used for homology searches are provided below and are described in the pertinent literature highlighted by the citations provided below.

It will be appreciated that other criteria can be used to generate more inclusive and more exclusive listings of the types set out in the tables. As those of skill will appreciate, narrow and broad searches both are useful. Thus, a skilled artisan can readily identify ORFs in contigs of the *Staphylococcus aureus* genome other than those listed in Tables 1-3, such as ORFs which are overlapping or encoded by the opposite strand of an identified ORF in addition to those ascertainable using the computer-based systems of the present invention.

As used herein, an "expression modulating fragment," EMF, means a series of nucleotide molecules which modulates the expression of an operably linked ORF or EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are fragments which induce the expression or an operably linked ORF in response to a specific regulatory factor or physiological event.

EMF sequences can be identified within the contigs of the *Staphylococcus aureus* genome by their proximity to the ORFs provided in Tables 1-3. An intergenic segment, or a fragment of the intergenic segment, from about 10 to 200 nucleotides in length, taken from any one of the ORFs of Tables 1-3 will modulate the expression of an operably linked ORF in a fashion similar to that found with the naturally linked ORF sequence. As used herein, an "intergenic segment" refers to fragments of the *Staphylococcus aureus* genome which are between two ORF(s) herein described. EMFs also can be identified using known EMFs as a target sequence or target motif in the computer-based systems of the present invention. Further, the two methods can be combined and used together.

The presence and activity of an EMF can be confirmed using an EMF trap vector. An EMF trap vector contains a cloning site linked to a marker sequence. A marker sequence encodes an identifiable phenotype, such as antibiotic resistance or a complementing nutrition auxotrophic factor, which can be identified or assayed when the EMF trap vector is placed within an appropriate host under appropriate conditions. As described above, a EMF will modulate the expression of an operably linked marker sequence. A more detailed discussion of various marker sequences is provided below.

A sequence which is suspected as being an EMF is cloned in all three reading frames in one or more restriction sites upstream from the marker sequence in the EMF trap vector. The vector is then transformed into an appropriate host using known procedures and the phenotype of the transformed host is examined under appropriate conditions. As described above, an EMF will modulate the expression of an operably linked marker sequence.

As used herein, a "diagnostic fragment," DF, means a series of nucleotide molecules which selectively hybridize to *Staphylococcus aureus* sequences. DFs can be readily identified by identifying unique sequences within contigs of the *Staphylococcus aureus* genome, such as by using well-known computer analysis software, and by generating and testing probes or amplification primers consisting of the DF sequence in an appropriate diagnostic format which determines amplification or hybridization selectivity.

The sequences falling within the scope of the present invention are not limited to the specific sequences herein described, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequences provided in SEQ ID NOS:1-5,191, a representative fragment thereof, or a nucleotide sequence at least 95%, preferably 99% and most preferably 99.9% identical to SEQ ID NOS:1-5,191, with a sequence from another isolate of the same species.

Furthermore, to accomodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the nucleic acid sequences mentioned above. In other words, in the coding region of an ORF, substitution of one codon for another which encodes the same amino acid is expressly contemplated.

Any specific sequence disclosed herein can be readily screened for errors by resequencing a particular fragment, such as an ORF, in both directions (*i.e.*, sequence both strands). Alternatively, error screening can be performed by sequencing corresponding polynucleotides of *Staphylococcus aureus* origin isolated by using part or all of the fragments in question as a probe or primer.

Each of the ORFs of the *Staphylococcus aureus* genome disclosed in Tables 1, 2 and 3, and the EMFs found 5' to the ORFs, can be used as polynucleotide reagents in numerous ways. For example, the sequences can be used as diagnostic probes or diagnostic amplification primers to detect the presence of a specific microbe in a sample, particular *Staphylococcus aureus*. Especially preferred in this regard are ORF such as those of Table 3, which do not match previously characterized sequences from other organisms and thus are most likely to be highly selective for *Staphylococcus aureus*. Also particularly preferred are ORFs that can be used to distinguish between strains of *Sta-*

phylococcus aureus, particularly those that distinguish medically important strain, such as drug-resistant strains.

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Information from the sequences of the present invention can be used to design antisense and triple helix-forming oligonucleotides. Polynucleotides suitable for use in these methods are usually 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription, for triple-helix formation, or to the mRNA itself, for antisense inhibition. Both techniques have been demonstrated to be effective in model systems, and the requisite techniques are well known and involve routine procedures. Triple helix techniques are discussed in, for example, Lee *et al.*, *Nucl. Acids Res.* **6**: 3073 (1979); Cooney *et al.*, *Science* **241**: 456 (1988); and Dervan *et al.*, *Science* **251**: 1360 (1991). Antisense techniques in general are discussed in, for instance, Okano, *J. Neurochem.* **56**: 560 (1991) and OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, FL (1988).

The present invention further provides recombinant constructs comprising one or more fragments of the *Staphylococcus aureus* genomic fragments and contigs of the present invention. Certain preferred recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a fragment of the *Staphylococcus aureus* genome has been inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. For vectors comprising the EMFs of the present invention, the vector may further comprise a marker sequence or heterologous ORF operably linked to the EMF.

Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Useful bacterial vectors include phagescript, PsiX174, pBluescript SK and KS (+ and -), pNH8a, pNH16a, pNH18a, pNH46a (available from Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (available from Pharmacia). Useful eukaryotic vectors include pWLneo, pSV2cat, pOG44, pXT1, pSG (available from Stratagene) pSVK3, pBPV, pMSG, pSVL (available from Pharmacia).

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

The present invention further provides host cells containing any one of the isolated fragments of the *Staphylococcus aureus* genomic fragments and contigs of the present invention, wherein the fragment has been introduced into the host cell using known methods. The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or a procaryotic cell, such as a bacterial cell.

A polynucleotide of the present invention, such as a recombinant construct comprising an ORF of the present invention, may be introduced into the host by a variety of well established techniques that are standard in the art, such as calcium phosphate transfection, DEAE, dextran mediated transfection and electroporation, which are described in, for instance, Davis, L. *et al.*, BASIC METHODS IN MOLECULAR BIOLOGY (1986).

A host cell containing one of the fragments of the *Staphylococcus aureus* genomic fragments and contigs of the present invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the Genetic Code, encode an identical polypeptide sequence.

Preferred nucleic acid fragments of the present invention are the ORFs depicted in Tables 2 and 3 which encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. This is particularly useful in producing small peptides and fragments of larger polypeptides. Such short fragments as may be obtained most readily by synthesis are useful, for example, in generating antibodies against the native polypeptide, as discussed further below.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily employ well-known methods for isolating polypeptides and proteins to isolate and purify polypeptides or proteins of the present invention produced naturally by a bacterial strain, or by other methods. Methods for isolation and purification that can be employed in this regard include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immu-

no-affinity chromatography.

The polypeptides and proteins of the present invention also can be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. Those skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, CV-1 cell, COS cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level.

"Recombinant," as used herein, means that a polypeptide or protein is derived from recombinant (e.g., microbial or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern different from that expressed in mammalian cells.

"Nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the polypeptides and proteins provided by this invention are assembled from fragments of the *Staphylococcus aureus* genome and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

"Recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. The expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic regulatory elements necessary for gene expression in the host, including elements required to initiate and maintain transcription at a level sufficient for suitable expression of the desired polypeptide, including, for example, promoters and, where necessary, an enhancers and a polyadenylation signal; (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate signals to initiate translation at the beginning of the desired coding region and terminate translation at its end. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

"Recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extra chromosomally. The cells can be prokaryotic or eukaryotic. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference in its entirety.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), alpha-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and, when desirable, provide amplification within the host.

Suitable prokaryotic hosts for transformation include strains of *Staphylococcus aureus*, *E. coli*, *B. subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. Others

may, also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (available from Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (available from Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter, where it is inducible, is derepressed or induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period to provide for expression of the induced gene product. Thereafter cells are typically harvested, generally by centrifugation, disrupted to release expressed protein, generally by physical or chemical means, and the resulting crude extract is retained for further purification.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described in Gluzman, *Cell* 23: 175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines.

Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

Recombinant polypeptides and proteins produced in bacterial culture is usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

An additional aspect of the invention includes *Staphylococcus aureus* polypeptides which are useful as immunodiagnostic antigens and/or immunoprotective vaccines, collectively "immunologically useful polypeptides". Such immunologically useful polypeptides may be selected from the ORFs disclosed herein based on techniques well known in the art and described elsewhere herein. The inventors have used the following criteria to select several immunologically useful polypeptides:

As is known in the art, an amino terminal type I signal sequence directs a nascent protein across the plasma and outer membranes to the exterior of the bacterial cell. Such outer membrane polypeptides are expected to be immunologically useful. According to Izard, J. W. et al., *Mol. Microbiol.* 13, 765-773; (1994), polypeptides containing type I signal sequences contain the following physical attributes: The length of the type I signal sequence is approximately 15 to 25 primarily hydrophobic amino acid residues with a net positive charge in the extreme amino terminus; the central region of the signal sequence must adopt an alpha-helical conformation in a hydrophobic environment; and the region surrounding the actual site of cleavage is ideally six residues long, with small side-chain amino acids in the -1 and -3 positions.

Also known in the art is the type IV signal sequence which is an example of the several types of functional signal sequences which exist in addition to the type I signal sequence detailed above. Although functionally related, the type IV signal sequence possesses a unique set of biochemical and physical attributes (Strom, M. S. and Lory, S., *J. Bacteriol.* 174, 7345-7351; 1992). These are typically six to eight amino acids with a net basic charge followed by an additional sixteen to thirty primarily hydrophobic residues. The cleavage site of a type IV signal sequence is typically after the initial six to eight amino acids at the extreme amino terminus. In addition, all type IV signal sequences contain a phenylalanine residue at the +1 site relative to the cleavage site.

Studies of the cleavage sites of twenty-six bacterial lipoprotein precursors has allowed the definition of a consensus amino acid sequence for lipoprotein cleavage. Nearly three-fourths of the bacterial lipoprotein precursors examined contained the sequence L-(A,S)-(G,A)-C at positions -3 to +1, relative to the point of cleavage (Hayashi, S. and Wu, H. C. *Lipoproteins in bacteria*. *J Bioenerg. Biomembr.* 22, 451-471; 1990).

It well known that most anchored proteins found on the surface of gram-positive bacteria possess a highly conserved carboxy terminal sequence. More than fifty such proteins from organisms such as *S. pyogenes*, *S. mutans*, *E. faecalis*, *S. pneumoniae*, and others, have been identified based on their extracellular location and carboxy terminal amino acid sequence (Fischetti, V. A. *Gram-positive commensal bacteria deliver antigens to elicit mucosal and systemic immunity*. *ASM News* 62, 405410; 1996). The conserved region is comprised of six charged amino acids at the extreme carboxy terminus coupled to 15-20 hydrophobic amino acids presumed to function as a transmembrane domain. Immediately adjacent to the transmembrane domain is a six amino acid sequence conserved in nearly all proteins ex-

aminated. The amino acid sequence of this region is L-P-X-T-G-X, where X is any amino acid.

Amino acid sequence similarities to proteins of known function by BLAST enables the assignment of putative functions to novel amino acid sequences and allows for the selection of proteins thought to function outside the cell wall. Such proteins are well known in the art and include "lipoprotein", "periplasmic", or "antigen".

5 An algorithm for selecting antigenic and immunogenic *Staphylococcus aureus* polypeptides including the foregoing criteria was developed by the present inventors. Use of the algorithm by the inventors to select immunologically useful *Staphylococcus aureus* polypeptides resulted in the selection of several ORFs which are predicted to be outermembrane-associated proteins. These proteins are identified in Table 4, below, and shown in the Sequence Listing as SEQ ID NOS:5,192 to 5,255. Thus the amino acid sequence of each of several antigenic *Staphylococcus aureus* polypeptides listed in Table 4 can be determined, for example, by locating the amino acid sequence of the ORF in the Sequence Listing. Likewise the polynucleotide sequence encoding each ORF can be found by locating the corresponding polynucleotide SEQ ID in Tables 1, 2, or 3, and finding the corresponding nucleotide sequence in the sequence listing.

As will be appreciated by those of ordinary skill in the art, although a polypeptide representing an entire ORF may be the closest approximation to a protein found *in vivo*, it is not always technically practical to express a complete ORF *in vitro*. It may be very challenging to express and purify a highly hydrophobic protein by common laboratory methods. As a result, the immunologically useful polypeptides described herein as SEQ ID NOS:5,192-5,255 may have been modified slightly to simplify the production of recombinant protein, and are the preferred embodiments. In general, nucleotide sequences which encode highly hydrophobic domains, such as those found at the amino terminal signal sequence, are excluded for enhanced *in vitro* expression of the polypeptides. Furthermore, any highly hydrophobic amino acid sequences occurring at the carboxy terminus are also excluded. Such truncated polypeptides include for example the mature forms of the polypeptides expected to exist in nature.

Those of ordinary skill in the art can identify soluble portions the polypeptide identified in Table 4, and in the case of truncated polypeptides sequences shown as SEQ ID NOS:5,192-5,255, may obtain the complete predicted amino acid sequence of each polypeptide by translating the corresponding polynucleotides sequences of the corresponding ORF listed in Tables 1,2 and 3 and found in the sequence listing.

Accordingly, polypeptides comprising the complete amino acid of an immunologically useful polypeptide selected from the group of polypeptides encoded by the ORFs identified in Table 4, or an amino acid sequence at least 95% identical thereto, preferably at least 97% identical thereto, and most preferably at least 99% identical thereto form an embodiment of the invention; in addition polypeptides comprising an amino acid sequence selected from the group of amino acid sequences shown in the sequence listing as SEQ ID NOS:5,191-5,255, or an amino acid sequence at least 95% identical thereto, preferably at least 97% identical thereto and most preferably at least 99% identical thereto, form an embodiment of the invention. Polynucleotides encoding the foregoing polypeptides also form part of the present invention.

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention, particularly those epitope-bearing portions (antigenic regions) identified in Table 4. The epitope-bearing portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See, for instance, Sutcliffe, J. G., Shinnick, T. M., Green, N. and Learner, R. A. (1983) "Antibodies that react with predetermined sites on proteins", Science, 219:660-666. Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson et al., Cell 37:767-778 (1984) at 777.

Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. Non-limiting examples of antigenic polypeptides or peptides that can be used to generate *S. aureus* specific antibodies include: a polypeptide comprising peptides shown in Table 4 below. These polypeptide fragments have been determined to bear antigenic epitopes of indicated *S. aureus* proteins by the analysis of the Jameson-Wolf antigenic index, a representative sample of which is shown in Figure 3.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means. See, e.g., Houghten, R. A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides:

specificity of antigen-antibody interaction at the level of individual amino acids. Proc. Natl. Acad. Sci. USA 82: 5131-5135; this "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten et al. (1986). Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art. See, for instance, Sutcliffe et al., supra; Wilson et al., supra;

5 Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).

Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. See, for instance, Geysen et al., supra. Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. et al. (1996) on Per-alkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a per-alkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Table 4 lists immunologically useful polypeptides identified by an algorithm which locates novel *Staphylococcus aureus* outermembrane proteins, as is described above. Also listed are epitopes or "antigenic regions" of each of the identified polypeptides. The antigenic regions, or epitopes, are delineated by two numbers x-y, where x is the number of the first amino acid in the open reading frame included within the epitope and y is the number of the last amino acid in the open reading frame included within the epitope. For example, the first epitope in ORF 168-6 is comprised of amino acids 36 to 45 of SEQ ID NO:5,192, as is described in Table 4. The inventors have identified several epitopes for each of the antigenic polypeptides identified in Table 4. Accordingly, forming part of the present invention are polypeptides comprising an amino acid sequence of one or more antigenic regions identified in Table 4. The invention further provides polynucleotides encoding such polypeptides.

The present invention further includes isolated polypeptides, proteins and nucleic acid molecules which are substantially equivalent to those herein described. As used herein, substantially equivalent can refer both to nucleic acid and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between reference and subject sequences. For purposes of the present invention, sequences having equivalent biological activity, and equivalent expression characteristics are considered substantially equivalent. For purposes of determining equivalence, truncation of the mature sequence should be disregarded.

The invention further provides methods of obtaining homologs from other strains of *Staphylococcus aureus*, of the fragments of the *Staphylococcus aureus* genome of the present invention and homologs of the proteins encoded by the ORFs of the present invention. As used herein, a sequence or protein of *Staphylococcus aureus* is defined as a homolog of a fragment of the *Staphylococcus aureus* fragments or contigs or a protein encoded by one of the ORFs of the present invention, if it shares significant homology to one of the fragments of the *Staphylococcus aureus* genome of the present invention or a protein encoded by one of the ORFs of the present invention. Specifically, by using the sequence disclosed herein as a probe or as primers, and techniques such as PCR cloning and colony/plaque hybridization, one skilled in the art can obtain homologs.

As used herein, two nucleic acid molecules or proteins are said to "share significant homology" if the two contain regions which possess greater than 85% sequence (amino acid or nucleic acid) homology. Preferred homologs in this regard are those with more than 90% homology. Especially preferred are those with 93% or more homology. Among especially preferred homologs those with 95% or more homology are particularly preferred. Very particularly preferred among these are those with 97% and even more particularly preferred among these are homologs with 99% or more homology. The most preferred homologs among these are those with 99.9% homology or more. It will be understood that, among measures of homology, identity is particularly preferred in this regard.

Region specific primers or probes derived from the nucleotide sequence provided in SEQ ID NOS:1-5,191 or from a nucleotide sequence at least 95%, particularly at least 99%, especially at least 99.5% identical to a sequence of SEQ ID NOS:1-5,191 can be used to prime DNA synthesis and PCR amplification, as well as to identify colonies containing cloned DNA encoding a homolog. Methods suitable to this aspect of the present invention are well known and have been described in great detail in many publications such as, for example, Innis et al., PCR PROTOCOLS, Academic Press, San Diego, CA (1990)).

When using primers derived from SEQ ID NOS:1-5,191 or from a nucleotide sequence having an aforementioned identity to a sequence of SEQ ID NOS:1-5,191, one skilled in the art will recognize that by employing high stringency conditions (e.g., annealing at 50-60°C in 6X SSPE and 50% formamide, and washing at 50-65°C in 0.5X SSPE) only sequences which are greater than 75% homologous to the primer will be amplified. By employing lower stringency

conditions (e.g., hybridizing at 35-37°C in 5X SSPC and 40-45% formamide, and washing at 42°C in 0.5X SSPC), sequences which are greater than 40-50% homologous to the primer will also be amplified.

When using DNA probes derived from SEQ ID NOS:1-5,191, or from a nucleotide sequence having an aforementioned identity to a sequence of SEQ ID NOS:1-5,191, for colony/plaque hybridization, one skilled in the art will recognize that by employing high stringency conditions (e.g., hybridizing at 50-65°C in 5X SSPC and 50% formamide, and washing at 50-65°C in 0.5X SSPC), sequences having regions which are greater than 90% homologous to the probe can be obtained, and that by employing lower stringency conditions (e.g., hybridizing at 35-37°C in 5X SSPC and 40-45% formamide, and washing at 42°C in 0.5X SSPC), sequences having regions which are greater than 35-45% homologous to the probe will be obtained.

Any organism can be used as the source for homologs of the present invention so long as the organism naturally expresses such a protein or contains genes encoding the same. The most preferred organism for isolating homologs are bacteria which are closely related to *Staphylococcus aureus*.

ILLUSTRATIVE USES OF COMPOSITIONS OF THE INVENTION

Each ORF provided in Tables 1 and 2 is identified with a function by homology to a known gene or polypeptide. As a result, one skilled in the art can use the polypeptides of the present invention for commercial, therapeutic and industrial purposes consistent with the type of putative identification of the polypeptide. Such identifications permit one skilled in the art to use the *Staphylococcus aureus* ORFs in a manner similar to the known type of sequences for which the identification is made; for example, to ferment a particular sugar source or to produce a particular metabolite. A variety of reviews illustrative of this aspect of the invention are available, including the following reviews on the industrial use of enzymes, for example, BIOCHEMICAL ENGINEERING AND BIOTECHNOLOGY HANDBOOK, 2nd Ed., Macmillan Publications, Ltd. NY (1991) and BIOCATALYSTS IN ORGANIC SYNTHESIS, Tramper *et al.*, Eds., Elsevier Science Publishers, Amsterdam, The Netherlands (1985). A variety of exemplary uses that illustrate this and similar aspects of the present invention are discussed below.

1. Biosynthetic Enzymes

Open reading frames encoding proteins involved in mediating the catalytic reactions involved in intermediary and macromolecular metabolism, the biosynthesis of small molecules, cellular processes and other functions includes enzymes involved in the degradation of the intermediary products of metabolism, enzymes involved in central intermediary metabolism, enzymes involved in respiration, both aerobic and anaerobic, enzymes involved in fermentation, enzymes involved in ATP proton motor force conversion, enzymes involved in broad regulatory function, enzymes involved in amino acid synthesis, enzymes involved in nucleotide synthesis, enzymes involved in cofactor and vitamin synthesis, can be used for industrial biosynthesis.

The various metabolic pathways present in *Staphylococcus aureus* can be identified based on absolute nutritional requirements as well as by examining the various enzymes identified in Table 1-3 and SEQ ID NOS:1-5,191.

Of particular interest are polypeptides involved in the degradation of intermediary metabolites as well as non-macromolecular metabolism. Such enzymes include amylases, glucose oxidases, and catalase.

Proteolytic enzymes are another class of commercially important enzymes. Proteolytic enzymes find use in a number of industrial processes including the processing of flax and other vegetable fibers, in the extraction, clarification and depectinization of fruit juices, in the extraction of vegetables' oil and in the maceration of fruits and vegetables to give unicellular fruits. A detailed review of the proteolytic enzymes used in the food industry is provided in Rombouts *et al.*, *Symbiosis* 21: 79 (1986) and Voragen *et al.* in BIOCATALYSTS IN AGRICULTURAL BIOTECHNOLOGY, Whitaker *et al.*, Eds., *American Chemical Society Symposium Series* 389: 93 (1989).

The metabolism of sugars is an important aspect of the primary metabolism of *Staphylococcus aureus*. Enzymes involved in the degradation of sugars, such as, particularly, glucose, galactose, fructose and xylose, can be used in industrial fermentation. Some of the important sugar transforming enzymes, from a commercial viewpoint, include sugar isomerases such as glucose isomerase. Other metabolic enzymes have found commercial use such as glucose oxidases which produces ketogulonic acid (KGA). KGA is an intermediate in the commercial production of ascorbic acid using the Reichstein's procedure, as described in Krueger *et al.*, *Biotechnology* 6(A), Rhine *et al.*, Eds., Verlag Press, Weinheim, Germany (1984).

Glucose oxidase (GOD) is commercially available and has been used in purified form as well as in an immobilized form for the deoxygenation of beer. See, for instance, Hartmeir *et al.*, *Biotechnology Letters* 1: 21 (1979). The most important application of GOD is the industrial scale fermentation of gluconic acid. Market for gluconic acids which are used in the detergent, textile, leather, photographic, pharmaceutical, food, feed and concrete industry, as described, for example, in Bigelis *et al.*, beginning on page 357 in GENE MANIPULATIONS AND FUNGI; Benett *et al.*, Eds., Academic Press, New York (1985). In addition to industrial applications, GOD has found applications in medicine for

quantitative determination of glucose in body fluids recently in biotechnology for analyzing syrups from starch and cellulose hydrosylates. This application is described in Owusu *et al.*, *Biochem. et Biophysica. Acta.* 872: 83 (1986), for instance.

The main sweetener used in the world today is sugar which comes from sugar beets and sugar cane. In the field of industrial enzymes, the glucose isomerase process shows the largest expansion in the market today. Initially, soluble enzymes were used and later immobilized enzymes were developed (Krueger *et al.*, *Biotechnology, The Textbook of Industrial Microbiology*, Sinauer Associated Incorporated, Sunderland, Massachusetts (1990)). Today, the use of glucose- produced high fructose syrups is by far the largest industrial business using immobilized enzymes. A review of the industrial use of these enzymes is provided by Jorgensen, *Starch* 40:307 (1988).

Proteinases, such as alkaline serine proteinases, are used as detergent additives and thus represent one of the largest volumes of microbial enzymes used in the industrial sector. Because of their industrial importance, there is a large body of published and unpublished information regarding the use of these enzymes in industrial processes. (See Faultman *et al.*, *Acid Proteases Structure Function and Biology*, Tang, J., ed., Plenum Press, New York (1977) and Godfrey *et al.*, *Industrial Enzymes*, MacMillan Publishers, Surrey, UK (1983) and Hepner *et al.*, *Report Industrial Enzymes* by 1990, Hel Hepner & Associates, London (1986)).

Another class of commercially usable proteins of the present invention are the microbial lipases, described by, for instance, Macrae *et al.*, *Philosophical Transactions of the Chiral Society of London* 310:227 (1985) and Poserke, *Journal of the American Oil Chemist Society* 61:1758 (1984). A major use of lipases is in the fat and oil industry for the production of neutral glycerides using lipase catalyzed inter-esterification of readily available triglycerides. Application of lipases include the use as a detergent additive to facilitate the removal of fats from fabrics in the course of the washing procedures.

The use of enzymes, and in particular microbial enzymes, as catalyst for key steps in the synthesis of complex organic molecules is gaining popularity at a great rate. One area of great interest is the preparation of chiral intermediates. Preparation of chiral intermediates is of interest to a wide range of synthetic chemists particularly those scientists involved with the preparation of new pharmaceuticals, agrochemicals, fragrances and flavors. (See Davies *et al.*, *Recent Advances in the Generation of Chiral Intermediates Using Enzymes*, CRC Press, Boca Raton, Florida (1990)). The following reactions catalyzed by enzymes are of interest to organic chemists: hydrolysis of carboxylic acid esters, phosphate esters, amides and nitriles, esterification reactions, trans-esterification reactions, synthesis of amides, reduction of alkanones and oxoalkanates, oxidation of alcohols to carbonyl compounds, oxidation of sulfides to sulfoxides, and carbon bond forming reactions such as the aldol reaction.

When considering the use of an enzyme encoded by one of the ORFs of the present invention for biotransformation and organic synthesis it is sometimes necessary to consider the respective advantages and disadvantages of using a microorganism as opposed to an isolated enzyme. Pros and cons of using a whole cell system on the one hand or an isolated partially purified enzyme on the other hand, has been described in detail by Bud *et al.*, *Chemistry in Britain* (1987), p. 127.

Amino transferases, enzymes involved in the biosynthesis and metabolism of amino acids, are useful in the catalytic production of amino acids. The advantages of using microbial based enzyme systems is that the amino transferase enzymes catalyze the stereo- selective synthesis of only L-amino acids and generally possess uniformly high catalytic rates. A description of the use of amino transferases for amino acid production is provided by Roselle-David, *Methods of Enzymology* 136:479 (1987).

Another category of useful proteins encoded by the ORFs of the present invention include enzymes involved in nucleic acid synthesis, repair, and recombination. A variety of commercially important enzymes have previously been isolated from members of *Staphylococcus aureus*. These include Sau3A and Sau96I.

2. Generation of Antibodies

As described here, the proteins of the present invention, as well as homologs thereof, can be used in a variety of procedures and methods known in the art which are currently applied to other proteins. The proteins of the present invention can further be used to generate an antibody which selectively binds the protein. Such antibodies can be either monoclonal or polyclonal antibodies, as well fragments of these antibodies, and humanized forms.

The invention further provides antibodies which selectively bind to one of the proteins of the present invention and hybridomas which produce these antibodies. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing polyclonal and monoclonal antibodies as well as hybridomas capable of producing the desired antibody are well known in the art (Campbell, A. M., *MONOCLONAL ANTIBODY TECHNOLOGY: LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY*, Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth *et al.*, *J. Immunol. Methods* 35: 1-21 (1980), Kohler and Milstein, *Nature* 256: 495-497 (1975)), the trioma technique, the human B- cell hybridoma technique (Kozbor *et al.*, *Immunology Today*

4: 72 (1983), pgs. 77-96 of Cole *et al.*, in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. (1985)).

Any animal (mouse, rabbit, *etc.*) which is known to produce antibodies can be immunized with the pseudogene polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or interperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of the protein encoded by the ORF of the present invention used for immunization will vary based on the animal which is immunized, the antigenicity of the peptide and the site of injection.

The protein which is used as an immunogen may be modified or administered in an adjuvant in order to increase the protein's antigenicity. Methods of increasing the antigenicity of a protein are well known in the art and include, but are not limited to coupling the antigen with a heterologous protein (such as globulin or galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells.

Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al.*, *Exp. Cell Res.* 175: 109-124 (1988)).

Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, A. M., *Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1984)).

Techniques described for the production of single chain antibodies (U. S. Patent 4,946,778) can be adapted to produce single chain antibodies to proteins of the present invention.

For polyclonal antibodies, antibody containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures.

The present invention further provides the above-described antibodies in detectably labelled form. Antibodies can be detectably labelled through the use of radioisotopes, affinity labels (such as biotin, avidin, *etc.*), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, *etc.*) fluorescent labels (such as FITC or rhodamine, *etc.*), paramagnetic atoms, *etc.* Procedures for accomplishing such labelling are well-known in the art, for example see Sternberger *et al.*, *J. Histochem. Cytochem.* 18:315 (1970); Bayer, E. A. *et al.*, *Meth. Enzym.* 62:308 (1979); Engval, E. *et al.*, *Immunol.* 109:129 (1972); Goding, J. W. *J. Immunol. Meth.* 13:215 (1976)).

The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and in situ assays to identify cells or tissues in which a fragment of the *Staphylococcus aureus* genome is expressed.

The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D. M. *et al.*, "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W. D. *et al.*, *Meth. Enzym.* 34 Academic Press, N. Y. (1974)). The immobilized antibodies of the present invention can be used for *in vitro*, *in vivo*, and in situ assays as well as for immunoaffinity purification of the proteins of the present invention.

3. Diagnostic Assays and Kits

The present invention further provides methods to identify the expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using one of the DFs, antigens or antibodies of the present invention.

In detail, such methods comprise incubating a test sample with one or more of the antibodies, or one or more of the DFs, or one or more antigens of the present invention and assaying for binding of the DFs, antigens or antibodies to components within the test sample.

Conditions for incubating a DF, antigen or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the DF or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the DFs, antigens or antibodies of the present invention. Examples of such assays can be found in Chard, T., *An Introduction to Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. *et al.*, *Techniques in Immunocytochemistry*, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry*, PCT publication WO95/32291, and *Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1985), all of which are hereby incorporated herein by reference.

The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based

on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the Dfs, antigens or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound DF, antigen or antibody.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody, antigen or DF.

Types of detection reagents include labelled nucleic acid probes, labelled secondary antibodies, or in the alternative, if the primary antibody is labelled, the enzymatic, or antibody binding reagents which are capable of reacting with the labelled antibody. One skilled in the art will readily recognize that the disclosed Dfs, antigens and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4. Screening Assay for Binding Agents

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents which bind to a protein encoded by one of the ORFs of the present invention or to one of the fragments and the *Staphylococcus aureus* fragment and contigs herein described.

In general, such methods comprise steps of:

- (a) contacting an agent with an isolated protein encoded by one of the ORFs of the present invention, or an isolated fragment of the *Staphylococcus aureus* genome; and
- (b) determining whether the agent binds to said protein or said fragment.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention.

Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides, for example see Hurby *et al.*, Application of Synthetic Peptides: Antisense Peptides, In Synthetic Peptides, A User's Guide, W. H. Freeman, NY (1992), pp. 289-307, and Kaspczak *et al.*, Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control.

One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods usually contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee *et al.*, Nucl. Acids Res. 6:3073 (1979); Cooney *et al.*, Science 241:456 (1988); and Dervan *et al.*, Science 251: 1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated

to be effective in model systems. Information contained in the sequences of the present invention can be used to design antisense and triple helix-forming oligonucleotides, and other DNA binding agents.

5. Pharmaceutical Compositions and Vaccines

The present invention further provides pharmaceutical agents which can be used to modulate the growth or pathogenicity of *Staphylococcus aureus*, or another related organism, *in vivo* or *in vitro*. As used herein, a "pharmaceutical agent" is defined as a composition of matter which can be formulated using known techniques to provide a pharmaceutical compositions. As used herein, the "pharmaceutical agents of the present invention" refers the pharmaceutical agents which are derived from the proteins encoded by the ORFs of the present invention or are agents which are identified using the herein described assays.

As used herein, a pharmaceutical agent is said to "modulate the growth or pathogenicity of *Staphylococcus aureus* or a related organism, *in vivo* or *in vitro*," when the agent reduces the rate of growth, rate of division, or viability of the organism in question. The pharmaceutical agents of the present invention can modulate the growth or pathogenicity of an organism in many fashions, although an understanding of the underlying mechanism of action is not needed to practice the use of the pharmaceutical agents of the present invention. Some agents will modulate the growth or pathogenicity by binding to an important protein thus blocking the biological activity of the protein, while other agents may bind to a component of the outer surface of the organism blocking attachment or rendering the organism more prone to act the bodies nature immune system. Alternatively, the agent may comprise a protein encoded by one of the ORFs of the present invention and serve as a vaccine. The development and use of vaccines derived from membrane associated polypeptides are well known in the art. The inventors have identified particularly preferred immunogenic *Staphylococcus aureus* polypeptides for use as vaccines. Such immunogenic polypeptides are described above and summarized in Table 4, below.

As used herein, a "related organism" is a broad term which refers to any organism whose growth or pathogenicity can be modulated by one of the pharmaceutical agents of the present invention. In general, such an organism will contain a homolog of the protein which is the target of the pharmaceutical agent or the protein used as a vaccine. As such, related organisms do not need to be bacterial but may be fungal or viral pathogens.

The pharmaceutical agents and compositions of the present invention may be administered in a convenient manner, such as by the oral, topical, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes. The pharmaceutical compositions are administered in an amount which is effective for treating and/or prophylaxis of the specific indication. In general, they are administered in an amount of at least about 1 mg/kg body weight and in most cases they will be administered in an amount not in excess of about 1 g/kg body weight per day. In most cases, the dosage is from about 0.1 mg/kg to about 10 g/kg body weight daily, taking into account the routes of administration, symptoms, etc.

The agents of the present invention can be used in native form or can be modified to form a chemical derivative. As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed in, among other sources, REMINGTON'S PHARMACEUTICAL SCIENCES (1980) cited elsewhere herein.

For example, such moieties may change an immunological character of the functional derivative, such as affinity for a given antibody. Such changes in immunomodulation activity are measured by the appropriate assay, such as a competitive type immunoassay. Modifications of such protein properties as redox or thermal stability, biological half-life, hydrophobicity, susceptibility to proteolytic degradation or the tendency to aggregate with carriers or into multimers also may be effected in this way and can be assayed by methods well known to the skilled artisan.

The therapeutic effects of the agents of the present invention may be obtained by providing the agent to a patient by any suitable means (e.g., inhalation, intravenously, intramuscularly, subcutaneously, enterally, or parenterally). It is preferred to administer the agent of the present invention so as to achieve an effective concentration within the blood or tissue in which the growth of the organism is to be controlled. To achieve an effective blood concentration, the preferred method is to administer the agent by injection. The administration may be by continuous infusion, or by single or multiple injections.

In providing a patient with one of the agents of the present invention, the dosage of the administered agent will vary depending upon such factors as the patient's age, weight, height, sex, general medical condition, previous medical history, etc. In general, it is desirable to provide the recipient with a dosage of agent which is in the range of from about 1 pg/kg to 10 mg/kg (body weight of patient), although a lower or higher dosage may be administered. The therapeutically effective dose can be lowered by using combinations of the agents of the present invention or another agent.

As used herein, two or more compounds or agents are said to be administered "in combination" with each other when either (1) the physiological effects of each compound, or (2) the serum concentrations of each compound can

be measured at the same time. The composition of the present invention can be administered concurrently with, prior to, or following the administration of the other agent.

The agents of the present invention are intended to be provided to recipient subjects in an amount sufficient to decrease the rate of growth (as defined above) of the target organism.

The administration of the agent(s) of the invention may be for either a "prophylactic" or "therapeutic" purpose. When provided prophylactically, the agent(s) are provided in advance of any symptoms indicative of the organisms growth. The prophylactic administration of the agent(s) serves to prevent, attenuate, or decrease the rate of onset of any subsequent infection. When provided therapeutically, the agent(s) are provided at (or shortly after) the onset of an indication of infection. The therapeutic administration of the compound(s) serves to attenuate the pathological symptoms of the infection and to increase the rate of recovery.

The agents of the present invention are administered to a subject, such as a mammal, or a patient, in a pharmaceutically acceptable form and in a therapeutically effective concentration. A composition is said to be "pharmacologically acceptable" if its administration can be tolerated by a recipient patient. Such an agent is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient.

The agents of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby these materials, or their functional derivatives, are combined in admixture with a pharmaceutically acceptable carrier vehicle. Suitable vehicles and their formulation, inclusive of other human proteins, e.g., human serum albumin, are described, for example, in REMINGTON'S PHARMACEUTICAL SCIENCES, 16th Ed., Osol, A., Ed., Mack Publishing, Easton PA (1980). In order to form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of one or more of the agents of the present invention, together with a suitable amount of carrier vehicle.

Additional pharmaceutical methods may be employed to control the duration of action. Control release preparations may be achieved through the use of polymers to complex or absorb one or more of the agents of the present invention. The controlled delivery may be effectuated by a variety of well known techniques, including formulation with macromolecules such as, for example, polyesters, polyamino acids, polyvinyl, pyrrolidone, ethylenevinylacetate, methylcellulose, carboxymethylcellulose, or protamine, sulfate, adjusting the concentration of the macromolecules and the agent in the formulation, and by appropriate use of methods of incorporation, which can be manipulated to effectuate a desired time course of release. Another possible method to control the duration of action by controlled release preparations is to incorporate agents of the present invention into particles of a polymeric material such as polyesters, polyamino acids, hydrogels, poly(lactic acid) or ethylene vinylacetate copolymers. Alternatively, instead of incorporating these agents into polymeric particles, it is possible to entrap these materials in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization with, for example, hydroxymethylcellulose or gelatine-microcapsules and poly(methylmethacrylate) microcapsules, respectively, or in colloidal drug delivery systems, for example, liposomes, albumin microspheres, microemulsions, nanoparticles, and nanocapsules or in macroemulsions. Such techniques are disclosed in REMINGTON'S PHARMACEUTICAL SCIENCES (1980).

The invention further provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

In addition, the agents of the present invention may be employed in conjunction with other therapeutic compounds.

6. Shot-Gun Approach to Megabase DNA Sequencing

The present invention further demonstrates that a large sequence can be sequenced using a random shotgun approach. This procedure, described in detail in the examples that follow, has eliminated the up front cost of isolating and ordering overlapping or contiguous subclones prior to the start of the sequencing protocols.

Certain aspects of the present invention are described in greater detail in the examples that follow. The examples are provided by way of illustration. Other aspects and embodiments of the present invention are contemplated by the inventors, as will be clear to those of skill in the art from reading the present disclosure.

ILLUSTRATIVE EXAMPLES**LIBRARIES AND SEQUENCING****1. Shotgun Sequencing Probability Analysis**

The overall strategy for a shotgun approach to whole genome sequencing follows from the Lander and Waterman (Landerman and Waterman, *Genomics* 2: 231 (1988)) application of the equation for the Poisson distribution. According to this treatment, the probability, P_0 , that any given base in a sequence of size L , in nucleotides, is not sequenced after a certain amount, n , in nucleotides, of random sequence has been determined can be calculated by the equation $P_0 = e^{-m}$, where m is L/n , the fold coverage.* For instance, for a genome of 2.8 Mb, $m=1$ when 2.8 Mb of sequence has been randomly generated (1X coverage). At that point, $P_0 = e^{-1} = 0.37$. The probability that any given base has not been sequenced is the same as the probability that any region of the whole sequence L has not been determined and, therefore, is equivalent to the fraction of the whole sequence that has yet to be determined. Thus, at one-fold coverage, approximately 37% of a polynucleotide of size L , in nucleotides has not been sequenced. When 14 Mb of sequence has been generated, coverage is 5X for a 2.8 Mb and the unsequenced fraction drops to .0067 or 0.67%. 5X coverage of a 2.8 Mb sequence can be attained by sequencing approximately 17,000 random clones from both insert ends with an average sequence read length of 410 bp.

Similarly, the total gap length, G , is determined by the equation $G = Le^{-m}$, and the average gap size, g , follows the equation, $g = L/n$. Thus, 5X coverage leaves about 240 gaps averaging about 82 bp in size in a sequence of a polynucleotide 2.8 Mb long.

The treatment above is essentially that of Lander and Waterman, *Genomics* 2: 231 (1988).

2. Random Library Construction

In order to approximate the random model described above during actual sequencing, a nearly ideal library of cloned genomic fragments is required. The following library construction procedure was developed to achieve this end.

Staphylococcus aureus DNA was prepared by phenol extraction. A mixture containing 600 μ g DNA in 3.3 ml of 300 mM sodium acetate, 10 mM Tris-HCl, 1 mM Na-EDTA, 30% glycerol was sonicated for 1 min. at 0°C in a Branson Model 450 Sonicator at the lowest energy setting using a 3 mm probe. The sonicated DNA was ethanol precipitated and redissolved in 500 μ l TE buffer.

To create blunt-ends, a 100 μ l aliquot of the resuspended DNA was digested with 5 units of BAL31 nuclease (New England BioLabs) for 10 min at 30°C in 200 μ l BAL31 buffer. The digested DNA was phenol-extracted, ethanol-precipitated, redissolved in 100 μ l TE buffer, and then size-fractionated by electrophoresis through a 1.0% low melting temperature agarose gel. The section containing DNA fragments 1.6-2.0 kb in size was excised from the gel, and the LGT agarose was melted and the resulting solution was extracted with phenol to separate the agarose from the DNA. DNA was ethanol precipitated and redissolved in 20 μ l of TE buffer for ligation to vector.

A two-step ligation procedure was used to produce a plasmid library with 97% inserts, of which >99% were single inserts. The first ligation mixture (50 μ l) contained 2 μ g of DNA fragments, 2 μ g pUC18 DNA (Pharmacia) cut with SmaI and dephosphorylated with bacterial alkaline phosphatase, and 10 units of T4 ligase (GIBCO/BRL) and was incubated at 14°C for 4 hr. The ligation mixture then was phenol extracted and ethanol precipitated, and the precipitated DNA was dissolved in 20 μ l TE buffer and electrophoresed on a 1.0% low melting agarose gel. Discrete bands in a ladder were visualized by ethidium bromide-staining and UV illumination and identified by size as insert (i), vector (v), v+i, v+2i, v+3i, etc. The portion of the gel containing v+i DNA was excised and the v+i DNA was recovered and resuspended into 20 μ l TE. The v+i DNA then was blunt-ended by T4 polymerase treatment for 5 min. at 37°C in a reaction mixture (50 μ l) containing the v+i linears, 500 μ M each of the 4 dNTPs, and 9 units of T4 polymerase (New England BioLabs), under recommended buffer conditions. After phenol extraction and ethanol precipitation the repaired v+i linears were dissolved in 20 μ l TE. The final ligation to produce circles was carried out in a 50 μ l reaction containing 5 μ l of v+i linears and 5 units of T4 ligase at 14°C overnight. After 10 min. at 70°C the following day, the reaction mixture was stored at -20°C.

This two-stage procedure resulted in a molecularly random collection of single-insert plasmid recombinants with minimal contamination from double-insert chimeras (<1%) or free vector (<3%).

Since deviation from randomness can arise from propagation the DNA in the host, *E. coli* host cells deficient in all recombination and restriction functions (A. Greener, *Strategies* 3 (1):5 (1990)) were used to prevent rearrangements, deletions, and loss of clones by restriction. Furthermore, transformed cells were plated directly on antibiotic diffusion plates to avoid the usual broth recovery phase which allows multiplication and selection of the most rapidly growing cells.

Plating was carried out as follows. A 100 μ l aliquot of Epicurian Coli SURE II Supercompetent Cells (Stratagene 200152) was thawed on ice and transferred to a chilled Falcon 2059 tube on ice. A 1.7 μ l aliquot of 1.42 M beta-

mercaptoethanol was added to the aliquot of cells to a final concentration of 25 mM. Cells were incubated on ice for 10 min. A 1 µl aliquot of the final ligation was added to the cells and incubated on ice for 30 min. The cells were heat pulsed for 30 sec. at 42° C and placed back on ice for 2 min. The outgrowth period in liquid culture was eliminated from this protocol in order to minimize the preferential growth of any given transformed cell. Instead the transformation mixture was plated directly on a nutrient rich SOB plate containing a 5 ml bottom layer of SOB agar (5% SOB agar: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 1.5% Difco Agar per liter of media). The 5 ml bottom layer is supplemented with 0.4 ml of 50 mg/ml ampicillin per 100 ml SOB agar. The 15 ml top layer of SOB agar is supplemented with 1 ml X-Gal (2%), 1 ml MgCl₂ (1 M), and 1 ml MgSO₄/100 ml SOB agar. The 15 ml top layer was poured just prior to plating. Our titer was approximately 100 colonies/10 µl aliquot of transformation.

All colonies were picked for template preparation regardless of size. Thus, only clones lost due to "poison" DNA or deleterious gene products would be deleted from the library, resulting in a slight increase in gap number over that expected.

3. Random DNA Sequencing

High quality double stranded DNA plasmid templates were prepared using an alkaline lysis method developed in collaboration with 5Prime → 3Prime Inc. (Boulder, CO). Plasmid preparation was performed in a 96-well format for all stages of DNA preparation from bacterial growth through final DNA purification. Average template concentration was determined by running 25% of the samples on an agarose gel. DNA concentrations were not adjusted.

Templates were also prepared from a *Staphylococcus aureus* lambda genomic library. An unamplified library was constructed in Lambda DASH II vector (Stratagene). *Staphylococcus aureus* DNA (> 100 kb) was partially digested in a reaction mixture (200 µl) containing 50 µg DNA, 1X Sau3AI buffer, 20 units Sau3AI for 6 min. at 23 C. The digested DNA was phenol-extracted and centrifuges over a 10- 40% sucrose gradient. Fractions containing genomic DNA of 15-25 kb were recovered by precipitation. One µl of fragments was used with 1 µl of DASHII vector (Stratagene) in the recommended ligation reaction. One µl of the ligation mixture was used per packaging reaction following the recommended protocol with the Gigapack II XL Packaging Extract Phage were plated directly without amplification from the packaging mixture (after dilution with 500 µl of recommended SM buffer and chloroform treatment). Yield was about 2.5x10⁹ pfu/µl.

An amplified library was prepared from the primary packaging mixture according to the manufacturer's protocol. The amplified library is stored frozen in 7% dimethylsulfoxide. The phage titer is approximately 1x10⁹ pfu/ml.

Mini-liquid lysates (0.1 µl) are prepared from randomly selected plaques and template is prepared by long range PCR. Samples are PCR amplified using modified T3 and T7 primers, and Elongase Supermix (LTI).

Sequencing reactions are carried out on plasmid templates using a combination of two workstations (BIOMEK 1000 and Hamilton Microlab 2200) and the Perkin-Elmer 9600 thermocycler with Applied Biosystems PRISM Ready Reaction Dye Primer Cycle Sequencing Kits for the M13 forward (M13-21) and the M13 reverse (M13RP1) primers. Dye terminator sequencing reactions are carried out on the lambda templates on a Perkin-Elmer 9600 Thermocycler using the Applied Biosystems Ready Reaction Dye Terminator Cycle Sequencing kits. Modified T7 and T3 primers are used to sequence the ends of the inserts from the Lambda DASH II library. Sequencing reactions are on a combination of AB 373 DNA Sequencers and ABI 377 DNA sequencers. All of the dye terminator sequencing reactions are analyzed using the 2X 9 hour module on the AB 377. Dye primer reactions are analyzed on a combination of ABI 373 and ABI 377 DNA sequencers. The overall sequencing success rate very approximately is about 85% for M13-21 and M13RP1 sequences and 65% for dye-terminator reactions. The average usable read length is 485 bp for M13-21 sequences, 445bp for M13RP1 sequences, and 375 bp for dye-terminator reactions.

4. Protocol for Automated Cycle Sequencing

The sequencing was carried out using Hamilton Microstation 2200, Perkin Elmer 9600 thermocyclers, ABI 373 and ABI 377 Automated DNA Sequencers. The Hamilton combines pre-aliquoted templates and reaction mixes consisting of deoxy- and dideoxynucleotides, the thermostable Taq DNA polymerase, fluorescently-labelled sequencing primers, and reaction buffer. Reaction mixes and templates were combined in the wells of a 96-well thermocycling plate and transferred to the Perkin Elmer 9600 thermocycler. Thirty consecutive cycles of linear amplification (i.e., one primer synthesis) steps were performed including denaturation, annealing of primer and template, and extension; i.e., DNA synthesis. A heated lid with rubber gaskets on the thermocycling plate prevents evaporation without the need for an oil overlay.

Two sequencing protocols were used: one for dye-labelled primers and a second for dye-labelled dideoxy chain terminators. The shotgun sequencing involves use of four dye-labelled sequencing primers, one for each of the four terminator nucleotide. Each dye-primer was labelled with a different fluorescent dye, permitting the four individual reactions to be combined into one lane of the 373 or 377 DNA Sequencer for electrophoresis, detection, and base-

calling. ABI currently supplies premixed reaction mixes in bulk packages containing all the necessary non-template reagents for sequencing. Sequencing can be done with both plasmid and PCR-generated templates with both dye-primers and dye-terminators with approximately equal fidelity, although plasmid templates generally give longer usable sequences.

Thirty-two reactions were loaded per ABI 373 Sequencer each day and 96 samples can be loaded on an ABI 377 per day. Electrophoresis was run overnight (ABI 373) or for 2 1/2 hours (ABI 377) following the manufacturer's protocols. Following electrophoresis and fluorescence detection, the ABI 373 or ABI 377 performs automatic lane tracking and base-calling. The lane-tracking was confirmed visually. Each sequence electropherogram (or fluorescence lane trace) was inspected visually and assessed for quality. Trailing sequences of low quality were removed and the sequence itself was loaded via software to a Sybase database (archived daily to 8mm tape). Leading vector polylinker sequence was removed automatically by a software program. Average edited lengths of sequences from the standard ABI 373 or ABI 377 were around 400 bp and depend mostly on the quality of the template used for the sequencing reaction.

INFORMATICS

1. Data Management

A number of information management systems for a large-scale sequencing lab have been developed. (For review see, for instance, Kerlavage *et al.*, *Proceedings of the Twenty-Sixth Annual Hawaii International Conference on System Sciences*, IEEE Computer Society Press, Washington D. C., 585 (1993)) The system used to collect and assemble the sequence data was developed using the Sybase relational database management system and was designed to automate data flow wherever possible and to reduce user error. The database stores and correlates all information collected during the entire operation from template preparation to final analysis of the genome. Because the raw output of the ABI 373 Sequencers was based on a Macintosh platform and the data management system chosen was based on a Unix platform, it was necessary to design and implement a variety of multi-user, client-server applications which allow the raw data as well as analysis results to flow seamlessly into the database with a minimum of user effort.

2. Assembly

An assembly engine (TIGR Assembler) developed for the rapid and accurate assembly of thousands of sequence fragments was employed to generate contigs. The TIGR assembler simultaneously clusters and assembles fragments of the genome. In order to obtain the speed necessary to assemble more than 10^4 fragments, the algorithm builds a hash table of 12 bp oligonucleotide subsequences to generate a list of potential sequence fragment overlaps. The number of potential overlaps for each fragment determines which fragments are likely to fall into repetitive elements. Beginning with a single seed sequence fragment, TIGR Assembler extends the current contig by attempting to add the best matching fragment based on oligonucleotide content. The contig and candidate fragment are aligned using a modified version of the Smith-Waterman algorithm which provides for optimal gapped alignments (Waterman, M. S., *Methods in Enzymology* 164: 765 (1988)). The contig is extended by the fragment only if strict criteria for the quality of the match are met. The match criteria include the minimum length of overlap, the maximum length of an unmatched end, and the minimum percentage match. These criteria are automatically lowered by the algorithm in regions of minimal coverage and raised in regions with a possible repetitive element. The number of potential overlaps for each fragment determines which fragments are likely to fall into repetitive elements. Fragments representing the boundaries of repetitive elements and potentially chimeric fragments are often rejected based on partial mismatches at the ends of alignments and excluded from the current contig. TIGR Assembler is designed to take advantage of clone size information coupled with sequencing from both ends of each template. It enforces the constraint that sequence fragments from two ends of the same template point toward one another in the contig and are located within a certain range of base pairs (definable for each clone based on the known clone size range for a given library).

3. Identifying Genes

The predicted coding regions of the *Staphylococcus aureus* genome were initially defined with the program zorf, which finds ORFs of a minimum length. The predicted coding region sequences were used in searches against a database of all *Staphylococcus aureus* nucleotide sequences from GenBank (release 92.0), using the BLASTN search method to identify overlaps of 50 or more nucleotides with at least a 95% identity. Those ORFs with nucleotide sequence matches are shown in Table 1. The ORFs without such matches were translated to protein sequences and compared to a non-redundant database of known proteins generated by combining the Swiss-prot, PIR and GenPept databases. ORFs of at least 80 amino acids that matched a database protein with BLASTP probability less than or equal to 0.01 are shown in Table 2. The table also lists assigned functions based on the closest overlaps in the databases.

ORFs of at least 120 amino acids that did not match protein or nucleotide sequences in the databases at these levels are shown in Table 3.

ILLUSTRATIVE APPLICATIONS

1. Production of an Antibody to a *Staphylococcus aureus* Protein

Substantially pure protein or polypeptide is isolated from the transfected or transformed cells using any one of the methods known in the art. The protein can also be produced in a recombinant prokaryotic expression system, such as *E. coli*, or can be chemically synthesized. Concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows.

2. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, G. and Milstein, C., *Nature* 256:495 (1975) or modifications of the methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, E., *Meth. Enzymol.* 70:419 (1980), and modified methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis, L. *et al.* *Basic Methods in Molecular Biology* Elsevier, New York. Section 21-2 (1989).

3. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein described above, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than other and may require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis, J. *et al.*, *J. Clin. Endocrinol. Metab.* 33:988-991 (1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, O. *et al.*, Chap. 19 in: *Handbook of Experimental Immunology*, Wier, D., ed, Blackwell (1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12M). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: *Manual of Clinical Immunology*, second edition, Rose and Friedman, eds., Amer. Soc. For Microbiology, Washington, D. C. (1980).

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. In addition, they are useful in various animal models of *Staphylococcal* disease known to those of skill in the art as a means of evaluating the protein used to make the antibody as a potential vaccine target or as a means of evaluating the antibody as a potential immunotherapeutic reagent.

3. Preparation of PCR Primers and Amplification of DNA

Various fragments of the *Staphylococcus aureus* genome, such as those of Tables 1-3 and SEQ ID NOS:1-5,191 can be used, in accordance with the present invention, to prepare PCR primers for a variety of uses. The PCR primers are preferably at least 15 bases, and more preferably at least 18 bases in length. When selecting a primer sequence, it is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approxi-

mately the same. The PCR primers and amplified DNA of this Example find use in the Examples that follow.

4. Gene expression from DNA Sequences Corresponding to ORFs

5 A fragment of the *Staphylococcus aureus* genome provided in Tables 1-3 is introduced into an expression vector using conventional technology. Techniques to transfer cloned sequences into expression vectors that direct protein translation in mammalian, yeast, insect or bacterial expression systems are well known in the art. Commercially available vectors and expression systems are available from a variety of suppliers including Stratagene (La Jolla, California), Promega (Madison, Wisconsin), and Invitrogen (San Diego, California). If desired, to enhance expression and facilitate proper protein folding, the codon context and codon pairing of the sequence may be optimized for the particular expression organism, as explained by Hatfield *et al.*, U. S. Patent No. 5,082,767, incorporated herein by this reference.

The following is provided as one exemplary method to generate polypeptide(s) from cloned ORFs of the *Staphylococcus aureus* genome fragment. Bacterial ORFs generally lack a poly A addition signal. The addition signal sequence can be added to the construct by, for example, splicing out the poly A addition sequence from pSG5 (Stratagene) using BglI and SalI restriction endonuclease enzymes and incorporating it into the mammalian expression vector pXT1 (Stratagene) for use in eukaryotic expression systems. pXT1 contains the LTRs and a portion of the gag gene of Moloney Murine Leukemia Virus. The positions of the LTRs in the construct allow efficient stable transfection. The vector includes the Herpes Simplex thymidine kinase promoter and the selectable neomycin gene. The *Staphylococcus aureus* DNA is obtained by PCR from the bacterial vector using oligonucleotide primers complementary to the *Staphylococcus aureus* DNA and containing restriction endonuclease sequences for PstI incorporated into the 5' primer and BglII at the 5' end of the corresponding *Staphylococcus aureus* DNA 3' primer, taking care to ensure that the *Staphylococcus aureus* DNA is positioned such that its followed with the poly A addition sequence. The purified fragment obtained from the resulting PCR reaction is digested with PstI, blunt ended with an exonuclease, digested with BglII, purified and ligated to pXT1, now containing a poly A addition sequence and digested BglII.

25 The ligated product is transfected into mouse NIH 3T3 cells using Lipofectin (Life Technologies, Inc., Grand Island, New York) under conditions outlined in the product specification. Positive transfectants are selected after growing the transfected cells in 600 ug/ml G418 (Sigma, St. Louis, Missouri). The protein is preferably released into the supernatant. However if the protein has membrane binding domains, the protein may additionally be retained within the cell or expression may be restricted to the cell surface. Since it may be necessary to purify and locate the transfected product, synthetic 15-mer peptides synthesized from the predicted *Staphylococcus aureus* DNA sequence are injected into mice to generate antibody to the polypeptide encoded by the *Staphylococcus aureus* DNA.

Alternatively and if antibody production is not possible, the *Staphylococcus aureus* DNA sequence is additionally incorporated into eukaryotic expression vectors and expressed as, for example, a globin fusion. Antibody to the globin moiety then is used to purify the chimeric protein. Corresponding protease cleavage sites are engineered between the globin moiety and the polypeptide encoded by the *Staphylococcus aureus* DNA so that the latter may be freed from the formed by simple protease digestion. One useful expression vector for generating globin chimerics is pSG5 (Stratagene). This vector encodes a rabbit globin. Intron II of the rabbit globin gene facilitates splicing of the expressed transcript, and the polyadenylation signal incorporated into the construct increases the level of expression. These techniques are well known to those skilled in the art of molecular biology. Standard methods are published in methods texts such as Davis *et al.*, cited elsewhere herein, and many of the methods are available from the technical assistance representatives from Stratagene, Life Technologies, Inc., or Promega. Polypeptides of the invention also may be produced using *in vitro* translation systems such as *in vitro* ExpressTM Translation Kit (Stratagene).

45 While the present invention has been described in some detail for purposes of clarity and understanding, one skilled in the art will appreciate that various changes in form and detail can be made without departing from the true scope of the invention.

All patents, patent applications and publications referred to above are hereby incorporated by reference.

50

55

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	Percent ident	HSP nt length	ORF nt length
1	1	1412	757	emb K17301 SAM1	S. aureus DNA for hid gene and for part of agr gene	100	663	663
1	2	3273	2452	emb K52543 SAAG	S. aureus agrA, agrB and hid genes	99	809	822
1	5	6418	5651	db D14711 STAH	Staphylococcus aureus HSP10 and HSP60 genes	98	223	768
5	1	807	439	emb X72700 SAPV	S. aureus genes for S and F components of Panton-Valentine leucocidins	81	216	369
5	4	5031	3571	emb X72700 SAPV	S. aureus genes for S and F components of Panton-Valentine leucocidins	95	424	1461
10	1	86	904	gb U25288	Staphylococcus aureus gyrase-like protein alpha and beta subunit (grlA and grlB) genes, complete cds	98	715	819
16	5	5302	6246	gb U35773	Staphylococcus aureus prolipoprotein diacylglycerol transferase (lgt) gene, complete cds	94	251	945
16	6	6249	7091	gb U35773	Staphylococcus aureus prolipoprotein diacylglycerol transferase (lgt) gene, complete cds	99	843	843
16	7	7084	7584	gb U35773	Staphylococcus aureus prolipoprotein diacylglycerol transferase (lgt) gene, complete cds	99	342	501
20	1	995	549	gb U19300	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	100	443	447
20	2	1011	841	gb U19300	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	91	137	171
20	3	2010	1798	gb U19300	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	100	110	213
20	4	5300	3825	gb H16714	Staphylococcus aureus peptidoglycan hydrolase gene, complete cds	100	948	1476
20	5	4788	4282	gb H16714	Staphylococcus aureus peptidoglycan hydrolase gene, complete cds	100	309	507
26	1	2	145	gb U41072	Staphylococcus aureus isoleucyl-tRNA synthetase (lles) gene, partial cds	100	126	144
26	2	84	557	gb U41072	Staphylococcus aureus isoleucyl-tRNA synthetase (lles) gene, partial cds	99	430	474
26	3	763	3531	emb X74219 SAIL	S. aureus gene for isoleucyl-tRNA synthetase	99	2769	2769
29	3	1261	4392	gb U66665	Staphylococcus aureus DNA fragment with class II promoter activity	100	117	332
31	14	14977	13463	emb X73889 SAP1	S. aureus genes p1 and p2	99	1351	1515
31	15	14241	13855	emb X73889 SAP1	S. aureus genes p1 and p2	98	258	387
38	17	14284	13112	gb U12715	S. aureus gdh gene encoding lipase (glycerol ester hydrolase)	100	372	1173
38	18	13434	15518	gb U12715	S. aureus gdh gene encoding lipase (glycerol ester hydrolase)	100	2085	2085

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
46	2	519	1727	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8D, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	1209	1209
46	3	1720	2295	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	576	576
46	4	2259	3182	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	97	924	924
46	5	3173	4498	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	1283	1326
46	6	4536	5720	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	1185	1185
46	7	6455	6120	[gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	99	278	336
48	1	2	955	[gb U23893	Staphylococcus aureus recA gene, complete cds	99	954	954
50	3	4465	2924	[emb X85029 SAAR	S. aureus AhpC gene	100	88	1542
50	4	4108	3515	[emb X85029 SAAR	S. aureus AhpC gene	98	540	594
54	3	5074	3392	[emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	100	1668	1683
54	4	4865	4122	[emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	99	720	744
54	5	5056	4562	[emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	100	463	495
54	6	11386	8300	[gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	100	3087	3087
58	3	1743	2819	[emb X87104 SADM	S. aureus mdr, pbp4 and tagD genes (SG511-95 isolate)	89	68	1077
58	4	2858	3280	[emb X91786 SAPB	S. aureus abcA, pbp4, and tagD genes	99	423	423
58	5	6005	4701	[emb X91786 SAPB	S. aureus abcA, pbp4, and tagD genes	99	1305	1305
58	6	5677	5378	[gb J29478	Staphylococcus aureus ABC transporter-like protein AbcA (abcA) gene, partial cds	100	300	300
58	7	5086	6840	[emb X91786 SAPB	S. aureus abcA, pbp4, and tagD genes	99	1755	1755
72	1	888	445	[gb J21856	S. aureus agr gene encoding an accessory gene regulator protein, complete cds	100	444	444
72	2	2457	1453	[emb X52343 SAAG	S. aureus agrA, agrB and hld genes	99	673	1005

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
82	1	357	3917	emb X64172 SARP	S.aureus rplL, orf202, rpoB(orf1) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	2396	3561
82	2	4027	7677	emb X89233 SARP	S.aureus DNA for rpoC gene	99	3171	3651
82	3	7745	8068	gb U20869	Staphylococcus aureus ribosomal protein S12 (rpsL) gene, complete cds, ribosomal protein S7 (rpsG) and ORF 1 genes, partial cds	100	320	324
82	4	8103	8579	gb U20869	Staphylococcus aureus ribosomal protein S12 (rpsL) gene, complete cds, ribosomal protein S7 (rpsG) and ORF 1 genes, partial cds	100	477	477
82	5	8618	8821	gb U20869	Staphylococcus aureus ribosomal protein S12 (rpsL) gene, complete cds, ribosomal protein S7 (rpsG) and ORF 1 genes, partial cds	100	154	204
84	1	18	191	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	164	174
84	2	189	893	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	94	705	705
84	3	887	1660	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	99	774	774
84	4	1584	3503	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	98	1920	1920
84	5	3394	4521	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	97	1128	1128
84	6	4519	5643	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	97	1125	1125
96	2	1245	3896	emb Z18852 SACF	S.aureus gene for clumping factor	83	660	2652
97	2	625	882	gb U41072	Staphylococcus aureus isoleucyl-tRNA synthetase (iles) gene, partial cds	97	68	758
111	1	3	452	gb U41499	Staphylococcus aureus ORF1, partial cds, ORF2, ORF3, autolysin (atl) genes, complete cds	100	450	450
111	2	526	1041	gb U41499	Staphylococcus aureus ORF1, partial cds, ORF2, ORF3, autolysin (atl) genes, complete cds	99	516	516
117	2	1278	1958	gb M83994	Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	100	61	681

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
118	4	3787	4254	dbj U06690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	99	467	468
130	4	2557	3640	emb X13290 SATN	Staphylococcus aureus multi-resistance plasmid pSKI DNA containing transposon Tn4003	78	956	1044
130	5	3813	4265	omb Z16422 SADI	S. aureus dfrB gene for dihydrofolate reductase	98	416	453
130	6	4309	5172	emb Z16422 SADI	S. aureus dfrB gene for dihydrofolate reductase	98	607	864
136	4	5296	6207	emb X71437 SAGY	S. aureus genes gyrB, gyrA and recF (partial)	97	838	912
136	5	11680	8987	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	100	2694	2694
136	6	12886	10940	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	99	1947	1947
136	7	12592	11765	gb S77055	recF cluster: dnaA=replicase assembly protein...gyrB-DNA gyrase beta subunit [Staphylococcus aureus, Y8886, Genomic, 3573 nt]	99	822	828
143	3	4171	2867	gb U36379	Staphylococcus aureus S-adenosylmethionine synthetase gene, complete cds	99	1305	1305
143	4	3100	4281	gb L42943	Staphylococcus aureus (clone KIN50) phosphoenolpyruvate carboxykinase (pckA) gene, complete cds	100	1170	1182
143	5	4254	4718	gb U51133	Staphylococcus aureus phosphoenolpyruvate carboxykinase (pckA) gene, complete cds	100	449	465
143	9	6977	7261	gb U51132	Staphylococcus aureus o-succinylbenzoic acid CoA ligase (menc), and o-succinylbenzoic acid synthetase (menc) genes, complete cds	100	75	285
143	10	9464	8161	gb U51132	Staphylococcus aureus o-succinylbenzoic acid CoA ligase (menc), and o-succinylbenzoic acid synthetase (menc) genes, complete cds	100	1104	1104
143	11	11232	9748	gb U51132	Staphylococcus aureus o-succinylbenzoic acid CoA ligase (menc), and o-succinylbenzoic acid synthetase (menc) genes, complete cds	100	1485	1485
143	12	10739	10320	gb U51132	Staphylococcus aureus o-succinylbenzoic acid CoA ligase (menc), and o-succinylbenzoic acid synthetase (menc) genes, complete cds	100	332	420
152	5	2454	3437	emb X58434 SAPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydroliipoamide acetyltransferase and dihydroliipoamide dehydrogenase	99	305	984
152	6	3513	4820	emb X58434 SAPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydroliipoamide acetyltransferase and dihydroliipoamide dehydrogenase	98	1308	1308
152	7	4818	6230	emb X58434 SAPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydroliipoamide acetyltransferase and dihydroliipoamide dehydrogenase	99	1413	1413
153	1	387	1526	gb S77055	recF cluster: dnaA=replicase assembly protein...gyrB-DNA gyrase beta subunit [Staphylococcus aureus, Y8886, Genomic, 3573 nt]	99	1140	1140
153	2	1877	2152	gb S77055	recF cluster: dnaA=replicase assembly protein...gyrB-DNA gyrase beta subunit [Staphylococcus aureus, Y8886, Genomic, 3573 nt]	100	276	276

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
153	3	2143	2289	gb S77055	recF cluster: dnaA-replisome assembly protein...gyrB-DNA gyrase beta subunit (Staphylococcus aureus, Y8886, Genomic, 5 genes, 3573 nt)	99	113	147
154	10	10792	9314	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	91	134	1479
154	11	9935	9615	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	99	229	321
154	12	9943	10167	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	94	123	225
154	13	10069	11501	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	99	1326	1413
159	2	2195	1212	gb J028079 STAP	Staphylococcus aureus gene for penicillin-binding protein 1, complete cds	100	71	984
161	3	2596	2270	gb H83998	Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	92	203	327
162	1	1406	705	gb U21221	Staphylococcus aureus hyaluronate lyase (hysA) gene, complete cds	100	702	702
163	4	1263	1772	gb U19770	Staphylococcus aureus pyrrolidone carboxyl peptidase (pcp) gene, complete cds	96	127	510
164	7	4774	9117	gb J086727 D867	Staphylococcus aureus DNA for DNA polymerase III, complete cds	99	3470	4344
168	7	7448	6447	gb U21636	Staphylococcus aureus cmp-binding-factor 1 (cbf1) and ORF X genes, complete cds	100	1002	1002
168	8	9538	7961	gb U21636	Staphylococcus aureus cmp-binding-factor 1 (cbf1) and ORF X genes, complete cds	99	1158	1578
173	6	9240	7801	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	100	1440	1440
173	7	11252	9522	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	99	1731	1731
173	8	8285	8704	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	100	420	420
173	9	10168	9839	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	100	330	330
173	10	11815	10829	embLX14827 SALA	Staphylococcus aureus lacC and lacD genes	100	987	987
173	11	12721	11774	embLX14827 SALA	Staphylococcus aureus lacC and lacD genes	100	948	948
173	12	12838	12305	gb M64724	S. aureus tagatase 6-phosphate isomerase gene, complete cds	100	534	534
173	13	13243	12773	gb H32103	Staphylococcus aureus lac repressor (lacR) gene, complete cds and lacA repressor (lacA), partial cds	100	471	471
173	14	14633	13866	gb H32103	Staphylococcus aureus lac repressor (lacR) gene, complete cds and lacA repressor (lacA), partial cds	100	768	768

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent Ident	HSP nt length	ORF nt length
178	1	2	655	gb U52961	Staphylococcus aureus holin-like protein LrgA (lrgA) and LrgB (lrgB) genes, complete cds	100	115	654
178	2	2201	1482	gb U52961	Staphylococcus aureus holin-like protein LrgA (lrgA) and LrgB (lrgB) genes, complete cds	100	720	720
178	3	2361	1909	gb U52961	Staphylococcus aureus holin-like protein LrgA (lrgA) and LrgB (lrgB) genes, complete cds	100	453	453
178	4	1551	1853	gb U52961	Staphylococcus aureus holin-like protein LrgA (lrgA) and LrgB (lrgB) genes, complete cds	100	303	303
178	5	3541	2777	gb U42945	Staphylococcus aureus lysS and lysR genes, complete cds	99	765	765
178	6	3294	3025	gb U42945	Staphylococcus aureus lysS and lysR genes, complete cds	99	270	270
181	1	1114	590	gb M61177	S. aureus sigma factor (plac) gene, complete cds	99	499	525
182	1	3	341	emb M61307 SASP	Staphylococcus aureus spa gene for protein A	98	277	319
182	2	690	2312	gb J01786	S. aureus spa gene coding for protein A, complete cds	97	1332	1623
182	3	5861	4251	emb M61307 SASP	Staphylococcus aureus spa gene for protein A	99	119	1611
185	1	3	824	gb U31979	Staphylococcus aureus chorismate synthase (aroC) and nucleoside diphosphate kinase (ndk) genes, complete cds, dehydrogenase synthase (aroB) and geranylgeranyl pyrophosphate synthetase homolog (gerCC) genes, partial cds	90	132	822
191	3	841	2760	emb X17679 SACO	Staphylococcus aureus coa gene for coagulase	99	1920	1920
191	4	2967	3143	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	99	177	177
191	5	5768	4566	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	99	250	1203
196	1	1741	872	gb U36472	Staphylococcus aureus lysyl-tRNA synthetase gene, complete cds, transfer RNA (tRNA) genes, 5S ribosomal RNA (5S rRNA) gene, 16S ribosomal RNA (16S rRNA) gene, 23S ribosomal RNA (23S rRNA) gene	99	870	870
198	3	1688	2011	emb X93205 SAPT	S. aureus ptbH and ptbI genes	99	324	324
198	4	2005	2310	emb X93205 SAPT	S. aureus ptbH and ptbI genes	97	304	306
202	1	163	1305	emb X97985 SA12	S. aureus orf1, 2, 3 & 4	99	1143	1143
202	2	1303	2175	emb X73889 SAP1	S. aureus genes P1 and P2	94	444	873
210	1	3114	1558	dbj D17366 STAA	Staphylococcus aureus atl gene for autolysin, complete cds and other ORFs	99	1552	1557
210	2	2939	2232	gb U41499	Staphylococcus aureus ORF1, partial cds, ORF2, ORF3, autolysin (atl) genes, complete cds	99	684	708
214	11	7439	7770	dbj D86240 D862	Staphylococcus aureus gene for unknown function and dlt operon dltA, dltB, dltC and dltD genes, complete cds	96	157	342

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	Percent ident	HSP nt length	ORF nt length
216	3	398	1318	emb X72700 SAPV	S. aureus genes for S and F components of Pantone-Valentine leucocidins	88	265	921
219	2	1810	1073	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	100	50	738
219	3	2979	2035	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	99	945	945
219	4	4359	3196	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	99	1164	1164
219	5	7044	5176	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	98	1869	1869
219	6	6557	5883	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	99	675	675
219	7	6801	6334	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP20; HSP70; HSP40; ORF35, complete cds	98	468	468
221	8	10816	10034	gb U19298	Staphylococcus aureus phosphatidylinositol-specific phospholipase C (pic) gene, complete cds	91	67	783
223	1	2855	1506	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	99	102	1350
234	1	2	1357	emb X97985 SA12	S. aureus orfs 1,2,3 & 4	100	176	1356
234	2	1694	2485	emb X97985 SA12	S. aureus orfs 1,2,3 & 4	100	792	792
234	3	2648	3148	emb X97985 SA12	S. aureus orfs 1,2,3 & 4	99	501	501
234	4	3120	4604	emb X97985 SA12	S. aureus orfs 1,2,3 & 4	99	1305	1485
236	6	3826	5322	gb U48826	Staphylococcus aureus elastin binding protein (ebpS) gene, complete cds	96	648	1497
248	1	2	403	emb X62286 SAP2	S. aureus DNA for penicillin-binding protein 2	100	103	402
248	2	388	852	gb U35426	Staphylococcus aureus penicillin-binding protein 2 (pbp2) gene, complete cds	99	465	465
253	2	1519	1093	gb U46541	Staphylococcus aureus sacA gene, complete cds	96	447	447
254	2	150	1815	gb U57060	Staphylococcus aureus sacA gene, complete cds	94	142	1686
254	3	1973	2728	gb U57060	Staphylococcus aureus sacA gene, complete cds	99	756	756
260	1	2	1900	gb U90693	Staphylococcus aureus glycerol ester hydrolase (lip) gene, complete cds	99	1213	1899
265	1	1	942	dbj D21131 STAS	Staphylococcus aureus gene for a participant in homogeneous expression of high-level methicillin resistance, complete cds	99	941	942

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	Percent Ident	HSP nt length	ORF nt length
265	2	688	476	dbj D21131 STAS	Staphylococcus aureus gene for a participant in homogeneous expression of high-level methicillin resistance, complete cds	99	213	213
265	3	2418	1765	dbj D21131 STAS	Staphylococcus aureus gene for a participant in homogeneous expression of high-level methicillin resistance, complete cds	98	69	654
266	1	2	1018	dbj D14711 STAH	Staphylococcus aureus HSP10 and HSP60 genes	98	743	1017
282	1	1	525	gb F72488	hmb-porphobilinogen synthase (Staphylococcus aureus, SA1959, Genomic, 1087 nt)	100	110	525
282	2	516	1502	gb F72488	hmb-porphobilinogen synthase (Staphylococcus aureus, SA1959, Genomic, 1087 nt)	100	952	987
284	1	3	170	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	98	84	168
284	2	282	1034	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	100	712	753
284	3	1028	2026	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	99	979	999
284	4	1990	2202	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	98	187	213
289	3	1536	1991	gb H32470	S. aureus Sau3AI-restriction-enzyme and Sau3AI-modification-enzyme genes, complete cds	99	338	456
303	1	2	868	gb L01055	Staphylococcus aureus gamma-hemolysin components A, B and C (hlgA, hlgB, hlgC) genes, complete cds	99	867	867
303	2	1409	2383	gb L01055	Staphylococcus aureus gamma-hemolysin components A, B and C (hlgA, hlgB, hlgC) genes, complete cds	100	975	975
303	3	2367	3161	gb L01055	Staphylococcus aureus gamma-hemolysin components A, B and C (hlgA, hlgB, hlgC) genes, complete cds	99	793	795
305	1	2707	1355	dbj D17366 STAA	Staphylococcus aureus atl gene for autolysin, complete cds and other ORFs	99	1343	1353
311	1	2628	1315	gb L42945	Staphylococcus aureus lvtS and lvtR genes, complete cds	98	1314	1314
312	6	7019	7870	gb L4017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	74	351	852
323	1	1998	1003	gb U31175	Staphylococcus aureus D-specific D-2-hydroxyacid dehydrogenase (ddh) gene, complete cds	98	996	996
326	1	1	237	emb Y00356 SASP	Staphylococcus aureus V8 serine protease gene	100	108	237
338	1	687	388	emb X64389 SALE	S. aureus leuF-P83 gene for F component of leucocidin R	98	259	300
338	2	1828	1088	emb X64389 SALE	S. aureus leuF-P83 gene for F component of leucocidin R	97	137	741

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	ORF nt length
342	2	579	1754	[gb U06462]	[Staphylococcus aureus SA4 FtsZ (ftsZ) gene, complete cds	100	1176
344	2	517	1248	[emb V01281 SNUU	[S. aureus mRNA for nuclease	98	732
349	1	457	230	[gb H20393]	[S. aureus bacteriophage phi-11 attachment site (attB)	96	172
353	1	1016	516	[gb H03994]	[Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	100	187
353	2	1582	1046	[gb H03994]	[Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	99	537
356	1	3	674	[gb U20503]	[Staphylococcus aureus MHC class II analog gene, complete cds	75	671
361	1	1	903	[gb L19298]	[Staphylococcus aureus phosphatidylinositol-specific phospholipase C (pic) gene, complete cds	98	747
361	2	1103	1507	[gb L19298]	[Staphylococcus aureus phosphatidylinositol-specific phospholipase C (pic) gene, complete cds	97	68
373	1	3	1148	[emb X62288 SAPE	[S. aureus DNA for penicillin-binding protein 2	99	1146
389	3	1904	1248	[emb X62282 SATS	[S. aureus target site DNA for IS431 insertion	97	349
400	1	1	540	[emb X61716 SMHL	[S. aureus hlb gene encoding sphingomyelinase	99	389
400	2	1693	1187	[emb X13404 SAHL	[Staphylococcus aureus hlb gene for beta-hemolysin	99	178
408	1	1810	1049	[gb S76213]	[esp23-alkaline shock protein 23 (methicillin resistant) [Staphylococcus aureus, 912. Genomic, 1360 ntl	99	163
418	1	2	217	[gb L41499]	[Staphylococcus aureus ORF1, partial cds, ORF2, ORF3, autolysin (atl) genes, complete cds	100	216
418	2	854	639	[dbj D17366 STAA	[Staphylococcus aureus atl gene for autolysin, complete cds and other ORFs	100	188
421	2	1262	2509	[gb L43098]	[Transposon Tn5404 and insertion sequences IS1181 and IS1182 (from Staphylococcus aureus) DNA	99	1248
422	1	2	325	[gb K02985]	[S. aureus (strain RN450) transposon Tn554 insertion site	96	200
427	1	865	434	[dbj D28879 STAP	[Staphylococcus aureus gene for penicillin-binding protein 1, complete cds	100	432
427	2	1829	1122	[dbj D28879 STAP	[Staphylococcus aureus gene for penicillin-binding protein 1, complete cds	100	151
435	1	2	808	[dbj D86240 D862	[Staphylococcus aureus gene for unknown function and dit operon ditA, ditB, ditC and ditD genes, complete cds	100	556
435	2	832	999	[dbj D86240 D862	[Staphylococcus aureus gene for unknown function and dit operon ditA, ditB, ditC and ditD genes, complete cds	100	134
436	1	1341	685	[emb X17688 SAFE	[S. aureus factor essential for expression of methicillin resistance (fema) gene, complete cds, and trpA gene, 3' end	97	657

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
436	2	2403	1657	emb X17688 SAFE	S. aureus factor essential for expression of methicillin resistance (fema) gene, complete cds, and trpA gene, 3' and	100	294	747
442	1	347	1300	emb X72700 SAVP	S. aureus genes for S and F components of Pantone-Valentine leucocidins	84	204	954
445	2	1906	2178	gb L01055	Staphylococcus aureus gamma-hemolysin components A, B and C (higa, hlgB, hlgC) genes, complete cds	98	187	273
447	1	167	1078	gb U19770	Staphylococcus aureus pyrrolidone carboxyl peptidase (pcp) gene, complete cds	100	514	912
447	2	1176	1784	gb U19770	Staphylococcus aureus pyrrolidone carboxyl peptidase (pcp) gene, complete cds	95	597	609
454	3	7309	4319	emb Z18652 SACF	S. aureus gene for clumping factor	75	653	2991
472	4	7896	5479	gb L25288	Staphylococcus aureus gyrase-like protein alpha and beta subunit (grIA and grIB) genes, complete cds	99	2418	2418
472	5	8120	6792	gb L25288	Staphylococcus aureus gyrase-like protein alpha and beta subunit (grIA and grIB) genes, complete cds	99	1328	1329
475	2	566	889	emb X52543 SAAG	S. aureus agrA, agrB and hld genes	100	76	324
481	4	1922	1560	emb X64172 SARP	S. aureus rplL, orf202, rpsB(rif) and rpsC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	250	363
481	5	1244	1534	emb X64172 SARP	S. aureus rplL, orf202, rpsB(rif) and rpsC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	224	291
487	2	1388	1188	gb H83994	Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	98	72	201
489	1	2737	1370	gb U21221	Staphylococcus aureus hyaluronate lyase (hyaA) gene, complete cds	99	1368	1368
503	2	1135	653	gb H83994	Staphylococcus aureus prolipoprotein signal peptidase (lsp) gene, complete cds	100	108	483
511	3	1613	2242	gb L14017	Staphylococcus aureus methicillin-resistance protein (mecR1) gene and unknown ORF, complete cds	84	323	630
511	4	3122	2700	gb S76213	esp23-saltline shock protein 23 (methicillin resistant) [Staphylococcus aureus, 912, Genomic, 1380 nt]	96	423	423
520	2	758	1297	emb X72014 SAFI	S. aureus fib gene for fibrinogen-binding protein	99	540	540
520	3	1436	1801	emb X72013 SAFI	S. aureus fib gene for fibrinogen-binding protein	99	221	366
526	1	2150	1092	dbj U17366 STAA	Staphylococcus aureus atl gene for autolysin, complete cds and other ORFs	99	641	1059

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match	Gene name	Percent ident	HSP nt length	ORF nt length
528	2	58	963	[gb L19300]	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	99	260	906
528	3	1098	2870	[gb L19300]	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	99	866	1773
530	1	3	434	[gb U31979]	Staphylococcus aureus chorismate synthase (aroC) and nucleoside diphosphate kinase (ndk) genes, complete cds, dehydroquinase synthase (aroB) and geranylgeranyl pyrophosphate synthetase homolog (gerCC) genes, partial cds	99	432	432
530	2	1211	2395	[gb U31979]	Staphylococcus aureus chorismate synthase (aroC) and nucleoside diphosphate kinase (ndk) genes, complete cds, dehydroquinase synthase (aroB) and geranylgeranyl pyrophosphate synthetase homolog (gerCC) genes, partial cds	91	1185	1185
530	3	2409	2801	[gb U31979]	Staphylococcus aureus chorismate synthase (aroC) and nucleoside diphosphate kinase (ndk) genes, complete cds, dehydroquinase synthase (aroB) and geranylgeranyl pyrophosphate synthetase homolog (gerCC) genes, partial cds	88	181	303
530	4	2690	3484	[gb L05004]	Staphylococcus aureus dehydroquinase synthase (aroB) gene, 3' end cds; 3-phosphoshikimate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	100	75	795
530	5	3482	4792	[gb L05004]	Staphylococcus aureus dehydroquinase synthase (aroB) gene, 3' end cds; 3-phosphoshikimate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	99	905	1311
530	6	4790	5380	[gb L05004]	Staphylococcus aureus dehydroquinase synthase (aroB) gene, 3' end cds; 3-phosphoshikimate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	100	196	591
539	1	3	338	[emb X76490 SAGL]	S. aureus (bb270) glnA and glnR genes	99	336	336
539	2	336	527	[emb X76490 SAGL]	S. aureus (bb270) glnA and glnR genes	100	189	192
554	1	727	365	[gb U73374]	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	100	54	363
554	2	2175	1252	[gb U73374]	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	99	918	924
554	3	1574	1374	[gb U73374]	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	96	122	201
584	2	1019	705	[gb U21221]	Staphylococcus aureus hyaluronate lyase (hvaA) gene, complete cds	99	306	315
587	3	1475	4288	[emb Z18852 SACF]	S. aureus gene for clumping factor	98	2588	2814
594	1	3881	1953	[dbj D28879 STAP]	Staphylococcus aureus gene for penicillin-binding protein 1, complete cds	99	1873	1929

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	Percent ident	HSP nt length	ORF nt length
605	1	2	745	dbj D96240 D862	Staphylococcus aureus gene for unknown function and dlt operon dltA, dltB, dltC and dltB genes, complete cds	98	338	744
609	1	1628	816	emb X76490 SAGL	S. aureus (b2270) glrA and glrB genes	100	495	813
614	1	1280	642	gb H312103	Staphylococcus aureus lac repressor (lacR) gene, complete cds and lacA repressor (lacA), partial cds	99	639	639
626	1	2508	1255	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	100	225	1251
626	2	3315	2284	gb H63176	Staphylococcus aureus helicase required for T181 replication (pcrA) gene, complete cds	99	838	1032
629	1	1999	1001	emb X17688 SAFE	S. aureus factor essential for expression of methicillin resistance (femA) gene, complete cds, and trpA gene, 3' end	99	990	999
629	2	1407	1195	emb X17688 SAFE	S. aureus factor essential for expression of methicillin resistance (femA) gene, complete cds, and trpA gene, 3' end	98	194	213
631	2	5126	3228	emb Z18852 SACF	S. aureus gene for clumping factor	82	489	1899
632	1	3	551	emb Z30588 SAST	S. aureus (RM4220) genes for potential ABC transporter and potential membrane spanning protein	99	549	549
632	2	529	1323	emb Z30588 SAST	S. aureus (RM4220) genes for potential ABC transporter and potential membrane spanning protein	99	795	795
651	1	1909	1070	gb L19300	Staphylococcus aureus DNA sequence encoding three ORFs, complete cds; prophage phi-11 sequence homology, 5' flank	99	478	840
657	2	1800	1105	gb L14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	84	456	696
662	1	908	456	emb X13404 SAHL	Staphylococcus aureus hlb gene for beta-hemolysin	100	369	453
662	2	230	475	emb X13404 SAHL	Staphylococcus aureus hlb gene for beta-hemolysin	100	246	246
662	3	746	1399	emb X13404 SAHL	Staphylococcus aureus hlb gene for beta-hemolysin	99	653	654
682	1	956	480	gb H63177	S. aureus sigma factor (plac) gene, complete cds	100	136	477
685	1	1182	592	gb U65000	Staphylococcus aureus type-I signal peptidase Spas (spas) gene, and type-I signal peptidase Spas (spas) gene, complete cds	98	534	591
685	2	1716	1153	gb U65000	Staphylococcus aureus type-I signal peptidase Spas (spas) gene, and type-I signal peptidase Spas (spas) gene, complete cds	96	564	564
697	1	3	527	gb H63177	S. aureus sigma factor (plac) gene, complete cds	100	195	525
697	2	485	784	gb H63177	S. aureus sigma factor (plac) gene, complete cds	97	280	300

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	Percent ident	HSP nt length	ORF nt length
710	1	15	503	dbj D86240 D862	Staphylococcus aureus gene for unknown function and dlt operon dltA, dltB, dltC and dltD genes, complete cds	99	217	489
733	1	26	205	gb H80252	Staphylococcus aureus norA199 gene (which mediates active efflux of fluoroquinolones), complete cds	97	140	180
741	1	1736	1197	dbj D83951 STAL	Staphylococcus aureus DNA for LukM component, LukF-PV like component, complete cds	81	522	540
752	1	1	636	emb Y00356 SASP	Staphylococcus aureus V8 serine protease gene	99	618	636
752	2	588	956	emb Y00356 SASP	Staphylococcus aureus V8 serine protease gene	99	340	369
756	1	1108	709	emb X01645 SAT0	Staphylococcus aureus (Wood 46) gene for alpha-toxin	98	567	600
777	1	1582	950	emb Z49245 SNA2	S. aureus partial sod gene for superoxide dismutase	99	429	633
780	1	1111	557	gb U20503	Staphylococcus aureus MHC class II analog gene, complete cds	86	550	555
784	1	73	687	gb U6329	Staphylococcus aureus novel antigen gene, complete cds	99	588	615
797	1	182	544	dbj D14711 STAH	Staphylococcus aureus HSP10 and HSP40 genes	98	363	363
798	1	532	302	emb X58434 SNPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	95	196	231
823	1	3	467	gb S77055	recF cluster: dnaA, replisome assembly protein, gyrase, DNA gyrase beta subunit [Staphylococcus aureus, Y886, Genomic, 5 genes, 3573 nt]	99	156	465
848	1	348	175	gb L25288	Staphylococcus aureus gyrase-like protein alpha and beta subunit (grIA and grIB) genes, complete cds	99	174	174
848	2	476	318	gb L25288	Staphylococcus aureus gyrase-like protein alpha and beta subunit (grIA and grIB) genes, complete cds	100	131	159
866	1	792	397	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L2/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	395	396
883	1	1	285	dbj D90119 STAW	S. aureus norA gene	99	131	285
884	1	606	334	emb X52543 SAG	S. aureus agrA, agrB and hld genes	98	265	273
884	2	716	522	emb X52543 SAG	S. aureus agrA, agrB and hld genes	100	195	195
912	2	517	681	emb Z30588 SAST	S. aureus (RM4220) genes for potential ABC transporter and potential membrane spanning protein	99	163	165
917	1	2	265	gb H64724	S. aureus tagatase 6-phosphate isomerase gene, complete cds	99	247	264
917	2	238	396	gb H64724	S. aureus tagatase 6-phosphate isomerase gene, complete cds	95	147	159
918	1	2426	1215	emb X93205 SAPT	S. aureus ptmH and ptmI genes	99	1212	1212

TABLE I

S. aureus - Containing regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	Percent ident	ORF nt length
967	1	1	411	[dbj U90119 STAN	S. aureus norA gene	97	395
991	1	672	337	[emb X52543 SAAG	S. aureus agrA, agrB and hid genes	99	336
1000	1	1117	845	[gb U14017	Staphylococcus aureus methicillin-resistance protein (mecA) gene and unknown ORF, complete cds	78	190
1001	1	498	265	[dbj D86240 D862	Staphylococcus aureus gene for unknown function and dit operon dltA, dltB, dltC and dltD genes, complete cds	99	234
1010	1	1	285	[gb U21221	Staphylococcus aureus hyaluronate lyase (hysA) gene, complete cds	99	224
1046	1	656	330	[emb X72700 SAPV	S. aureus genes for S and F components of Pantone-Valentine leucocidin	85	205
1060	1	480	286	[emb X58434 SAPD	S. aureus pdsA, pdsC and pdsD genes for pyruvate decarboxylase, dihydrolypsoamide acetyltransferase and dihydrolypsoamide dehydrogenase	99	180
1073	1	1176	589	[gb X02985	S. aureus (strain RN450) transposon Tn554 insertion site	100	131
1079	1	3	230	[dbj D86240 D862	Staphylococcus aureus gene for unknown function and dit operon dltA, dltB, dltC and dltD genes, complete cds	99	228
1079	2	218	484	[dbj D86240 D862	Staphylococcus aureus gene for unknown function and dit operon dltA, dltB, dltC and dltD genes, complete cds	100	267
1079	1	460	645	[dbj D86240 D862	Staphylococcus aureus gene for unknown function and dit operon dltA, dltB, dltC and dltD genes, complete cds	100	186
1092	1	289	146	[emb X58434 SAPD	S. aureus pdsA, pdsC and pdsD genes for pyruvate decarboxylase, dihydrolypsoamide acetyltransferase and dihydrolypsoamide dehydrogenase	98	124
1143	1	1	243	[gb U63177	S. aureus sigma factor (pdsC) gene, complete cds	99	243
1157	1	2	136	[emb Z48003 SADN	S. aureus gene for DNA polymerase III	97	127
1189	1	720	361	[gb S74031	norA-NorA (ISP794) [Staphylococcus aureus, NCTC 8325, Insertion, 1820 nt]	99	360
1190	1	2	283	[gb U21854	S. aureus agr gene encoding an accessory gene regulator protein, complete cds	100	282
1190	2	1127	888	[emb X52543 SANG	S. aureus agrA, agrB and hid genes	100	240
1225	1	2	163	[emb X17679 SACO	Staphylococcus aureus coa gene for coagulase	97	124
1243	1	2	529	[dbj D86240 D862	Staphylococcus aureus gene for unknown function and dit operon dltA, dltB, dltC and dltD genes, complete cds	99	495
1244	1	1	210	[gb S74031	norA-NorA (ISP794) [Staphylococcus aureus, NCTC 8325, Insertion, 1820 nt]	100	210
1301	1	41	472	[emb X76490 SNub	S. aureus (bb270) glnA and glnR genes	99	299

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	percent ident	HSP nt length	ORF nt length
1315	1	18	326	emb X64172 SARP	S.aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	98	277	309
1519	1	2	175	dbj D28879 STAP	Staphylococcus aureus gene for penicillin-binding protein 1, complete cds	98	139	174
1863	1	1346	675	dbj D46240 D862	Staphylococcus aureus gene for unknown function and dlt operon dltA, dltB, dltC and dltD genes, complete cds	98	672	672
1797	1	644	324	gb U71374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	99	321	321
1857	1	1	192	gb H90536	Staphylococcus aureus alpha-hemolysin gene, 3' end	98	192	192
1923	1	2	181	emb X17688 SAFE	S.aureus factor essential for expression of methicillin resistance (femA) gene, complete cds, and trpA gene, 3' end	100	180	180
1957	1	2	346	gb U60589	Staphylococcus aureus novel antigen gene, complete cds	99	345	345
1988	1	1	402	dbj D86240 D862	Staphylococcus aureus gene for unknown function and dlt operon dltA, dltB, dltC and dltD genes, complete cds	100	402	402
2100	1	414	208	gb H63177	S.aureus sigma factor (plac) gene, complete cds	99	207	207
2199	1	1	402	gb U66664	Staphylococcus aureus DNA fragment with class II promoter activity	99	131	402
2537	1	308	156	emb X17688 SAFE	S.aureus factor essential for expression of methicillin resistance (femA) gene, complete cds, and trpA gene, 3' end	99	153	153
2891	1	2	400	gb J25426	Staphylococcus aureus penicillin-binding protein 2 (pbp2) gene, complete cds	99	399	399
2950	1	778	398	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP70; HSP70; HSP40; ORF35, complete cds	100	358	381
2971	1	3	398	gb U51132	Staphylococcus aureus o-succinylbenzoic acid CoA ligase (menE), and o-succinylbenzoic acid synthetase (menC) genes, complete cds	97	272	396
2978	1	618	328	gb J31979	Staphylococcus aureus chorismate synthase (aroC) and nucleoside diphosphate kinase (ndk) genes, complete cds, dehydroquinase synthase (aroB) and geranylgeranyl pyrophosphate synthetase homolog (gerCC) genes, partial cds	98	250	291
2985	1	832	464	emb X17679 SACO	Staphylococcus aureus coa gene for coagulase	98	347	369
3006	1	2170	1784	gb U11775	Staphylococcus aureus methicillin-resistant ATCC 33552 clone RNN30 16S-23S rRNA spacer region	87	82	387
3008	1	474	238	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP70; HSP70; HSP40; ORF35, complete cds	88	178	237
3008	2	451	281	dbj D30690 STAN	Staphylococcus aureus genes for ORF37; HSP70; HSP70; HSP40; ORF35, complete cds	97	320	171

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP length	ORF nt length
3011	1	793	398	emb X62992 SAFM	S. aureus fnbA gene for fibronectin binding protein B	93	72	396
3019	1	2	235	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	97	234	234
3023	1	81	233	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	87	100	153
3029	1	90	287	gb U51133	Staphylococcus aureus phosphoenolpyruvate carboxylase (pckA) gene, complete cds	100	135	198
3039	1	18	164	gb U51133	Staphylococcus aureus phosphoenolpyruvate carboxylase (pckA) gene, complete cds	97	135	147
3039	2	70	327	gb U51133	Staphylococcus aureus phosphoenolpyruvate carboxylase (pckA) gene, complete cds	77	183	259
3056	1	3	215	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	213	213
3059	1	1	261	db J03690 STAN	Staphylococcus aureus genes for ORF37, HSP20, HSP70, HSP40, ORF35, complete cds	98	234	261
3073	1	27	284	gb U06451	Staphylococcus aureus proline permease homolog (putP) gene, complete cds	99	229	258
3074	1	2	397	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	96	250	396
3088	1	3	239	db J03672 DR67	Staphylococcus aureus DNA for DNA polymerase III, complete cds	95	215	237
3097	1	444	244	emb Z18003 SARH	S. aureus gene for DNA polymerase III	97	160	201
3102	1	307	155	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	97	142	153
3121	1	568	398	emb X58434 SAPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	100	88	171
3135	1	463	233	emb X89233 SARP	S. aureus DNA for rpoC gene	98	192	231
3133	1	2	175	emb Z18852 SACF	S. aureus gene for clumping factor	96	154	174
3160	1	420	211	db J010489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	89	197	210
3176	1	1	378	emb X58434 SAPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	96	91	378
3192	1	420	211	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	98	72	210
3210	1	3	143	gb J076714	Staphylococcus aureus peptidoglycan hydrolase gene, complete cds	96	141	141

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	percent ident	ORF nt length
3232	3	2106	1282	gb L14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	71	257
353A	1	2	374	emb X89233 SARP	S. aureus DNA for rpoC gene	99	346
3543	1	392	634	gb L11530	Staphylococcus aureus transfer RNA sequence with two rRNAs	99	102
3555	1	637	320	emb Z18852 SACF	S. aureus gene for clumping factor	99	307
3559	1	3	182	emb X17679 SACF	Staphylococcus aureus coa gene for coagulase	100	141
3559	2	95	113	emb X17679 SACF	Staphylococcus aureus coa gene for coagulase	98	174
3563	1	278	141	gb U35713	Staphylococcus aureus prolipoprotein diacylglycerol transferase (lgt) gene, complete cds	100	79
3563	2	527	363	gb U35713	Staphylococcus aureus prolipoprotein diacylglycerol transferase (lgt) gene, complete cds	98	162
3566	1	3	422	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	98	175
3588	1	2	262	gb L43098	Transposon Tn5404 and insertion sequences IS1181 and IS1182 (from Staphylococcus aureus) DNA	99	253
3603	1	3	350	gb J03479	S. aureus enzyme 116-lac (lacF), enzyme 11-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	99	345
3600	1	758	381	emb Z18852 SACF	S. aureus gene for clumping factor	72	346
3602	1	788	396	emb Z18852 SACF	S. aureus gene for clumping factor	98	319
3656	1	1013	528	emb Z18852 SACF	S. aureus gene for clumping factor	84	403
3682	1	3	236	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	231
3682	2	224	415	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	112
3693	1	758	423	emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	100	229
3702	1	593	354	gb L11510	Staphylococcus aureus transfer RNA sequence with two rRNAs	54	81
3725	1	924	463	emb Z18852 SACF	S. aureus gene for clumping factor	71	367
3761	1	809	450	gb L14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	85	333
3767	1	1	402	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	98	387

TABLE 1

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	ORF nt length
3775	1	2	286	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	227
3786	1	456	229	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	100	204
3786	2	542	366	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	95	123
3798	1	3	251	emb X17679 SNCO	Staphylococcus aureus coa gene for coagulase	99	249
3813	1	793	398	gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	98	396
3819	1	184	402	emb X68425 SA23	S. aureus gene for 23S rRNA	99	161
3844	1	932	468	gb U48826	Staphylococcus aureus elastin binding protein (ebpS) gene, complete cds	87	204
3845	1	1	381	emb X58434 SNPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydroliponamide acetyltransferase and dihydroliponamide dehydrogenase	94	356
3856	1	798	400	gb U44017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	76	192
3859	1	1049	573	emb Z18852 SACF	S. aureus gene for clumping factor	85	347
3871	1	650	327	gb N76714	Staphylococcus aureus peptidoglycan hydrolase gene, complete cds	100	299
3876	1	2	253	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	100	217
3877	1	572	288	gb J03479	S. aureus enzyme III-lac (lacP), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacO) genes, complete cds	97	209
3878	1	1	237	emb X58434 SNPD	S. aureus pdhA, pdhC and pdhD genes for pyruvate decarboxylase, dihydroliponamide acetyltransferase and dihydroliponamide dehydrogenase	96	155
3888	1	3	173	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	98	171
3891	1	1	183	emb X89233 SARP	S. aureus DNA for rpoC gene	100	170
3893	2	181	357	emb X89233 SARP	S. aureus DNA for rpoC gene	98	79
3894	1	3	485	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	450
3895	1	836	420	gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	99	411
3905	1	48	239	gb U05004	Staphylococcus aureus dehydroquinase synthase (aroB) gene, 3' end cds; 3-phosphonitrate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	100	159
3905	2	188	400	gb U05004	Staphylococcus aureus dehydroquinase synthase (aroB) gene, 3' end cds; 3-phosphonitrate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	97	88

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
3910	1	3	emb X58434 SAPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	99	278	357
3915	1	330	gb U14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	75	175	330
3964	1	691	emb Z48003 SADN	S. aureus gene for DNA polymerase III	100	295	345
4007	1	199	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	98	163	192
4036	1	3	dbj D10489 STAG	Staphylococcus aureus genes for DNA gyrase A and B, complete cds	99	339	369
4046	1	692	emb Z18852 SACF	S. aureus gene for clumping factor	87	221	345
4060	1	1	emb Z18852 SACF	S. aureus gene for clumping factor	96	271	375
4061	1	860	emb Z48003 SADN	S. aureus gene for DNA polymerase III	99	429	429
4062	1	606	gb U14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	75	198	303
4085	1	58	gb U11786	Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRIV42 16S-23S rRNA spacer region	98	127	345
4088	1	2	gb U143098	Transposon Tn504 and insertion sequences IS1181 and IS1182 (from Staphylococcus aureus) DNA	99	227	300
4093	1	2	emb X58434 SAPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	99	276	276
4097	1	1	emb Z18852 SACF	S. aureus gene for clumping factor	74	307	402
4116	1	22	gb U05004	Staphylococcus aureus dehydroquinase synthase (aroH) gene, 3' end cds; 3-phosphoshikimate-1-carboxyvinyltransferase (aroA) gene, complete cds; ORF3, complete cds	98	157	381
4125	1	240	gb U73374	Staphylococcus aureus type 8 capsule genes, cap8A, cap8B, cap8C, cap8D, cap8E, cap8F, cap8G, cap8H, cap8I, cap8J, cap8K, cap8L, cap8M, cap8N, cap8O, cap8P, complete cds	100	86	162
4149	1	35	gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	99	200	213
4151	1	629	gb U14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	87	150	264
4154	1	754	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	297	357
4179	1	1	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	98	240	294

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
4203	1	1	255	emb X892133 SARP	S. aureus DNA for rpoC gene	99	239	255
4206	1	1	303	emb Z18832 SACF	S. aureus gene for clumping factor	100	236	301
4206	2	195	344	emb Z18852 SACF	S. aureus gene for clumping factor	95	65	150
4208	1	108	314	emb X58434 SAPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	89	76	207
4216	1	656	330	emb X58434 SAPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	98	326	327
4226	1	594	298	gb L11530	Staphylococcus aureus transfer RNA sequence with two rRNAs	97	132	297
4260	1	216	383	gb U11784	Staphylococcus aureus methicillin-resistant ATCC 33952 clone RNV40 16S-23S rRNA spacer region	83	141	168
4272	1	355	179	emb Z48003 SAUN	S. aureus gene for DNA polymerase III	100	164	177
4276	1	4	177	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	99	150	174
4277	1	1	270	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	265	270
4282	1	691	377	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	98	282	315
4291	1	379	191	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	183	189
4295	1	3	329	emb X16457 SAST	Staphylococcus aureus gene for staphylocoagulase	94	144	327
4313	1	435	280	gb L11530	Staphylococcus aureus transfer RNA sequence with two rRNAs	100	94	156
4315	1	3	185	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	100	158	183
4315	2	101	310	gb J03479	S. aureus enzyme III-lac (lacF), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	98	75	210
4327	1	1	294	gb J43098	Transposon Tn5404 and insertion sequences IS1181 and IS1182 (from Staphylococcus aureus) DNA	98	294	294
4360	1	603	319	gb J02910	Staphylococcus aureus ATCC 25923 16S rRNA gene, partial sequence	100	116	285
4364	1	3	146	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	95	140	144
4388	1	167	310	emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	73	119	144

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	percent ident	HSP nt length	ORF nt length
4401	1	2	313	emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	97	243	312
4421	1	36	281	dbj D12572 STA2	Staphylococcus aureus rna gene for 23S ribosomal RNA	100	112	246
4426	1	3	293	emb Z18852 SACF	S. aureus gene for clumping factor	85	185	291
4428	1	493	248	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	139	246
4462	1	2	271	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	99	270	270
4466	1	1	240	emb Z18852 SACF	S. aureus gene for clumping factor	99	231	240
4469	1	1	312	gb J03479	S. aureus enzyme III-lac (lacE), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	99	265	312
4485	1	3	263	gb U43098	Transposon Tn5404 and insertion sequences IS1181 and IS1182 (from Staphylococcus aureus) DNA	98	259	261
4492	1	74	400	gb U86227	Staphylococcus aureus DNA gyrase B subunit (gyrB) RecF homologue (recF) and DNA gyrase A subunit (gyrA) gene, complete cds	85	104	327
4497	1	515	269	emb Z18852 SACF	S. aureus gene for clumping factor	99	213	267
4529	1	2	172	emb X64172 SARP	S. aureus rplL, orf202, rpoB(rif) and rpoC genes for ribosomal protein L7/L12, hypothetical protein ORF202, DNA-directed RNA polymerase beta & beta' chains	100	151	171
4547	1	1	300	emb X62992 SAFN	S. aureus fnbB gene for fibronectin binding protein B	100	157	300
4554	1	318	160	emb Z18852 SACF	S. aureus gene for clumping factor	84	126	159
4565	1	9	227	emb Z18852 SACF	S. aureus gene for clumping factor	84	213	219
4569	1	79	222	emb Z18852 SACF	S. aureus gene for clumping factor	98	127	144
460R	1	22	216	emb X58434 SAPD	S. aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	92	168	195
4614	1	464	234	emb Z18852 SACF	S. aureus gene for clumping factor	86	169	231
4623	1	105	302	gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	99	152	198
4632	1	18	206	gb J03479	S. aureus enzyme III-lac (lacE), enzyme II-lac (lacE), and phospho-beta-galactosidase (lacG) genes, complete cds	98	183	189
4646	1	1	222	emb Z18852 SACF	S. aureus gene for clumping factor	84	100	222
46R7	1	2	166	gb J04151	S. aureus fibronectin-binding protein (fnbA) mRNA, complete cds	98	156	165

TABLE I

S. aureus - Coding regions containing known sequences

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	percent ident	HSP nt length	ORF nt length
4695	1	313	158	gt L14017	Staphylococcus aureus methicillin-resistance protein (mecR) gene and unknown ORF, complete cds	75	155	156
4703	1	1	153	emb X58434 SAPD	S.aureus pdhB, pdhC and pdhD genes for pyruvate decarboxylase, dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase	98	103	153

TABLE I

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
20	6	5089	4679	gi 511839	ORF1 [Staphylococcus bacteriophage phi 11]	100	100	411
149	3	2032	1577	pir 849703 n497	int gene activator RlnA - bacteriophage phi 11	100	100	456
149	5	2109	1912	gi 166161	Bacteriophage phi-11 int gene activator [Staphylococcus acteriophage phi 11]	100	100	198
349	2	538	409	gi 166159	integrase (int) [Staphylococcus bacteriophage phi 11]	100	100	130
398	1	1372	707	gi 166159	integrase (int) [Staphylococcus bacteriophage phi 11]	100	99	666
398	2	783	1001	gi 455128	exciolnase (xla) [Staphylococcus bacteriophage phi 11]	100	100	219
502	4	1914	1744	gi 1204912	H influenzae predicted coding region H10660 [Haemophilus influenzae]	100	71	171
849	1	2	262	gi 1373002	polyprotein [bean common mosaic virus]	100	46	261
1349	1	277	140	gi 143359	protein synthesis initiation factor 2 (infB) [Bacillus subtilis] gi 49319	100	82	138
2880	1	21	308	gi 862933	protein kinase C inhibitor-1 [Homo sapiens]	100	98	288
3085	1	428	216	gi 1354211	PET112-like protein [Bacillus subtilis]	100	100	213
4168	2	571	398	gi 1354211	PET112-like protein [Bacillus subtilis]	100	100	174
331	1	2	247	gi 426473	nusG gene product [Staphylococcus carnosus]	98	95	246
207	2	1272	1463	gi 460259	enolase [Bacillus subtilis]	97	90	192
331	2	395	850	gi 581638	lil1 protein [Staphylococcus carnosus]	97	73	456
366	1	39	215	gi 166161	Bacteriophage phi-11 int gene activator [Staphylococcus acteriophage phi 11]	97	95	177
680	3	718	936	gi 426473	nusA gene product [Staphylococcus carnosus]	97	97	219
1578	1	284	144	gi 1339950	large subunit of NAMH-dependent glutamate synthase [Plectonoma boryanum]	97	79	141
157	1	321	518	gi 1022726	unknown [Staphylococcus haemolyticus]	96	88	198
205	33	16470	16147	gi 1165302	S10 [Bacillus subtilis]	96	91	324
3919	1	48	401	gi 971784	Cip-like ATP-dependent protease binding subunit [Bos taurus]	96	81	354
4133	1	830	417	gi 1022726	unknown [Staphylococcus haemolyticus]	96	84	414
4168	1	708	355	gi 1354211	PET112-like protein [Bacillus subtilis]	96	95	354
4207	1	312	157	gi 602031	similar to triethyamine DH [Mycoplasma capricolum] pir 549950 549950 probable triethyamine dehydrogenase (EC 5.99.7) - Mycoplasma capricolum (SQC3) (fragment)	96	86	156

TABLE 2

TABLE 2

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4227	2	152	331	gi 871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	96	81	180
4416	1	570	286	gi 1022226	unknown [Staphylococcus haemolyticus]	96	84	285
22	1	858	430	gi 511070	UreO [Staphylococcus xylosus]	93	88	429
22	7	4362	4036	gi 581787	urease gamma subunit [Staphylococcus xylosus]	93	79	327
82	6	8794	9114	pir JC0008 JC00	ribosomal protein S7 - Bacillus stearothermophilus	95	83	321
154	9	9280	7838	gi 1334211	PER112-like protein (Bacillus subtilis)	95	92	1443
186	3	2798	2055	gi 1514656	serine O-acetyltransferase [Staphylococcus xylosus]	95	87	744
205	5	4406	6014	gi 142462	ribosomal protein S11 [Bacillus subtilis]	95	85	393
205	7	5017	4793	gi 142459	Initiation factor I [Bacillus subtilis]	95	84	225
205	21	11365	10991	gi 1044974	ribosomal protein L14 [Bacillus subtilis]	95	93	375
259	5	7288	6644	sp P47995 VSEA	HYPOTHEETICAL PROTEIN IN SECA 3-REGION [ORFI] (FRAGMENT)	95	85	645
302	3	795	1097	gi 40186	homologous to E.coli ribosomal protein L27 [Bacillus subtilis] 140592 L27 ribosomal protein [Bacillus subtilis] r[C21895]C21895 ribosomal protein L27 - Bacillus subtilis PUS637 BL27_DACSU 50S RIBOSOMAL PROTEIN L27 (BL30) (BL24). 140175 L24 gene prod	95	89	303
310	1	579	1523	gi 1177684	chorismate mutase [Staphylococcus xylosus]	95	92	945
414	1	2	163	pir C48396 C483	ribosomal protein L34 - Bacillus stearothermophilus	95	90	162
4185	2	125	277	gi 1276841	glutamate synthase [GOGAT] [Porphyra purpurea]	95	86	153
22	2	1028	723	gi 511069	UreF [Staphylococcus xylosus]	94	91	306
22	5	5046	3310	gi 410516	urease alpha subunit [Staphylococcus xylosus]	94	85	1737
60	4	815	1372	gi 666116	glucose kinase [Staphylococcus xylosus]	94	87	558
205	18	10012	9536	gi 1044978	ribosomal protein S8 [Bacillus subtilis]	94	78	477
326	4	3378	2542	gi 557492	dihydroxymephoic acid (DHNA) synthetase [Bacillus subtilis] g 143186 dihydroxynaphthoic acid (DHNA) synthetase [Bacillus ubtilis]	94	85	837
414	3	737	955	gi 467386	chlorphen and furan oxidation [Bacillus subtilis]	94	77	219
426	3	2260	1823	gi 1263908	putative [Staphylococcus epidermidis]	94	87	438
534	1	2	355	gi 633650	enzyme II(mannitol) [Staphylococcus carnosus]	94	84	354
1017	1	2	229	gi 149435	putative [Lactococcus lactis]	94	73	228
3098	1	330	184	gi 413952	ipa-28d gene product [Bacillus subtilis]	94	50	147

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
332	1	630	316	gi 1022725	unknown [Staphylococcus haemolyticus]	94	84	315
42	5	2089	2259	pir A48396 A483	ribosomal protein L33 - Bacillus stearothermophilus	93	81	171
101	2	1745	1303	gi 155345	arsenic efflux pump protein [Pseudomonas aeruginosa]	93	82	363
205	24	12227	11865	ep P14577 RL16	50S RIBOSOMAL PROTEIN L16	93	83	363
259	4	8291	5673	gi 499335	secA protein [Staphylococcus carnosus]	93	85	2619
275	1	2226	1114	gi 633650	enzyme II (mannitol) [Staphylococcus carnosus]	93	86	1113
444	6	6207	5733	gi 1022726	unknown [Staphylococcus haemolyticus]	93	81	435
491	1	152	622	gi 46912	ribosomal protein L13 [Staphylococcus carnosus]	93	88	471
607	6	1674	2033	gi 1022726	unknown [Staphylococcus haemolyticus]	93	81	360
653	1	973	488	gi 580890	translation initiation factor IF3 (AA 1-172) [Bacillus thuringiensis]	93	77	486
1864	1	3	194	gi 306553	ribosomal protein small subunit (Homo sapiens)	93	93	192
2997	1	28	300	gi 143390	carbamyl phosphate synthetase [Bacillus subtilis]	93	82	273
3233	2	907	596	gi 1022725	unknown [Staphylococcus haemolyticus]	93	84	312
3781	2	794	621	gi 1022725	unknown [Staphylococcus haemolyticus]	93	88	174
16	1	3	374	gi 142781	putative cytoplasmic protein; putative [Bacillus subtilis] ep P37954 DVRB_BACSU EXCINUCLEASE ABC SUBUNIT B (DNA PROTEIN) FRAGMENT1	92	83	372
31	7	5915	6124	gi 1136430	KIAA0185 protein (Homo sapiens)	92	46	210
56	19	26483	27391	gi 467401	unknown [Bacillus subtilis]	92	80	909
69	6	5882	6130	gi 530200	ltrophoblastin [Ovis aries]	92	53	249
145	3	2568	2038	gi 1022725	unknown [Staphylococcus haemolyticus]	92	80	531
171	3	2760	2362	gi 517475	p-aminic acid transaminase [Staphylococcus haemolyticus]	92	86	399
205	12	7495	6962	gi 49189	secY gene product [Staphylococcus carnosus]	92	85	534
205	19	10812	10255	gi 1044976	ribosomal protein L5 [Bacillus subtilis]	92	82	558
219	1	710	357	gi 1303812	YqeV [Bacillus subtilis]	92	88	354
344	3	1575	1805	gi 1405474	CspC protein [Bacillus cereus]	92	85	231
699	1	20	361	gi 1413999	lipa-75d gene product [Bacillus subtilis]	92	81	342
1343	1	2	160	pir A45634 A454	ribosomal protein L19 - Bacillus stearothermophilus	92	84	159

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1958	1	524	264	gi 407908	teflact (Staphylococcus xyloosus)	92	80	261
357R	2	718	386	gi 1339950	large subunit of NADH-dependent glutamate synthase (Plectonema boryanum)	92	78	333
358S	1	644	324	gi 1339950	large subunit of NADH-dependent glutamate synthase (Plectonema boryanum)	92	81	321
3640	1	4	402	gi 1022726	unknown (Staphylococcus haemolyticus)	92	81	399
4362	1	14	178	gi 450688	hcdH gene of Ecoprr1 gene product (Escherichia coli) pif S18437 S18437 hcdH protein - Escherichia coli pif S09829 S09829 hypothetical protein A - Escherichia coli (S09 40-520)	92	76	165
4446	1	358	182	gi 1022725	unknown (Staphylococcus haemolyticus)	92	82	177
4549	1	462	232	gi 1022726	unknown (Staphylococcus haemolyticus)	92	80	231
4626	1	3	224	gi 1022725	unknown (Staphylococcus haemolyticus)	92	84	222
2	4	3980	4531	gi 535349	CodM (Bacillus subtilis)	91	74	552
28	1	2	1126	gi 1001376	hypothetical protein (Synecocystis sp.)	91	78	1125
60	5	1354	1701	gi 1226043	lorf2 downstream of glucose kinase (Staphylococcus xyloosus)	91	80	348
101	1	1989	1036	gi 150728	arsenic efflux pump protein (Plasmid p1258)	91	80	954
187	2	412	1194	gi 142559	ATP synthase alpha subunit (Bacillus megaterium)	91	79	783
205	22	11579	11298	gi 40149	S17 protein (AA 1-87) (Bacillus subtilis)	91	83	282
206	7	8184	10262	gi 1072418	glcA gene product (Staphylococcus carnosus)	91	83	2079
306	2	3885	2326	gi 143012	GMP synthetase (Bacillus subtilis)	91	78	1560
306	3	5119	3826	gi 467399	UMP dehydrogenase (Bacillus subtilis)	91	79	1494
310	3	2194	3207	gi 1177685	ccpA gene product (Staphylococcus xyloosus)	91	81	1014
343	4	2974	3150	gi 1949974	sucrose repressor (Staphylococcus xyloosus)	91	82	177
480	3	1606	3042	gi 433991	ATP synthase subunit beta (Bacillus subtilis)	91	85	1437
536	3	2026	1280	gi 143366	adenylosuccinate lyase (Pur-B) (Bacillus subtilis) pif C19326 WZBDS adenylsuccinate lyase (EC 4.3.2.2) - Bacillus subtilis	91	79	747
552	1	1064	615	gi 297874	fructose-bisphosphate aldolase (Staphylococcus carnosus) pif A49943 A49943 fructose-bisphosphate aldolase (EC 4.1.2.13) - Staphylococcus carnosus (strain TM300)	91	79	450
637	1	1	1536	gi 143597	CTP synthetase (Bacillus subtilis)	91	79	1536
859	1	21	359	gi 385178	unknown (Bacillus subtilis)	91	66	339

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1327	1	339	530	gi 496558	lorK (Bacillus subtilis)	91	71	192
2515	1	466	275	gi 511070	ureC (Staphylococcus xylosum)	91	85	192
2594	1	2	202	gi 146824	beta-cyathionase (Escherichia coli)	91	75	201
3764	1	847	425	gi 1022725	unknown (Staphylococcus haemolyticus)	91	78	423
4011	1	127	495	gi 1022726	unknown (Staphylococcus haemolyticus)	91	79	369
4227	1	1	177	gi 296464	ATPase (Lactococcus lactis)	91	66	177
42	3	815	1033	gi 520401	catalase (Haemophilus influenzae)	90	86	219
51	8	3717	4607	gi 580899	OppF gene product (Bacillus subtilis)	90	74	891
129	3	5317	4001	gi 1146206	glutamate dehydrogenase (Bacillus subtilis)	90	76	1317
184	17	16628	16933	sp P05766 RS15_305	RIBOSOMAL PROTEIN S15 (BS18)	90	74	306
171	5	2983	2819	gi 517475	D-amino acid transaminase (Staphylococcus haemolyticus)	90	78	165
205	4	4497	3550	gi 1142463	RNA polymerase alpha-core-subunit (Bacillus subtilis)	90	76	948
205	6	4748	4410	gi 1044989	ribosomal protein S13 (Bacillus subtilis)	90	73	339
205	10	7165	6404	gi 49189	secY gene product (Staphylococcus carnosus)	90	81	762
205	11	6645	6472	gi 49189	secY gene product (Staphylococcus carnosus)	90	78	174
205	27	13692	13345	gi 786157	Ribosomal Protein S19 (Bacillus subtilis)	90	79	348
205	31	15858	15496	gi 1165303	L3 (Bacillus subtilis)	90	79	363
260	5	7023	5723	gi 1161380	IcaA (Staphylococcus epidermidis)	90	78	1251
299	6	3378	3947	gi 467440	phosphoribosylpyrophosphate synthetase (Bacillus subtilis) gi 40218 PRPP synthetase (AA 1-317) (Bacillus subtilis)	90	78	570
320	2	1025	1717	gi 312443	carbamoyl-phosphate synthase (glutamine-hydrolysing) (Bacillus aldolyticus)	90	75	693
330	4	1581	1769	gi 986963	beta-tubulin (Sporidiobolus pararoseus)	90	80	189
369	1	954	523	pir S34762 S347	L-serine dehydratase beta chain - Clostridium sp.	90	77	432
557	1	3	188	gi 1511569	M. Jannaschii predicted coding region MJ1624 (Methanococcus jannaschii)	90	54	186
663	2	667	1200	gi 143786	tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Bacillus subtilis) p CJ70481 YND5 tryptophan--tRNA ligase (EC 6.1.1.2) - Bacillus ubtilla	90	73	534
717	1	1	261	gi 143065	hubst (Bacillus stearothermophilus)	90	79	261
745	4	1059	865	gi 1205433	M. influenzae predicted coding region H11190 (Haemophilus influenzae)	90	81	195

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
1007	1	386	565	gi1143366	[adenylosuccinate lyase (Pur-8) (Bacillus subtilis)] pir[C99326]W28505	90	77	180
					[adenylosuccinate lyase (EC 4.3.2.2) - Bacillus subtilis]			
1054	1	579	331	gi11033122	[ORF_729 (Escherichia coli)]	90	50	249
1156	1	117	707	gi1147776	[CipP (Bacillus subtilis)]	90	80	591
1180	1	408	205	gi11377831	[unknown (Bacillus subtilis)]	90	74	204
1253	1	1	462	gi140046	[phosphoglucose isomerase A (AA 1-449) (Bacillus stearothermophilus)]	90	75	462
					[ir S15936 NUBSA glucose-6-phosphate isomerase (EC 5.3.1.9) A - cillus stearothermophilus]			
2951	1	3	269	gi1144816	[formyltetrahydrofolate synthetase (FTHFS) (ctg start codon) (EC 3.4.3)]	90	76	267
					[Moorella thermoacetica]			
3140	1	327	166	gi11070014	[protein-dependent (Bacillus subtilis)]	90	52	162
4594	1	3	233	gi11871784	[Cip-like ATP-dependent protease binding subunit (Bos taurus)]	90	76	231
87	1	1028	1750	gi1467327	[unknown (Bacillus subtilis)]	89	75	723
112	1	2	505	gi1153741	[ATP-binding protein (Streptococcus mutans)]	89	77	504
118	1	120	398	gi11303804	[VqeO (Bacillus subtilis)]	89	75	279
129	4	3545	3757	gi1460257	[triose phosphate isomerase (Bacillus subtilis)]	89	84	213
164	12	11667	12755	gi139954	[IP2 (aa 1-741) (Bacillus stearothermophilus)]	89	80	1089
205	13	7875	7405	gi1216338	[ORF for L15 ribosomal protein (Bacillus subtilis)]	89	76	471
205	32	16152	15823	gi11165303	[L3 (Bacillus subtilis)]	89	80	330
270	3	2407	2207	gi1419021419	[arsenate reductase (EC 1.-.-.-) - Staphylococcus xylosus plasmid pSX267]	89	81	201
395	2	157	672	gi1520574	[glutamate racemase (Staphylococcus haemolyticus)]	89	80	516
494	1	3	839	gi1396259	[protease (Staphylococcus epidermidis)]	89	77	837
510	1	1	444	gi140046	[phosphoglucose isomerase A (AA 1-449) (Bacillus stearothermophilus)]	89	74	444
					[ir S15936 NUBSA glucose-6-phosphate isomerase (EC 5.3.1.9) A - cillus stearothermophilus]			
615	1	2124	1210	gi11303812	[VqeV (Bacillus subtilis)]	89	74	915
841	1	18	341	gi11165303	[L3 (Bacillus subtilis)]	89	80	334
1111	1	352	813	gi147146	[thermonuclease (Staphylococcus intermedius)]	89	70	462
1875	1	2	256	gi1205108	[ATP-dependent protease binding subunit (Haemophilus influenzae)]	89	82	255
2963	1	11	367	gi1467458	[cell division protein (Bacillus subtilis)]	89	83	357

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3020	1	90	362	gi 1239988	hypothetical protein [Bacillus subtilis]	89	66	273
3565	1	2	400	gi 1256635	dihydroxy-acid dehydratase [Bacillus subtilis]	89	75	399
3586	1	105	314	gi 580832	ATP synthase subunit gamma [Bacillus subtilis]	89	82	210
3629	1	794	399	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	89	78	396
3688	1	2	400	gi 1146206	glutamate dehydrogenase [Bacillus subtilis]	89	75	399
3699	1	794	399	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	89	75	396
4016	1	428	216	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	89	71	213
4177	1	471	301	gi 149426	putative [Lactococcus lactis]	89	76	171
4436	1	601	302	gi 1022725	unknown [Staphylococcus haemolyticus]	89	80	300
4635	1	320	162	gi 1022725	unknown [Staphylococcus haemolyticus]	89	73	159
2	2	1330	2676	gi 520754	putative [Bacillus subtilis]	88	76	1347
42	2	468	848	sp P42321 CATA_	CATALASE (EC 1.11.1.6)	88	76	381
53	5	6389	4722	gi 474177	alpha-D-1,4-glucosidase [Staphylococcus xylosum]	88	80	1668
56	16	18018	18617	gi 467411	recombination protein [Bacillus subtilis]	88	77	600
60	3	376	843	gi 666116	glucose kinase [Staphylococcus xylosum]	88	77	468
70	2	1583	1245	gi 44095	replication initiator protein [Listeria monocytogenes]	88	74	339
82	8	11514	12719	pir A60663 A606	translation elongation factor Tu - Bacillus subtilis	88	79	1206
103	7	4179	4391	gi 167181	serine/threonine kinase receptor [Brassica napus]	88	77	213
114	8	7732	8232	gi 1022726	unknown [Staphylococcus haemolyticus]	88	72	501
118	2	308	2011	gi 1303804	YqeQ [Bacillus subtilis]	88	77	1704
141	3	657	1136	gi 1405446	transketolase [Bacillus subtilis]	88	72	480
168	7	5871	6116	gi 1118002	dihydropterate synthase [Staphylococcus haemolyticus]	88	78	246
165	3	1428	2231	gi 40053	phenylalanyl-tRNA synthetase alpha subunit [Bacillus subtilis] tr S1730 YP88A phenylalanine--tRNA ligase (EC 6.1.1.20) alpha ein - Bacillus subtilis	88	80	804
205	28	15027	14185	gi 1165306	L2 [Bacillus subtilis]	88	82	843
225	1	1569	898	gi 1303840	YqfS [Bacillus subtilis]	88	70	672
235	1	2	1975	gi 452309	valyl-tRNA synthetase [Bacillus subtilis]	88	76	1974

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
339	3	2060	1566	gi1118002	[dihydropterate synthase (Staphylococcus haemolyticus)]	88	73	495
443	4	4325	2928	gi1558559	[pyrimidine nucleoside phosphorylase (Bacillus subtilis)]	88	73	1398
532	1	3	419	gi141797	[valyl-tRNA synthetase (Bacillus stearothermophilus) sp P19311 SV_BACST VALYL-TRNA SYNTHETASE (EC 6.1.1.9) VALINE--TRNA LIGASE) (VALASE)]	88	78	417
534	3	2504	2968	gi153049	[mannitol-specific enzyme-III (Staphylococcus carnosus) pir JQ0088 JQ0088 phosphotransferase system enzyme II (EC 7.1.69), mannitol-specific factor III - Staphylococcus carnosus sp P17874 PTMA_STACA PTS SYSTEM, MANNITOL-SPECIFIC IIA COMPONENT EIIA-MTU) I	88	82	465
705	2	584	399	gi1710018	[nitrite reductase (nirB) (Bacillus subtilis)]	88	70	185
1000	2	1824	1309	gi11022726	[unknown (Staphylococcus haemolyticus)]	88	78	516
1299	1	587	324	gi1401786	[phosphomannomutase (Mycoplasma pirum)]	88	55	264
1341	2	170	400	gi139963	[ribosomal protein L20 (AA 1-119) (Bacillus stearothermophilus) ir S05348 RS920 ribosomal protein L20 - Bacillus stearothermophilus	88	82	231
1386	1	41	214	pir B47154 B471	[signal recognition particle 54K chain homolog Ffh - Bacillus subtilis]	88	71	174
1386	2	183	533	pir B47154 B471	[signal recognition particle 54K chain homolog Ffh - Bacillus subtilis]	88	73	351
2949	1	704	399	gi1535350	[CodX (Bacillus subtilis)]	88	73	306
2984	1	5	169	gi1218277	[O-acetylserine(thiol) lyase (Spinacia oleracea)]	88	70	165
3035	1	1	138	gi1493083	[dihydroxyacetone kinase (Citrobacter freundii)]	88	67	138
3089	1	3	152	gi1606055	[ORP_F746 (Escherichia coli)]	88	88	150
3917	1	817	410	gi143378	[pyruvate decarboxylase (E-1) beta subunit (Bacillus subtilis) gi11377836 pyruvate decarboxylase E-1 beta subunit (Bacillus subtilis)]	88	77	408
4199	1	680	342	gi11405454	[aconitase (Bacillus subtilis)]	88	82	339
4201	1	734	369	gi1515938	[glutamate synthase (ferredoxin) (Synecocystis sp.) pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	88	84	366
4274	1	1	336	gi1515938	[glutamate synthase (ferredoxin) (Synecocystis sp.) pir S46957 S46957 glutamate synthase (ferredoxin) (EC 1.4.7.1) - ynechocystis sp.	88	84	336
4308	1	794	399	gi1146206	[glutamate dehydrogenase (Bacillus subtilis)]	88	71	396
2	5	4570	6000	gi1535350	[CodX (Bacillus subtilis)]	87	70	1431
52	8	6781	6482	gi11064791	[function unknown (Bacillus subtilis)]	87	66	300

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
73	3	1584	2480	gi142992	glycerol kinase (gpk) [EC 2.7.1.30] (Bacillus subtilis) pir[B45868]B45868 glycerol kinase (EC 2.7.1.30) Bacillus subtilis sp P18157 GPK_BACSU GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL -PHOSPHOTRANSFERASE) (GLYCEROKINASE) (GK)	87	72	897
98	12	8813	9100	gi1467433	unknown [Bacillus subtilis]	87	62	288
124	4	4265	2988	gi1556886	serine hydroxymethyltransferase [Bacillus subtilis] pir[S49363]S49363 serine hydroxymethyltransferase - Bacillus ubtilla	87	77	1278
124	6	4457	4032	gi1556883	unknown [Bacillus subtilis]	87	66	426
148	5	3741	4559	gi1467460	unknown [Bacillus subtilis]	87	70	819
164	13	12710	13810	gi139954	lip2 (aa 1-741) [Bacillus stearothermophilus]	87	72	1101
177	2	1104	2126	gi1467385	unknown [Bacillus subtilis]	87	78	1023
199	1	1982	1158	gi143527	iron-sulfur protein [Bacillus subtilis]	87	77	925
199	2	4717	2933	pir[A22763]A227	succinate dehydrogenase (EC 1.3.99.1) flavoprotein - Bacillus subtilis	87	80	1785
205	23	11782	11543	gi11044972	ribosomal protein l29 [Bacillus subtilis]	87	78	240
205	25	11275	12607	gi1165309	S3 [Bacillus subtilis]	87	75	669
222	1	2033	1107	gi11177249	rec233 gene product [Bacillus subtilis]	87	70	527
236	3	1635	1333	gi1146198	ferredoxin [Bacillus subtilis]	87	80	303
246	5	2585	2292	gi1467373	ribosomal protein S18 [Bacillus subtilis]	87	77	294
260	2	4189	3422	gi1161382	icaC [Staphylococcus epidermidis]	87	72	768
320	3	1696	2391	gi1312443	carbamoyl-phosphate synthase (glutamine-hydrolyzing) [Bacillus aldylyticus]	87	80	696
380	4	1165	1383	gi1142570	ATP synthase c subunit [Bacillus firmus]	87	80	219
414	4	900	1073	gi1467386	chlorophen and furan oxidation [Bacillus subtilis]	87	77	174
425	2	1003	794	gi11046166	lipin repressor [Mycoplasma genitalium]	87	69	210
448	1	1255	722	gi1405134	acetate kinase [Bacillus subtilis]	87	75	534
480	1	1	713	gi1142559	ATP synthase alpha subunit [Bacillus megaterium]	87	79	711
481	1	2	352	sp Q06797 RLJ_B	[50S RIBOSOMAL PROTEIN L1 (BLI)]	87	72	351
677	2	359	955	gi1460911	fructose-bisphosphate aldolase [Bacillus subtilis]	87	78	597
677	3	934	1284	gi1460911	fructose-bisphosphate aldolase [Bacillus subtilis]	87	78	351
876	1	3	452	gi1146247	asparaginyl-tRNA synthetase [Bacillus subtilis]	87	79	450

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
1376	1	426	214	gi 1085555	F46H6.4 gene product [Caenorhabditis elegans]	87	75	213
2206	1	3	374	gi 215098	excisionase [Bacteriophage 154a]	87	72	372
2938	1	3	290	gi 508979	GTP-binding protein [Bacillus subtilis]	87	69	288
3081	2	126	308	gi 467399	IMP dehydrogenase [Bacillus subtilis]	87	72	183
3535	1	3	401	gi 1405454	aconitase [Bacillus subtilis]	87	80	399
4238	1	547	275	gi 603769	HutU protein, urocanase [Bacillus subtilis]	87	73	273
4	8	10427	8736	gi 603769	HutU protein, urocanase [Bacillus subtilis]	86	72	1692
22	6	4190	3738	gi 410515	urease beta subunit [Staphylococcus xylosum]	86	73	453
54	2	2480	1572	gi 289287	UDP-glucose pyrophosphorylase [Bacillus subtilis]	86	70	909
124	3	2336	1713	gi 556887	uracil phosphoribosyltransferase [Bacillus subtilis] pir 549364 549364 uracil phosphoribosyltransferase - Bacillus subtilis	86	74	624
148	3	1349	3448	gi 467458	cell division protein [Bacillus subtilis]	86	75	2100
148	4	3638	3859	gi 467460	unknown [Bacillus subtilis]	86	73	222
152	3	1340	2086	gi 1377835	pyruvate decarboxylase E-1 alpha subunit [Bacillus subtilis]	86	75	747
164	18	17347	19467	gi 1184680	polynucleotide phosphorylase [Bacillus subtilis]	86	72	2121
180	2	554	1159	gi 143467	ribosomal protein S4 [Bacillus subtilis]	86	80	606
205	3	2966	2592	gi 142464	ribosomal protein L17 [Bacillus subtilis]	86	77	375
205	26	13364	12990	gi 40107	ribosomal protein L22 [Bacillus stearothermophilus] Ir S10612 S10612. ribosomal protein L22 - Bacillus stearothermophilus	86	75	375
246	7	3463	3140	gi 467375	ribosomal protein S6 [Bacillus subtilis]	86	70	324
299	3	1196	1540	gi 39656	spoVG gene product [Bacillus megaterium]	86	70	345
299	7	3884	4345	gi 467440	phosphoribosylpyrophosphate synthetase [Bacillus subtilis] gi 40218 PRPP synthetase (AA 1-317) [Bacillus subtilis]	86	78	462
304	5	2170	2523	gi 666983	putative ATP binding subunit [Bacillus subtilis]	86	65	354
310	2	1487	1678	gi 1177684	chorismate mutase [Staphylococcus xylosum]	86	71	192
337	5	2086	3405	gi 487434	isocitrate dehydrogenase [Bacillus subtilis]	86	78	1320
339	2	1489	1109	gi 1118003	dihydropyrimidinase [Staphylococcus haemolyticus]	86	77	381
358	2	2124	3440	gi 1146219	28.2% of identity to the Escherichia coli GTP-binding protein Era; putative [Bacillus subtilis]	86	73	1317

TABLE 2

5. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	patch gene name	% sim	% idenc	length (nt)
404	2	1015	2058	gi 1303817	YqfA [Bacillus subtilis]	86	78	1044
581	2	661	452	gi 40056	phoP gene product [Bacillus subtilis]	86	71	210
642	2	338	1075	gi 1176399	EpIF [Staphylococcus epidermidis]	86	72	738
770	1	622	347	gi 143328	phoP protein [put.]; putative [Bacillus subtilis]	86	69	276
865	1	1777	890	gi 1146247	asparaginyl-tRNA synthetase [Bacillus subtilis]	86	74	888
868	2	963	1133	gi 1002911	transmembrane protein [Saccharomyces cerevisiae]	86	59	171
904	1	1	162	gi 1303912	YqfW [Bacillus subtilis]	86	72	162
989	1	35	433	gi 1303993	YqfL [Bacillus subtilis]	86	76	399
1212	1	296	350	gi 414014	lipa-90d gene product [Bacillus subtilis]	86	70	147
1323	1	2	148	gi 40041	pyruvate dehydrogenase (lipoamide) [Bacillus stearothermophilus] ir S1078 PSSPF pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) pha chain - Bacillus stearothermophilus	86	75	147
3085	2	540	310	gi 1354211	PER112-like protein [Bacillus subtilis]	86	86	231
1847	1	1	228	gi 296464	ATPase [Lactococcus lactis]	86	63	228
4487	1	476	240	gi 1022726	unknown [Staphylococcus haemolyticus]	86	73	237
4583	1	372	187	gi 1022725	unknown [Staphylococcus haemolyticus]	86	79	186
25	5	4287	5039	gi 1502421	3-ketoacyl-acyl carrier protein reductase [Bacillus subtilis]	85	64	753
56	21	30627	29395	gi 1408507	pyridine nucleoside transport protein [Bacillus subtilis]	85	69	1231
68	2	332	1192	gi 467376	unknown [Bacillus subtilis]	85	74	861
73	2	880	1707	gi 142992	glycerol kinase (gipK) (EC 2.7.1.30) [Bacillus subtilis] pir B45868 B45868 glycerol kinase (EC 2.7.1.30) - Bacillus subtilis sp P18157 GLPK_BACSU GLYCEROL KINASE (EC 2.7.1.30) (ATP:GLYCEROL -PHOSPHOTRANSFERASE) (GLYCEROL KINASE) (GK)	85	72	828
106	4	1505	3490	gi 143766	lthrSv (EC 6.1.1.3) [Bacillus subtilis]	85	74	1986
128	2	1153	2202	gi 311924	glyceraldehyde-3-phosphate dehydrogenase [Clostridium pasteurianum] pir S14254 S14254 glyceraldehyde-3-phosphate dehydrogenase (EC 2.1.12) - Clostridium pasteurianum	85	75	1050
129	4	6466	5252	gi 1064807	ORTHONINE AMINOTRANSFERASE [Bacillus subtilis]	85	73	1215
138	6	3475	5673	gi 1072419	gicB gene product [Staphylococcus carnosus]	85	74	2199
189	1	2	169	gi 467385	unknown [Bacillus subtilis]	85	65	168

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
205	15	8624	8106	gi11044981	ribosomal protein S5 (Bacillus subtilis)	85	75	519
205	20	10928	10596	pir A02819 RSBS	ribosomal protein L34 - Bacillus stearothermophilus	85	72	333
220	6	6450	6101	gi148980	secA gene product (Bacillus subtilis)	85	66	390
231	4	4877	3159	gi11002520	MutS (Bacillus subtilis)	85	70	1719
243	9	8013	8783	gi1414011	lpa-87r gene product (Bacillus subtilis)	85	72	771
249	2	5894	3186	gi1405454	aconitase (Bacillus subtilis)	85	73	2709
302	1	140	475	gi140173	homolog of E.coli ribosomal protein L21 (Bacillus subtilis) ir S18439 S18439 Ribosomal protein L21 - Bacillus subtilis p P26908 RL21_BACSU 50S RIBOSOMAL PROTEIN L21 (BL20).	85	72	336
333	1	5445	2968	gi1442360	ClpC adenosine triphosphatase (Bacillus subtilis)	85	69	2478
364	6	6082	8196	gi1871784	Clp-like ATP-dependent protease binding subunit [Bos taurus]	85	68	2115
448	2	1992	1339	gi1405134	acetate kinase (Bacillus subtilis)	85	68	654
747	1	1251	853	gi11373157	orf-X, hypothetical protein; Method: conceptual translation supplied by author (Bacillus subtilis)	85	73	399
886	2	159	467	gi1541768	hemolysase (Yersinia enterocolitica)	85	55	309
1089	1	1208	606	pir B47154 B471	signal recognition particle 54K chain homolog Ffh - Bacillus subtilis	85	71	603
1163	1	816	409	gi1304155	(diaminopimelate decarboxylase (Bacillus methanolicus) ep p41023 DCDA_BACMT DIAMINOPIMELATE DECARBOXYLASE (EC 4.1.1.20) DAP DECARBOXYLASE)	85	62	408
1924	1	487	251	gi1215098	excisionase (Bacteriophage 154a)	85	73	237
2932	1	776	390	gi11041099	Pyruvate Kinase (Bacillus licheniformis)	85	71	387
3030	1	3	275	gi142370	pyruvate formate-lyase (AA 1-760) [Escherichia coli] ir S01788 S01788 formate C-acetyltransferase (EC 2.3.1.54) - cherichia coli	85	74	273
3111	1	595	299	gi163568	flimb deformity protein (Gallus gallus)	85	85	297
3778	1	630	316	gi1391840	beta-subunit of HDT (Pseudomonas fragi)	85	67	315
3835	1	1	387	gi1204472	type I restriction enzyme ECOR124/3 I M protein (Haemophilus influenzae)	85	56	387
4042	1	3	386	gi118178	formate acetyltransferase (Chlamydomonas reinhardtii) ir S24997 S24997 formate C-acetyltransferase (EC 2.3.1.54) - Chlamydomonas reinhardtii	85	70	384
4053	1	35	340	gi1204472	type I restriction enzyme ECOR124/3 I M protein (Haemophilus influenzae)	85	56	306
4108	1	2	181	gi1072418	glcA gene product (Staphylococcus carnosus)	85	61	180
4300	1	575	330	gi1151932	fructose enzyme II (Rhodospirillum rubrum)	85	59	246

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% Ident	Length (nt)
492	1	627	355	gi 1022725	unknown [Staphylococcus haemolyticus]	85	74	273
408	1	2	235	gi 871784	Clp-like ATP-dependent protease binding subunit (Bos taurus)	85	62	234
440	1	578	291	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	85	68	288
455	1	2	253	gi 450688	hadM gene of Ecopri gene product [Escherichia coli] pIR[S38437]S38437 hadM protein - Escherichia coli pIR[S09625]S09625 hypothetical protein A - Escherichia coli (S08 40-520)	85	52	252
461	1	481	242	gi 1256635	dihydroxy-acid dehydratase [Bacillus subtilis]	85	65	240
4	10	10061	10591	gi 46982	fosB gene product [Staphylococcus epidermidis]	84	68	531
13	2	1348	1172	gi 142450	ahrC protein [Bacillus subtilis]	84	56	177
16	4	1803	4652	gi 1277198	DNA repair protein [Deinococcus radiodurans]	84	67	2850
22	3	1535	1128	gi 511069	UreF [Staphylococcus xylosum]	84	73	408
23	7	5055	5306	gi 603320	Yer082p [Saccharomyces cerevisiae]	84	61	252
53	11	11597	11145	gi 1303940	YqjW [Bacillus subtilis]	84	68	453
53	12	14059	12770	gi 142613	branched chain alpha-keto acid dehydrogenase E2 [Bacillus subtilis]	84	71	1290
70	1	1332	982	gi 46647	ORF (repE) [Staphylococcus aureus]	84	68	351
73	4	2512	4311	gi 142993	glycerol-3-phosphate dehydrogenase (gipD) (EC 1.1.99.5) [Bacillus subtilis]	84	74	1800
94	7	4324	6096	gi 467427	methylol-tRNA synthetase [Bacillus subtilis]	84	66	1773
100	9	9501	8680	gi 1340128	ORF1 [Staphylococcus aureus]	84	78	822
117	3	1934	3208	gi 1237019	Srb [Bacillus subtilis]	84	68	1275
148	6	4720	5670	gi 467462	cysteine synthetase A [Bacillus subtilis]	84	69	951
152	4	2064	2456	gi 143377	pyruvate decarboxylase (E-1) alpha subunit [Bacillus subtilis] pIR[B36718]B36718 pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) lpha chain - Bacillus subtilis	84	70	393
169	7	3634	3861	gi 1001342	hypothetical protein [Synecocystis sp.]	84	66	228
171	4	2992	2657	gi 517475	D-amino acid transaminase [Staphylococcus haemolyticus]	84	71	336
186	6	6941	6216	gi 467475	unknown [Bacillus subtilis]	84	70	726
205	9	6261	5692	gi 216340	ORF for adenylate kinase [Bacillus subtilis]	84	71	570
224	2	915	1391	gi 288269	beta-fructofuranosidase [Staphylococcus xylosum]	84	70	477

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
251	1	92	388	gi11303790	ygef [Bacillus subtilis]	84	65	297
282	3	1526	2836	gi1143040	glutamate-1-semialdehyde 2,1-aminotransferase [Bacillus subtilis] pir1042728 042728 glutamate-1-semialdehyde 2,1-aminomutase [EC 4.3.8] - Bacillus subtilis	84	75	1311
307	5	3138	2859	gi11070014	protein-dependent [Bacillus subtilis]	84	62	180
320	4	2343	4229	gi1143390	carbamyl phosphate synthetase [Bacillus subtilis]	84	70	1887
372	1	3	286	gi11022725	unknown [Staphylococcus haemolyticus]	84	70	294
413	2	2201	1341	gi11256146	ybbQ [Bacillus subtilis]	84	65	861
439	1	3	392	gi11046173	osmotically inducible protein [Mycoplasma genitalium]	84	53	390
461	3	1362	2270	gi1140211	threonine synthase [thrC] (AA 1-352) [Bacillus subtilis] ir1A25364 A25364 threonine synthase [EC 4.2.99.2] - Bacillus btills	84	69	909
487	1	3	299	gi1144531	integrin-like protein alpha intip [Candida albicans]	84	46	297
491	2	624	905	pir1508564 R3NS	ribosomal protein S9 - Bacillus stearothermophilus	84	69	282
491	3	816	1031	pir1508564 R3NS	ribosomal protein S9 - Bacillus stearothermophilus	84	77	198
548	1	3	341	gi1141231	inacII peptidase [Bacillus caldolyticus]	84	74	339
728	2	2701	1748	gi11912445	DNA polymerase [Bacillus caldolyticus]	84	68	954
769	1	3	257	gi11510953	cobalamin biosynthesis protein N [Methanococcus jannaschii]	84	38	255
954	1	308	156	gi11405454	aconitase [Bacillus subtilis]	84	57	151
957	1	3	395	gi1143402	recombination protein (itg start codon) [Bacillus subtilis] gi11303923 RecN [Bacillus subtilis]	84	68	393
975	1	3	452	gi1185934	ClpB [Synechococcus sp.]	84	70	450
1585	1	3	257	gi11510140	ligandopeptidase F [Lactococcus lactis]	84	56	255
2954	1	3	323	gi11603769	HutU protein, urocanase [Bacillus subtilis]	84	73	321
2996	1	650	348	gi1181178	formate acetyltransferase [Chlamydomonas reinhardtii] ir1S24997 S24997 formate C-acetyltransferase [EC 2.3.1.34] - Chlamydomonas reinhardtii	84	65	303
3766	1	737	375	gi11517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	84	72	363
4022	1	2	169	gi1146206	glutamate dehydrogenase [Bacillus subtilis]	84	54	168
4058	1	620	312	gi1151932	fructose enzyme II [Rhodospirillum rubrum]	84	71	309
4108	2	106	351	gi11072418	glcA gene product [Staphylococcus carnosus]	84	77	246

TABLE 2

: aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
4183	1	3	308	gi 603769	[hutU protein, urecanase (Bacillus subtilis)]	84	72	306
4726	1	55	234	gi 146208	[glutamate synthase large subunit (EC 2.6.1.53) (Escherichia coli) pifA29617A29617 glutamate synthase (NADPH) (EC 1.4.1.13) large hain - Escherichia coli]	84	73	180
22	4	2043	1576	gi 393297	[urease accessory protein (Bacillus sp.)]	83	64	468
53	13	14722	13745	gi 142612	[branched chain alpha-keto acid dehydrogenase E1-beta (Bacillus subtilis)]	83	68	978
57	16	13357	12872	gi 143132	[lactate dehydrogenase (AC 1.1.1.27) (Bacillus caldolyticus) pifB29704B29704 L-lactate dehydrogenase (EC 1.1.1.27) - Bacillus caldolyticus]	83	66	486
66	3	3119	2274	gi 1103894	[yqhM (Bacillus subtilis)]	83	63	846
66	5	6118	4643	gi 1212730	[yqhK (Bacillus subtilis)]	83	68	1476
70	3	1864	1523	gi 44095	[replication initiator protein (Listeria monocytogenes)]	83	73	342
90	1	377	1429	gi 155571	[alcohol dehydrogenase 1 (adhA) (EC 1.1.1.1) (Zymomonas mobilis) pifA35260A35260 alcohol dehydrogenase (EC 1.1.1.1) I - Zymomonas mobilis]	83	70	1053
95	2	708	2162	gi 506381	[phospho-beta-glucosidase (Bacillus subtilis)]	83	70	1455
137	1	68	694	gi 467191	[initiation protein of replicon (Bacillus subtilis)]	83	77	677
140	4	3209	2742	gi 1634107	[kdsB (Escherichia coli)]	83	65	468
142	3	3468	2989	gi 1212776	[luminase synthase (b-subunit) (Bacillus amyloquelens)]	83	69	480
161	12	5749	6686	gi 903307	[ORF75 (Bacillus subtilis)]	83	64	948
164	9	9880	11070	gi 49316	[ORF2 gene product (Bacillus subtilis)]	83	66	1191
164	14	14148	14546	gi 580902	[ORF6 gene product (Bacillus subtilis)]	83	60	399
170	2	3144	2467	gi 520844	[orf4 (Bacillus subtilis)]	83	64	678
186	2	2029	1370	gi 289284	[cysteineyl-tRNA synthetase (Bacillus subtilis)]	83	72	660
205	14	7822	7607	gi 216337	[ORF for L30 ribosomal protein (Bacillus subtilis)]	83	74	216
237	6	3683	4540	gi 1510488	[limdoleglycerol-phosphate synthase (cyclase) (Methanococcus jannaschii)]	83	60	858
301	1	985	638	gi 467419	[unknown (Bacillus subtilis)]	83	65	348
302	4	1421	2743	gi 508979	[GTP-binding protein (Bacillus subtilis)]	83	68	1323
321	4	3933	3571	gi 39844	[fumarate (citric) (aa 1-462) (Bacillus subtilis)]	83	68	363
367	1	2	352	gi 1039479	[ORF1 (Lactococcus lactis)]	83	54	351

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
387	1	3	662	gi 806281	DNA polymerase I [Bacillus stearothermophilus]	83	70	660
527	2	916	1566	gi 396259	protease [Staphylococcus epidermidis]	83	67	651
533	1	355	179	gi 142455	alanine dehydrogenase (EC 1.4.1.1) [Bacillus stearothermophilus] pir B94261 B94261 alanine dehydrogenase (EC 1.4.1.1) - Bacillus stearothermophilus	83	66	177
536	4	1617	1438	gi 143366	adenylosuccinate lyase (Pur-B) [Bacillus subtilis]; tr C39336 W2BS05 adenylosuccinate lyase (EC 4.3.2.2) - Bacillus subtilis	83	67	180
652	1	2	859	gi 520753	DNA topoisomerase I [Bacillus subtilis]	83	72	858
774	2	200	361	gi 1522665	M. jannaschii predicted coding region MJEC128 [Methanococcus jannaschii]	83	58	162
897	1	120	296	gi 1064807	ORTHONINE AMINOTRANSFERASE [Bacillus subtilis]	83	76	177
1213	1	3	491	gi 289288	lexA [Bacillus subtilis]	83	67	489
2529	1	296	150	gi 143786	cryptophany- tRNA synthetase (EC 6.1.1.2) [Bacillus subtilis] pir J70481 J70481 tRNA synthetase (EC 6.1.1.2) - Bacillus subtilis	83	69	147
2973	1	649	326	gi 1109687	Pro2 [Bacillus subtilis]	83	58	324
1009	1	728	366	gi 182532	ORF_0294 [Escherichia coli]	83	65	363
3035	2	45	305	gi 950062	hypothetical yeast protein 1 [Mycoplasma capricolum] pir 848578 848578 hypothetical protein - Mycoplasma capricolum SC3 (fragment)	83	59	261
1906	1	67	309	gi 1153197	thioredoxin reductase [Escherichia coli]	83	61	243
445R	1	540	271	gi 397526	clumping factor [Staphylococcus aureus]	83	78	270
4570	1	444	223	gi 1022726	unknown [Staphylococcus haemolyticus]	83	74	222
4654	1	97	261	gi 1072419	glcB gene product [Staphylococcus carnosus]	83	79	165
16	2	295	1191	gi 153854	luva402 protein [Streptococcus pneumoniae]	82	67	897
16	3	1193	1798	gi 153854	luva402 protein [Streptococcus pneumoniae]	82	70	606
38	12	9644	8724	gi 1204400	N-acetylneuraminate lyase [Haemophilus influenzae]	82	58	921
42	4	988	2019	gi 841192	catalase [Bacteroides fragilis]	82	70	1032
51	6	2590	3489	gi 143607	sporulation protein [Bacillus subtilis]	82	69	900
56	11	12270	11925	gi 39431	oligo-1,6-glucosidase [Bacillus cereus]	82	60	1656
56	15	17673	18014	gi 467410	unknown [Bacillus subtilis]	82	66	342
61	2	881	3313	gi 143148	transfer RNA-Leu synthetase [Bacillus subtilis]	82	70	2433

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
82	7	9162	11318	gi 48240	elongation factor G (AA 1-691) (Thermus aquaticus thermophilus) trlS1928 EFWG translation elongation factor G - Thermus aquaticus p P13551 EFG_THETH ELONGATION FACTOR G (EF-G).	82	64	2157
85	2	5470	3260	gi 143369	phosphoribosylformyl glycylamide synthetase II (pur-Q) (Bacillus subtilis)	82	66	2211
102	6	3662	5380	gi 1256635	dihydroxy-acid dehydratase (Bacillus subtilis)	82	65	1719
117	4	3282	3493	pir A47154 A471	orf1 5' of fth - Bacillus subtilis	82	53	252
128	6	4377	5933	gi 460258	phosphoglycerate mutase (Bacillus subtilis)	82	66	1557
129	2	1229	2182	gi 403373	glycerophosphoryl diester phosphodiesterase (Bacillus subtilis) pir S37251 S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis	82	52	954
170	1	2	1441	gi 1377831	unknown (Bacillus subtilis)	82	67	1440
177	1	3	1094	gi 467386	thiophen and furan oxidation (Bacillus subtilis)	82	65	1092
184	4	3572	4039	gi 153566	ORF (19K protein) (Enterococcus faecalis)	82	59	468
189	8	4455	4225	gi 1001878	Capt protein (Listeria monocytogenes)	82	73	231
206	19	32166	20707	gi 473916	lipopeptide antibiotics (iturin A (Bacillus subtilis) ap P39144 P314_BACSU LIPOPEPTIDE ANTIBIOTICS ITURIN A AND SURFACTIN BIOSYNTHESIS PROTEIN.	82	50	660
221	2	805	1722	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen (Streptococcus pyogenes)	82	63	918
223	4	3866	3651	gi 439619	(Salmonella typhimurium IS200 insertion sequence from SARA17, art1a1, gene product (Salmonella typhimurium)	82	69	216
260	3	5207	4296	gi 1161381	Icsp (Staphylococcus epidermidis)	82	61	912
315	3	4864	2855	gi 143397	quinol oxidase (Bacillus subtilis)	82	67	2010
321	10	8520	7945	gi 142981	ORF5: This ORF includes a region (aa23-103) containing a potential non-sulphur centre homologous to a region of Rhodospirillum rubrum nd Chromatium vinosum: putative (Bacillus stearothermophilus) pir PQ0299 PQ0299 hypothetical protein 5 (glcA 3' region) -	82	82	576
331	3	1055	1342	gi 436574	ribosomal protein L1 (Bacillus subtilis)	82	71	288
370	2	262	618	gi 1301793	YqeL (Bacillus subtilis)	82	59	357
404	4	3053	4024	gi 1303821	YqfE (Bacillus subtilis)	82	68	972
405	4	4840	3073	gi 1303913	YqhX (Bacillus subtilis)	82	67	1368
436	3	4096	2864	gi 149521	tryptophan synthase beta subunit (Lactococcus lactis) pir S3129 S3129 tryptophan synthase (EC 4.2.1.20) beta chain - Lactococcus lactis subsp. lactis	82	67	1233

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
441	4	3394	2573	gi1142952	glyceraldehyde-3-phosphate dehydrogenase [Bacillus teurothermophilus]	82	67	822
444	12	10415	11227	gi11204354	spore germination and vegetative growth protein [Haemophilus influenzae]	82	67	813
446	1	3	191	gi1143387	aspartate transcarbamylase [Bacillus subtilis]	82	66	189
462	3	1007	1210	gi1142521	deoxyribodipyrimidine photolyase [Bacillus subtilis] p1r[AJ7192]AJ7192 uvrB protein - Bacillus subtilis sp14951[UVRC_BACSU EXCINUCLEASE ABC SUBUNIT C]	82	64	204
537	1	1560	784	gi11853767	UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Bacillus subtilis]	82	61	777
680	2	407	700	gi11426472	secE gene product [Staphylococcus carnosus]	82	69	294
724	2	565	386	gi1143373	phosphoribosyl aminimidazole carboxy foray] ormyltransferase/inosine monophosphate cyclohydrolyase (Pur-H3J) Bacillus subtilis]	82	68	180
763	1	422	213	gi11467458	cell division protein [Bacillus subtilis]	82	35	210
818	1	564	283	gi11064787	function unknown [Bacillus subtilis]	82	69	282
858	1	175	1176	gi1143043	uroporphyrinogen decarboxylase [Bacillus subtilis] p1r[B47045]B47045 uroporphyrinogen decarboxylase (EC 4.1.1.37) - acillus subtilis	82	71	1002
895	1	3	599	gi11027507	ATP binding protein [Borrelia burgdorferi]	82	72	597
939	1	10	399	gi1143795	transfer RNA-Tyr synthetase [Bacillus subtilis]	82	60	390
961	1	1	306	gi11577647	gamma-hemolysin [Staphylococcus aureus]	82	69	306
1192	1	307	155	gi1146974	NH3-dependent NAD synthetase [Escherichia coli]	82	71	153
1317	1	49	375	gi11407908	Elisac [Staphylococcus xyloosus]	82	72	327
1341	1	1	150	gi1139962	ribosomal protein L35 (AA 1-66) [Bacillus stearothermophilus] 1r[S05347]R38S35 ribosomal protein L35 - Bacillus stearothermophilus	82	68	150
2590	2	567	349	gi1534655	ATPase subunit epsilon [Bacillus stearothermophilus] sp1P42009[ATPE_BACST ATP SYNTHASE EPSILON CHAIN (EC 3.6.1.34)]	82	47	219
3024	1	45	224	gi11467402	unknown [Bacillus subtilis]	82	64	180
3045	1	276	139	gi11467335	ribosomal protein L9 [Bacillus subtilis]	82	60	138
3045	2	558	400	gi11467335	ribosomal protein L9 [Bacillus subtilis]	82	82	159
3091	1	474	238	gi11499335	secA protein [Staphylococcus carnosus]	82	78	237
3107	1	416	210	gi1546918	orvY 3' of conk [Bacillus subtilis] 526. Peptide Partial, 140 aa] p1r[S43612]S43612 hypothetical protein Y - Bacillus subtilis sp1P40098[THRD_BACSU HYPOTHETICAL PROTEIN IN CONK 3' REGION (ORF7) FRAGMENT]	82	64	207

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
432	1	2	319	gi 42086	nitrate reductase alpha subunit [Escherichia coli] p P09152 NARO_SCOJ1 RESPIRATORY NITRATE REDUCTASE 1 ALPHA CHAIN (EC 7.99.4). (SUB 2-1247)	82	75	318
23	3	3275	2574	gi 1199573	spas [Shingomonas sp.]	81	64	702
42	1	638	321	gi 466778	lysine specific permease [Escherichia coli]	81	59	318
48	5	4051	4350	gi 1045937	H. genitalium predicted coding region MG246 [Mycoplasma genitalium]	81	62	300
51	4	1578	2579	pir S16649 S166	dcIAC protein - Bacillus subtilis	81	55	1002
53	2	354	1494	gi 1303961	YqjJ [Bacillus subtilis]	81	67	1131
53	8	9419	7971	gi 146930	6-phosphogluconate dehydrogenase [Escherichia coli]	81	66	1449
54	9	10757	10119	gi 143016	permease [Bacillus subtilis]	81	65	639
54	10	13360	11786	gi 143015	gluconate kinase [Bacillus subtilis]	81	64	1575
57	17	13983	13366	pit A25805 A258	L-lactate dehydrogenase (EC 1.1.1.27) - Bacillus subtilis	81	74	618
81	2	2708	2217	gi 1222302	IN10-related protein [Haemophilus influenzae]	81	54	492
86	1	745	374	gi 414017	jpa-91d gene product [Bacillus subtilis]	81	70	372
103	6	6438	4861	gi 971342	nitrate reductase beta subunit [Bacillus subtilis] sp P42176 NARO_BACSU NITRATE REDUCTASE BETA CHAIN (EC 1.7.99.4)	81	64	1578
120	15	10845	12338	gi 1524392	GbsA [Bacillus subtilis]	81	67	1494
128	5	3676	4413	g 143319	triose phosphate isomerase [Bacillus megaterium]	81	64	718
131	9	10308	9280	g 299163	alanine dehydrogenase [Bacillus subtilis]	81	68	1039
143	6	6088	5471	g 439619	[Salmonella typhimurium] IS200 insertion sequence from SRA17, att1a1, gene product [Salmonella typhimurium]	81	61	618
169	1	43	825	gi 897795	30S ribosomal protein (pediococcus acidilactici) sp P99668 RS2_PEDAC 30S RIBOSOMAL PROTEIN S2	81	65	783
230	1	450	226	gi 1125826	short region of weak similarity to tyrosine-protein kinase receptors in a fibronectin type III-like domain [Caenorhabditis elegans]	81	54	225
233	5	2000	2677	gi 467404	unknown [Bacillus subtilis]	81	61	678
241	2	3081	2149	gi 16510	succinate-CoA ligase (GDP-forming) [Arabidopsis thaliana] tr S0579 S0579 succinate-CoA ligase (GDP-forming) (EC 6.2.1.4) pha chain - Arabidopsis thaliana (fragment)	81	69	913
256	1	1	981	pir S09411 S094	spoIIIE protein - Bacillus subtilis	81	65	981
259	3	3752	2691	sp P28367 RF2_B	PROBABLE PEPTIDE CHAIN RELEASE FACTOR 2 (RF-2) (FRAGMENT)	81	65	1062

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
275	2	1728	3581	gi 726480	L-glutamine-O-fructose-6-phosphate amidotransferase [Bacillus subtilis]	81	68	1854
285	1	1466	715	gi 1204844	H. influenzae predicted coding region H10594 [Haemophilus influenzae]	81	63	732
296	1	99	1406	gi 467328	adenylosuccinate synthetase [Bacillus subtilis]	81	67	1308
302	9	5590	5889	gi 147485	queA [Escherichia coli]	81	64	300
317	2	1137	1176	gi 154961	resolvase [Transposon Tn917]	81	54	240
343	2	1034	1342	gi 405955	lyeod [Escherichia coli]	81	60	309
360	2	1404	2471	gi 1204570	aspartyl-tRNA synthetase [Haemophilus influenzae]	81	67	1058
364	5	6251	5706	gi 1204652	methylated-DNA--protein-cysteine methyltransferase [Haemophilus influenzae]	81	63	546
372	2	1707	1135	gi 467416	unknown [Bacillus subtilis]	81	65	573
392	1	43	603	pir S09411 S094	spoIIIE protein - Bacillus subtilis	81	65	561
404	9	5252	6154	gi 1606745	bex [Bacillus subtilis]	81	65	903
426	2	1727	1119	gi 139453	manganese superoxide dismutase [Bacillus caldotenax] ir S22053 S22053 superoxide dismutase (EC 1.15.1.1) (Mn) - Bacillus licheniformis	81	66	609
440	7	4653	5889	pir C37083 C370	hypothetical protein II (neph 3' region) - Salmonella typhimurium (fragment)	81	57	237
625	3	1105	2070	gi 1262360	protein kinase PknB [Mycobacterium leprae]	81	56	966
754	2	504	1064	gi 11303902	lyqH [Bacillus subtilis]	81	71	561
842	1	86	430	gi 1405446	transketolase [Bacillus subtilis]	81	68	345
953	1	798	400	gi 1205429	dipeptide transport ATP-binding protein [Haemophilus influenzae]	81	57	399
961	2	252	401	gi 487686	synergohymenotrophic toxin [Staphylococcus intermedius] pir S46944 S46944 synergohymenotrophic toxin - Staphylococcus intermedius	81	72	150
1035	1	1	189	gi 1046138	M. genitalium predicted coding region MG433 [Mycoplasma genitalium]	81	43	189
1280	1	670	449	gi 559164	helicase [Autographa californica nuclear polyhedrosis virus] sp P24307 V143 NPVAC HELICASE	81	43	222
3371	1	68	241	gi 1322245	mevalonate pyrophosphate decarboxylase [Rattus norvegicus]	81	62	174
3715	1	475	239	gi 537137	ORF_1388 [Escherichia coli]	81	58	237
3908	1	2	325	gi 439619	[Salmonella typhimurium] S200 insertion sequence from SARA17, att1.1, gene product [Salmonella typhimurium]	81	68	324
3940	1	3	401	gi 1296464	ATPase [Lactococcus lactis]	81	69	399

TABLE 2

3. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match reason	match gene name	% sim	% ident	length (nt)
3954	1	1	318	gi 1224069	amidase [Moraxella catarrhalis]	81	68	318
4049	1	337	170	gi 603768	putI protein, imidazole-5-propionate hydrolase [Bacillus subtilis] gi 603768 MutI protein, imidazole-5-propionate hydrolase [Bacillus subtilis]	81	68	168
4209	1	1	324	gi 401373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis] pir[S37251]S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis	81	58	324
4371	1	627	322	gi 216677	indolepyruvate decarboxylase [Enterobacter cloacae] pir[S16013]S16013 indolepyruvate decarboxylase (EC 4.1.1.-) - nterobacter cloacae	81	72	306
4387	1	19	228	gi 460689	TVG [Thermactinomyces vulgaris]	81	59	210
4391	1	581	306	gi 1524193	unknown [Mycobacterium tuberculosis]	81	67	276
4425	1	3	341	gi 143015	glucuronate kinase [Bacillus subtilis]	81	66	339
9	1	1593	847	gi 1064786	function unknown [Bacillus subtilis]	80	62	747
17	1	544	311	gi 559164	helicase [Autographa californica nuclear polyhedrosis virus] sp[P24307]V143_HPVAC HELICASE	80	40	234
45	2	1159	2448	gi 1109684	ProV [Bacillus subtilis]	80	63	1290
45	5	4032	4733	gi 1109687	ProZ [Bacillus subtilis]	80	55	702
54	8	10266	9502	gi 563952	glucuronate permease [Bacillus licheniformis]	80	62	765
62	12	8852	7545	gi 854655	Na/H antiporter system [Bacillus alcalophilus]	80	62	1308
62	14	8087	8683	gi 559713	ORF [Homo sapiens]	80	68	597
67	16	13781	14122	gi 305002	ORF_F356 [Escherichia coli]	80	65	342
70	13	11495	10296	gi 1303995	YqkN [Bacillus subtilis]	80	64	1200
98	9	6336	7130	gi 467428	unknown [Bacillus subtilis]	80	68	795
98	10	7294	7833	gi 467430	unknown [Bacillus subtilis]	80	64	540
98	11	7820	8737	gi 467431	high level kanamycin resistance [Bacillus subtilis]	80	61	918
109	16	14154	14813	gi 580875	lpa-57d gene product [Bacillus subtilis]	80	63	660
112	15	14294	16636	gi 1072361	pyruvate-formate-lyase [Clostridium pasteurianum]	80	65	2343
139	1	1448	726	gi 506699	CapC [Staphylococcus aureus]	80	58	723
139	2	2179	1488	gi 506698	CapB [Staphylococcus aureus]	80	59	732
174	4	3271	2870	gi 1146242	aspartate 1-decarboxylase [Bacillus subtilis]	80	61	402

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
177	3	2102	2842	gi 467385	unknown [Bacillus subtilis]	80	70	741
184	5	6124	5912	gi 161953	85-kDa surface antigen [Trypanosoma cruzi]	80	46	213
186	4	5368	3875	gi 289282	glutaryl-CoA synthetase [Bacillus subtilis]	80	65	1494
205	30	15796	15140	gi 40103	ribosomal protein L4 [Bacillus stearothermophilus]	80	66	657
207	1	140	1315	gi 460259	enolase [Bacillus subtilis]	80	67	1176
211	3	1078	1590	gi 410131	ORF7 [Bacillus subtilis]	80	61	513
235	2	1962	2255	gi 143797	valyl-tRNA synthetase [Bacillus stearothermophilus] op[PI1931]STV_BACST VALYL-TRNA SYNTHETASE (EC 6.1.1.9) VALINE--TRNA LIGASE) (VALURS)	80	55	294
239	1	1	1263	gi 143000	proton glutamate symport protein [Bacillus stearothermophilus] p[IS26247]IS26247 glutamate/aspartate transport protein - Bacillus stearothermophilus	80	59	1263
272	5	2724	2461	gi 709993	hypothetical protein [Bacillus subtilis]	80	54	264
301	3	1446	1111	gi 467418	unknown [Bacillus subtilis]	80	58	336
310	4	5697	4501	gi 1177686	acuC gene product [Staphylococcus xylosum]	80	67	1197
310	6	5258	7006	gi 348053	acetyl-CoA synthetase [Bacillus subtilis]	80	67	1749
310	7	7410	9133	gi 1103865	formyl-tetrahydrofolate synthetase [Streptococcus mutans]	80	67	1704
325	3	1114	1389	gi 110325	outer capsid protein [Rotavirus sp.]	80	40	276
337	1	1268	636	gi 537049	ORF_0470 [Escherichia coli]	80	55	633
374	2	929	1228	gi 1405448	YneF [Bacillus subtilis]	80	70	300
375	5	3062	2331	gi 467448	unknown [Bacillus subtilis]	80	68	270
388	1	267	587	gi 1064791	function unknown [Bacillus subtilis]	80	65	321
394	1	9	659	gi 304976	matches PS00017: ATP GTP_A and PS00301: EFACITOR_GTP: similar to longation factor G. TetM/TetO tetracycline-resistance proteins [Escherichia coli]	80	65	651
456	1	625	1263	gi 1145183	putative [Bacillus subtilis]	80	65	639
475	1	1	654	gi 288269	Beta-fructofuranosidase [Staphylococcus xylosum]	80	66	654
544	2	1449	2240	gi 529754	IspeC [Streptococcus pyogenes]	80	50	792
622	4	1623	1871	gi 1483545	unknown [Mycobacterium tuberculosis]	80	65	249
719	1	1	1257	gi 1064791	function unknown [Bacillus subtilis]	80	68	1257
739	1	107	838	gi 666983	putative ATP binding subunit [Bacillus subtilis]	80	61	732

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
745	2	581	414	gi1511600	coenzyme PQQ synthase protein III (Methanococcus jannaschii)	80	61	168
822	1	17	679	gi1410141	ORF17 [Bacillus subtilis]	80	68	663
827	2	991	836	gi1205301	leukotoxin secretion ATP-binding protein [Haemophilus influenzae]	80	54	156
1044	1	3	149	gi160632	vps2 [Marburg virus]	80	55	147
1220	2	571	413	pir161072 EPSG	gallicidin precursor - Staphylococcus gallinarum	80	74	159
2519	1	75	275	gi1147556	cbp [Escherichia coli]	80	45	201
2947	1	503	279	gi1184680	polynucleotide phosphorylase [Bacillus subtilis]	80	62	225
3120	1	2	226	gi1517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus yogenes]	80	65	225
3191	1	294	148	gi151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir144756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	80	59	147
3560	2	285	434	gi1217130	photosystem 1 core protein B [Synchococcus vulcanus]	80	70	150
3655	1	47	346	gi1415855	deoxyribose aldolase [Mycoplasma hominis]	80	56	300
3658	2	324	584	gi1551531	2-nitropropane dioxygenase [Williopsis saturnus]	80	54	261
3769	1	798	400	gi11339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	80	68	399
3781	1	692	348	gi166412	NADH-glutamate synthase [Medicago sativa]	80	62	345
3988	1	48	287	gi11204696	fructose-permease IIBC component [Haemophilus influenzae]	80	69	240
4030	1	571	287	gi11009366	respiratory nitrate reductase [Bacillus subtilis]	80	60	285
4092	1	547	275	gi1170207	orf6 [Lactobacillus sake]	80	69	273
4103	1	680	342	gi139956	IIIGlc [Bacillus subtilis]	80	65	339
4231	1	692	348	gi1289287	UDP-glucose pyrophosphorylase [Bacillus subtilis]	80	65	345
4265	1	595	299	gi1603768	HutC protein, imidazolone-5-propiolate hydrolase [Bacillus subtilis] gi1603768 HutC protein, imidazolone-5-propiolate hydrolase [Bacillus subtilis]	80	63	297
4504	1	498	250	gi11339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	80	68	249
2	6	5998	6798	gi1535351	[Cody [Bacillus subtilis]	79	63	801
4	7	8295	7037	gi1503768	HutC protein, imidazolone-5-propiolate hydrolase [Bacillus subtilis] gi1503768 HutC protein, imidazolone-5-propiolate hydrolase [Bacillus subtilis]	79	64	1245
25	6	5273	5515	pir1436728 A367	acyl carrier protein - Rhizobium meliloti	79	65	243

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
59	2	1173	1424	gi1147923	threonine dehydratase 2 (EC 4.2.1.16) [Escherichia coli]	79	75	252
60	1	1	204	gi1666115	orf1 upstream of glucose kinase [Staphylococcus xylosoe] pir[S52351]S52351 hypothetical protein 1 - Staphylococcus xylosoe	79	60	204
81	1	3002	1590	gi1466882	pps1; B1496_C2.189 [Mycobacterium leprae]	79	64	1413
85	7	7023	6505	gi1143364	phosphoribosyl aminimidazole carboxylase I (PUR-E) [Bacillus subtilis]	79	60	519
89	6	5660	4554	gi1144906	product homologous to E. coli thioresoxin reductase; J. Biol. Chem. 1988) 263:9015-9019, and to F52a protein of alkyl hydroperoxide oxidase from S. typhimurium; J. Biol. Chem. (1990) 265:10535-10540; pen reading frame A [Clostridium pasteurianum]	79	35	1107
102	11	7489	8571	gi1143093	ketol-acid reductoisomerase [Bacillus subtilis] ap137253 [LVC-BACSU KETOL-ACID REDUCTOISOMERASE (EC 1.1.1.46) ACETOHYDROXY-ACID ISOMEROREDUCTASE] (ALPHA-KETO-BETA-HYDROXYLACIL DUCTOISOMERASE)	79	64	1083
102	14	11190	12563	gi1149428	putative [Lactococcus lactis]	79	65	1374
127	9	7792	9372	gi1458688	PrfC/RF3 [Dichelobacter nodosus]	79	68	1581
139	3	2540	1983	gi1506697	[CapA [Staphylococcus aureus]	79	55	558
144	2	1644	1156	gi11498296	peptide methionine sulfoxide reductase [Streptococcus pneumoniae]	79	47	489
148	2	529	1098	gi1467457	hypoxanthine-guanine phosphoribosyltransferase [Bacillus subtilis] gi1467457 hypoxanthine-guanine phosphoribosyltransferase [Bacillus subtilis]	79	59	570
150	1	965	591	gi1755602	unknown [Bacillus subtilis]	79	61	375
176	1	1039	587	gi1297874	fructose-bisphosphate aldolase [Staphylococcus carnosus] pir149943149943 fructose-bisphosphate aldolase (EC 4.1.2.13) - taphylococcus carnosus (strain TN300)	79	65	453
186	7	7584	6874	gi1314298	[ORF5; putative Sms protein; similar to Sms proteins from Haemophilus influenzae and Escherichia coli [Listeria monocytogenes]	79	64	711
205	16	8887	8498	gi1044980	ribosomal protein L18 [Bacillus subtilis]	79	70	390
211	1	1	519	gi13303994	YqkW [Bacillus subtilis]	79	62	519
223	2	4183	2803	gi1488430	alcohol dehydrogenase 2 [Entamoeba histolytica]	79	60	1383
243	8	8915	7896	gi1580883	[ipa-8d gene product [Bacillus subtilis]	79	60	1020
279	4	3721	4329	gi1413930	[ipa-6d gene product [Bacillus subtilis]	79	59	609
300	1	11	1393	gi1403372	[glycerol 3-phosphate perasease [Bacillus subtilis]	79	62	1383
307	3	2930	1935	gi1950062	hypothetical yeast protein 1 [Mycoplasma capricolum] pir15485781548578 hypothetical protein - Mycoplasma capricolum SGC3 (fragment)	79	60	996

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
352	6	10106	8886	gi 216854	p47K [Pseudomonas chlororaphis]	79	59	1221
412	1	1153	578	gi 143177	putative [Bacillus subtilis]	79	51	576
481	3	621	1124	gi 786163	Ribosomal Protein L10 [Bacillus subtilis]	79	66	504
516	1	702	352	gi 805090	NisP [Lactococcus lactis]	79	48	351
525	2	2457	1426	gi 143371	phosphoribosyl aminimidazole synthetase (Pur-H) [Bacillus subtilis] pir H29326 AJ85CL phosphoribosylformylglycinamide cyclo-ligase EC 6.3.3.1 - Bacillus subtilis	79	61	1032
538	4	1448	2825	gi 1370207	orf6 [Lactobacillus sakei]	79	67	624
570	1	2	421	gi 476160	arginine permease substrate-binding subunit [Listeria monocytogenes]	79	61	420
645	8	2663	3241	gi 153898	transport protein [Salmonella typhimurium]	79	62	579
683	1	75	374	gi 3064795	function unknown [Bacillus subtilis]	79	82	300
816	3	4700	3987	gi 1407784	orf-1; novel antigen [Staphylococcus aureus]	79	62	714
2929	1	3	401	gi 1524397	glycine betaine transporter OpuD [Bacillus subtilis]	79	61	399
2937	1	357	202	pir 552915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	79	58	156
2940	1	768	385	gi 149429	putative [Lactococcus lactis]	79	72	384
2946	1	570	286	gi 143267	2-oxoglutarate dehydrogenase (odhA; EC 1.2.4.2) [Bacillus subtilis]	79	61	285
2999	1	3	212	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	79	59	210
3022	1	514	332	gi 450686	3-phosphoglycerate kinase [Thermotoga maritima]	79	61	183
3064	1	3	314	gi 1204436	pyruvate formate-lyase [Haemophilus influenzae]	79	60	312
3083	1	2	220	gi 1149662	hypD gene product [Clostridium perfringens]	79	56	219
3126	1	701	411	g- 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	79	55	291
3181	1	607	326	gi 1339950	large subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	79	59	282
3345	1	3	476	gi 871784	Cip-like ATP-dependent protease binding subunit [Bos taurus]	79	63	474
3718	1	536	270	pir C36889 C368	leuB protein, inactive - Lactococcus lactis subsp. lactis (strain IL1403)	79	71	287
3724	2	159	401	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	79	64	243
3836	1	608	312	gi 1524193	unknown [Mycobacterium tuberculosis]	79	65	297
3941	1	2	334	gi 415855	deoxyribose aldolase [Mycoplasma hominis]	79	54	333
4113	1	3	341	gi 143015	gluconate kinase [Bacillus subtilis]	79	63	339

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
4501	1	406	209	gi 1022726	unknown [Staphylococcus haemolyticus]	79	66	198
4612	1	2	238	gi 460689	TVG [Thermactinomyces vulgaris]	79	58	237
2	1	2	1213	gi 520753	DNA topoisomerase I [Bacillus subtilis]	78	64	1212
8	2	2266	1220	gi 216151	DNA polymerase (gene 1; tlg start codon) [Bacteriophage SP02] gi 575197 SP02 DNA polymerase (aa 1-648) [Bacteriophage SP02] p1c[A21498][DJBS2 DNA-directed RNA polymerase (EC 2.7.7.7) - phage P02]	78	72	1047
9	2	1340	1089	gi 1084787	function unknown [Bacillus subtilis]	78	57	252
32	8	6803	7702	gi 146974	NH3-dependent NAD synthetase [Escherichia coli]	78	63	900
36	4	2941	3138	gi 290503	glutamate permease [Escherichia coli]	78	53	198
53	15	17684	16221	gi 1303941	VgtV [Bacillus subtilis]	78	58	1464
57	14	10520	12067	gi 1072418	glcA gene product [Staphylococcus carnosus]	78	65	1548
66	7	6798	5812	gi 1212729	VgtV [Bacillus subtilis]	78	67	987
67	4	4029	4376	gi 468612	nika [Escherichia coli]	78	71	348
91	9	10058	10942	gi 467380	stage 0 sporulation [Bacillus subtilis]	78	50	885
102	12	8574	10130	gi 149426	putative [Lactococcus lactis]	78	61	1557
112	6	3540	4463	gi 854234	cymD gene product [Klebsiella oxytoca]	78	56	924
124	2	1888	1061	gi 405622	unknown [Bacillus subtilis]	78	60	828
130	3	1805	2260	gi 1256636	putative [Bacillus subtilis]	78	71	456
133	1	751	377	gi 168060	lamB [Emicella nidulans]	78	59	375
166	4	7125	6163	gi 451216	Mannosephosphate isomerase [Streptococcus mutans]	78	63	963
186	1	1586	795	gi 289284	cysteinyl-tRNA synthetase [Bacillus subtilis]	78	63	792
195	4	2749	2315	gi 1353874	unknown [Rhodospirillum rubrum]	78	58	435
199	3	4279	3623	gi 143525	succinate dehydrogenase cytochrome b-558 subunit [Bacillus subtilis] p1c[A29843][DEBSC succinate dehydrogenase (EC 1.3.99.1) cytochrome b-558 - Bacillus subtilis]	78	57	657
199	4	7209	5557	gi 142521	deoxyribodipyrimidine photolyase [Bacillus subtilis] p1c[A27192][A27192 uvrB protein - Bacillus subtilis sp p14931][UVRC_BACSU EXCINUCLEASE ABC SUBUNIT C]	78	62	1653
223	3	3831	3523	gi 139596	[Escherichia coli 15200 insertion sequence from ECOR6], partial 1, ene product [Escherichia coli]	78	47	309

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
289	4	1865	2149	gi1467439	temperature sensitive cell division [Bacillus subtilis]	78	62	285
321	9	7734	7315	gi1142979	ORF3 is homologous to an ORF downstream of the spot gene of E.coli; RFP [Bacillus stearothermophilus]	78	55	420
352	4	3714	3944	gi1349050	actin 1 [Pneumocystis carinii]	78	42	231
352	5	7592	6093	gi1903587	NADH dehydrogenase subunit 5 [Bacillus subtilis] ap139755 [NDHP_BACSU NADH DEHYDROGENASE SUBUNIT 5 (EC 1.6.5.3) NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5]	78	58	1500
376	1	2	583	gi1551693	dehydrobiotin synthase [Bacillus sphaericus]	78	34	582
424	2	1595	1768	gi11524117	alpha-acetolactate decarboxylase [Lactococcus lactis]	78	68	174
470	1	1914	988	gi11030068	NAD(P)H oxidoreductase, isoflavone reductase homologue [Solanum tuberosum]	78	63	927
558	1	762	562	gi11511588	bi-functional protein [Methanococcus jannaschii]	78	60	201
670	3	1152	1589	gi1122759	unknown [Bacillus subtilis]	78	64	438
714	1	64	732	gi1143460	37 kd minor sigma factor (rpoF, sigE; ttg start codon) [Bacillus subtilis]	78	57	669
814	1	3	368	gi11377833	unknown [Bacillus subtilis]	78	59	366
981	1	1381	692	gi1143802	CerC2 [Bacillus subtilis]	78	64	690
995	2	978	727	gi1296947	uridine kinase [Escherichia coli]	78	64	252
1045	1	3	401	gi11407784	orf-1; novel antigen [Staphylococcus aureus]	78	61	399
1163	2	368	186	gi11410117	diaminopimelate decarboxylase [Bacillus subtilis]	78	54	183
2191	1	794	399	gi1215098	excisionase [Bacteriophage 134a]	78	65	396
2933	1	2	181	gi11204436	pyruvate formate-lyase [Haemophilus influenzae]	78	73	180
3041	2	129	317	gi1624612	GLT [Escherichia coli]	78	53	189
3581	1	105	401	gi1763186	3-ketoacyl-CoA thiolase [Saccharomyces cerevisiae]	78	55	297
3709	1	3	230	gi1460689	TVG [Thermactinomyces vulgaris]	78	58	228
3974	1	528	265	gi1558839	unknown [Bacillus subtilis]	78	65	264
3980	1	3	401	gi139956	ITGlc [Bacillus subtilis]	78	62	399
4056	1	647	354	gi11256635	di-hydroxy-acid dehydratase [Bacillus subtilis]	78	55	294
4114	1	630	316	gi150937215093	hypothetical protein - Trypanosoma brucei	78	62	315
4185	1	3	179	gi11319950	large subunit of NADH-dependent glutamate synthase [Plectonena boryanum]	78	58	177

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4235	1	655	129	gi 558839	unknown [Bacillus subtilis]	78	60	327
4352	1	541	302	gi 603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis] gi 603768 HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis]	78	63	240
4368	1	612	307	gi 1353678	heavy-metal transporting P-type ATPase [Proteus mirabilis]	78	59	306
4461	1	428	216	gi 1276841	glutamate synthase (GOGAT) [Porphyra purpurea]	78	36	213
4530	1	474	238	gi 139956	ITOLC [Bacillus subtilis]	78	65	237
3	2	2969	2073	gi 1109684	ProV [Bacillus subtilis]	77	56	897
12	2	2426	1965	gi 467335	ribosomal protein L9 [Bacillus subtilis]	77	59	462
27	1	2	388	gi 1212728	Yqhi [Bacillus subtilis]	77	63	387
39	2	590	1252	gi 40054	phenylalanyl-tRNA synthetase beta subunit (AA 1-604) [Bacillus subtilis]	77	60	663
42	6	2704	2931	gi 606241	30S ribosomal subunit protein S14 [Escherichia coli] sp P03370 SS14_ECOLI 30S RIBOSOMAL PROTEIN S14. (SUB 2-101)	77	65	228
46	18	15459	16622	gi 297798	mitochondrial formate dehydrogenase precursor (Solanum tuberosum) pir JQ2272 JQ2272 formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato	77	55	1164
100	4	4562	4002	gi 1340128	ORF1 [Staphylococcus aureus]	77	54	561
102	8	5378	5713	gi 1311482	acetolactate synthase [Thermus aquaticus]	77	57	336
109	7	4742	5383	gi 710637	unknown [Bacillus subtilis]	77	56	642
117	1	2	1228	gi 1337015	ORF4 [Bacillus subtilis]	77	53	1227
124	10	8323	7888	gi 405819	thymidine kinase [Bacillus subtilis]	77	63	636
147	3	1146	985	gi 849027	hypothetical 15.9-kDa protein [Bacillus subtilis]	77	37	162
152	10	7354	7953	gi 1204583	spermidine/putrescine transport ATP-binding protein [Haemophilus influenzae]	77	55	600
169	2	1004	1282	gi 473825	elongation factor EF-Ts' [Escherichia coli]	77	58	279
184	2	380	1147	gi 216314	esterase [Bacillus stearothermophilus]	77	60	768
189	7	3296	3868	gi 853809	ORF3 [Clostridium perfringens]	77	48	573
193	1	132	280	gi 1303788	Yqeh [Bacillus subtilis]	77	54	159
195	8	8740	8414	gi 1499620	M. jannaschii predicted coding region M0798 [Methanococcus jannaschii]	77	44	327
205	8	5428	5204	gi 216340	ORF for adenylate kinase [Bacillus subtilis]	77	61	225

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
205	29	14795	14502	gi 786155	Ribosomal Protein L23 (Bacillus subtilis)	77	62	294
211	5	1908	2084	gi 410132	ORF6 (Bacillus subtilis)	77	47	177
217	5	3478	4416	gi 496254	fibrinectin/fibrinogen-binding protein (Streptococcus pyogenes)	77	54	939
222	1	267	998	gi 140784	orf-1, novel antigen (Staphylococcus aureus)	77	57	732
223	2	1819	1346	gi 467408	unknown (Bacillus subtilis)	77	61	474
243	3	2661	2299	gi 516155	unconventional myosin (Sus scrofa)	77	32	363
299	1	68	769	gi 467436	unknown (Bacillus subtilis)	77	54	702
301	4	1468	1283	gi 950071	ATP-bind. pyruvate kinase (Mycoplasma capricolum) p1r[S48605] (fragment) hypothetical protein - Mycoplasma capricolum SOC3 (fragment)	77	48	186
302	5	2741	3211	gi 508980	phsB (Bacillus subtilis)	77	57	471
302	7	3835	4863	gi 147783	ruvA protein (Escherichia coli)	77	60	1029
307	9	5402	4797	gi 1070015	protein-dependent (Bacillus subtilis)	77	60	606
312	1	99	1391	gi 143165	malic enzyme (EC 1.1.1.38) (Bacillus stearothermophilus) p1r[A33307] (DBSSX) malate dehydrogenase oxaloacetate-decarboxylating (EC 1.1.1.38) - Bacillus stearothermophilus	77	62	1293
312	2	1541	2443	gi 1398855	carboxyltransferase beta subunit (Synecococcus PCC7942)	77	58	901
321	5	5666	4596	gi 39844	fumarate (citG) [aa 1-462] (Bacillus subtilis)	77	65	1071
354	1	47	568	gi 154634	YmaR (Bacillus subtilis)	77	57	532
365	1	2	1021	gi 143374	phosphoribosyl glycineamide synthetase (PUR-D, gta start codon) Bacillus subtilis	77	62	1020
374	1	1	708	gi 1405446	transketolase (Bacillus subtilis)	77	61	708
385	1	1128	565	gi 533099	endonuclease III (Bacillus subtilis)	77	63	566
392	2	594	1940	gi 556014	UDP-N-acetyl muramate-alanine ligase (Bacillus subtilis) sp P40778 MURC_BACSU UDP-N-ACETYLURAMATE--ALANINE LIGASE (EC 3.2.8) (UDP-N-ACETYLURAMATE--ALANINE SYNTHETASE) (FRAGMENT)	77	65	1347
405	5	4079	3570	gi 1303912	Yqhw (Bacillus subtilis)	77	64	510
487	4	1302	1472	gi 432427	ORF1 gene product (Acinetobacter calcoaceticus)	77	48	171
522	1	2	562	gi A01179 SYNS	tyrosine--tRNA ligase (EC 6.1.1.1) - Bacillus stearothermophilus	77	63	561

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
523	2	1587	1351	gi11387979	44% identity over 302 residues with hypothetical protein from <i>Synechocystis</i> sp. accession D64006.CD; expression induced by environmental stress; some similarity to glycosyl transferases; two potential membrane-spanning helices [Bacillus subtilis]	77	48	237
536	2	983	612	gi1103366	adenylosuccinate lyase (PUB-8) [Bacillus subtilis] pir[C29336]W285DS adenylosuccinate lyase (EC 4.3.2.2) - Bacillus subtilis	77	61	372
548	2	339	872	gi1143387	isopartate transcarbamylase [Bacillus subtilis]	77	56	534
597	1	2	481	gi1904198	hypothetical protein [Bacillus subtilis]	77	33	480
633	2	1747	1313	gi1387577	ORF1A [Bacillus subtilis]	77	64	435
642	1	85	360	gi146971	epip gene product [Staphylococcus epidermidis]	77	61	276
659	1	125	1219	gi1072381	glutamy-aminopeptidase [Lactococcus lactis]	77	62	1095
670	4	1587	1820	gi1122760	unknown [Bacillus subtilis]	77	58	234
789	1	2	391	gi11377823	aminopeptidase [Bacillus subtilis]	77	65	390
815	1	10	573	gi11303861	YggM [Bacillus subtilis]	77	49	564
899	1	1	325	gi1130444	H. influenza predicted coding region H10594 [Haemophilus influenzae]	77	55	225
1083	1	3	188	gi1460828	B969 [Saccharomyces cerevisiae]	77	66	186
1942	1	415	209	gi1160047	p101/acidic basic repeat antigen [Plasmodium falciparum] pir[A29232]A29232 101K malaria antigen precursor - Plasmodium falciparum (strain Camp)	77	38	207
2559	1	1	171	gi11499034	H. jamaensis predicted coding region M2055 [Methanococcus jamaensis]	77	61	171
2933	2	243	401	gi142370	pyruvate formate-lyase (AA 1-760) [Escherichia coli] ir[IS01788]S01788 formate C-acetyltransferase (EC 2.3.1.54) - Escherichia coli	77	72	159
2966	1	56	292	gi11524397	glycine betaine transporter Opud [Bacillus subtilis]	77	45	337
2976	1	614	309	gi140003	oxoglutarate dehydrogenase (NADP+) [Bacillus subtilis] p[E23125]OD01_BACSU 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (EC 2.4.2) (ALPHA-KETOGLUTARATE DEHYDROGENASE)	77	60	306
2979	2	678	400	gi11204354	spore germination and vegetative growth protein [Haemophilus influenzae]	77	61	279
2988	1	601	377	gi1438465	Probable operon with orfP. Possible alternative initiation codon, aaaa 2151-2153. Homology with acetyltransferases.; putative Bacillus subtilis	77	55	225
2990	1	331	167	gi1142562	ATP synthase epsilon subunit [Bacillus megaterium] pir[B28599]PB85EM H-transporting ATP synthase (EC 3.6.1.34) pilon chain - Bacillus megaterium	77	63	165
3032	1	3	389	gi1488430	alcohol dehydrogenase 2 [Entamoeba histolytica]	77	56	387
3057	1	1	195	gi1468764	mcR gene product [Rhizobium meliloti]	77	50	195

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
4008	1	726	400	gi 503768	Hut1 protein, imidazolone-5-propiolate hydrolase (Bacillus subtilis) gi 503768 Hut1 protein, imidazolone-5-propiolate hydrolase Bacillus subtilis	77	52	327
4048	1	703	386	gi 216278	Gramicidin S synthetase 1 (Bacillus brevis)	77	55	318
4110	1	3	368	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	77	61	366
4115	1	1	348	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen (Streptococcus pyogenes)	77	65	348
4225	1	590	297	gi 1322245	nevalonate pyrophosphate decarboxylase (Rattus norvegicus)	77	60	294
4611	2	494	327	gi 508979	GTP-binding protein (Bacillus subtilis)	77	57	168
4668	1	361	182	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	77	61	180
25	1	2	1627	gi 1150620	HsaA (Streptococcus pneumoniae)	76	58	1626
38	5	1488	2537	pir A43577 A435	regulatory protein pfor - Clostridium perfringens	76	57	1050
52	5	2962	4041	gi 1161061	diacylglycerol kinase (Methylobacterium extorquens)	76	62	1080
56	20	27389	27955	gi 467402	unknown (Bacillus subtilis)	76	56	567
57	15	12046	12219	gi 1206040	weak similarity to keratin (Caenorhabditis elegans)	76	40	174
91	2	1062	2261	gi 475715	acetyl coenzyme A acetyltransferase (thiolase) [Clostridium acetobutylicum]	76	57	1200
98	2	818	1624	gi 467422	unknown (Bacillus subtilis)	76	62	807
98	5	2965	3228	gi 897793	ly98 gene product (Pediococcus acidilactici)	76	52	264
98	8	5922	6326	gi 467427	methionyl-tRNA synthetase (Bacillus subtilis)	76	53	405
104	3	1122	1885	gi 216151	DNA polymerase (gene L, tto start codon) [Bacteriophage SP02] gi 579197 SP02 DNA polymerase (aa 1-648) [Bacteriophage SP02] pir A21698 DJBPS2 DNA-directed DNA polymerase (EC 2.7.7) - phage P02	76	63	564
124	9	8134	7055	gi 853776	peptide chain release factor 1 (Bacillus subtilis) pir S55437 S55437 peptide chain release factor 1 - Bacillus subtilis	76	58	1080
164	5	2832	3311	gi 1204976	prolyl-CRNA synthetase [Haemophilus influenzae]	76	53	480
168	2	2617	1841	gi 117253	putative ATP-binding protein of ABC-type (Bacillus subtilis)	76	58	177
189	2	163	888	gi 467384	unknown (Bacillus subtilis)	76	63	726
235	3	2253	3518	gi 142936	polyl-polyglutamate synthetase (Bacillus subtilis) pir B40646 B40646 folC - Bacillus subtilis	76	53	1266
236	1	335	925	gi 1146197	putative (Bacillus subtilis)	76	54	591
237	8	5323	5541	gi 1279261	P1303.6 [Caenorhabditis elegans]	76	47	219

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
263	5	5490	4585	gi 1510348	[dihydrodipicolinate synthase (Methanococcus jannaschii)]	76	49	906
304	3	1051	1794	gi 666982	[putative membrane spanning subunit (Bacillus subtilis)] p1r[S52382] p1r[S52382] [Bacillus subtilis]	76	60	744
312	4	3611	4624	gi 143312	[6-phospho-1-fructokinase (99% start codon; EC 2.7.1.11) (Bacillus carotermophilus)]	76	56	1014
343	1	2	1036	gi 405956	[yeeE (Escherichia coli)]	76	59	1035
347	1	409	1701	gi 396304	[acetylornithine decarboxylase (Escherichia coli)]	76	72	1293
358	1	672	1907	gi 1146215	[39.0% identity to the Escherichia coli S1 ribosomal protein; putative (Bacillus subtilis)]	76	58	1236
371	1	1	222	gi 537084	[alternate gene name mgt; CG Site No. 497 (Escherichia coli)] p1r[S56468] S56468 mgtA protein - Escherichia coli	76	61	222
379	4	4331	4858	gi 143268	[dihydroisopentenyl transsuccinylase (odhB; EC 2.3.1.61) (Bacillus subtilis)]	76	61	528
404	5	4022	4492	gi 1303823	[YqG (Bacillus subtilis)]	76	60	471
411	1	2	307	gi 186025	[ORF YKL07W (Saccharomyces cerevisiae)]	76	55	306
472	3	4356	2854	gi 1405464	[Alat (Bacillus subtilis)]	76	57	1503
546	1	273	995	gi 153821	[streptococcal pyrogenic exotoxin type C (speC) precursor Streptococcus pyogenes]	76	36	723
588	1	1054	557	gi 1002520	[MutS (Bacillus subtilis)]	76	61	498
591	1	16	735	gi 885934	[CipB (Synecococcus sp.)]	76	44	720
602	2	175	798	gi 1486422	[OppD homologue (Rhizobium sp.)]	76	52	624
619	2	547	290	gi 330613	[major capsid protein (Human cytomegalovirus)]	76	47	258
660	4	2568	3302	gi 904199	[hypothetical protein (Bacillus subtilis)]	76	55	735
677	1	452	228	gi 40177	[spoOF gene product (Bacillus subtilis)]	76	58	225
962	1	24	206	gi 142443	[adenylosuccinate synthetase (Bacillus subtilis)] sp1p79726[PURA-BACSU ADENYLOSUCCINATE SYNTHETASE (EC 6.3.4.4) INP-ASPARTATE LIGASE]	76	67	183
978	1	1158	580	gi 1511333	[M. jannaschii predicted coding region M3322 (Methanococcus jannaschii)]	76	56	579
997	1	486	244	gi 467154	[No definition line found (Mycobacterium leprae)]	76	38	243
1563	1	529	266	gi 1303984	[YqkG (Bacillus subtilis)]	76	52	264
2184	1	361	182	gi 506706	[CapJ (Staphylococcus aureus)]	76	38	180
2572	1	1	387	gi 153898	[transport protein (Salmonella typhimurium)]	76	65	387

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
2942	1	29	400	gi 210020	nitrite reductase (nirB) [Bacillus subtilis]	76	59	372
2957	1	377	216	gi 1511251	hypothetical protein (SP-P4204) [Methanococcus jannaschii]	76	47	162
2980	1	554	279	gi 1405464	Alst [Bacillus subtilis]	76	53	276
3015	1	649	326	gi 108115	ornithine acetyltransferase [Bacillus subtilis]	76	61	324
3124	1	13	174	gi 882705	ORF_0401 [Escherichia coli]	76	65	162
3179	1	3	161	gi 168477	ferredoxin-dependent glutamate synthase (Zea mays) pir[A38596]A38596 glutamate synthase (ferredoxin) (EC 1.4.7.1) - size	76	53	159
3789	1	2	379	gi 39956	110Lc [Bacillus subtilis]	76	55	378
3892	1	3	314	gi 1510198	ferritinocellin binding protein [Methanococcus jannaschii]	76	52	312
3928	1	798	400	gi 143016	permease [Bacillus subtilis]	76	59	399
4159	1	757	386	gi P80544 HMRP_	METHICILLIN-RESISTANT SURFACE PROTEIN (FRAGMENTS)	76	66	372
4204	1	17	331	gi 296464	ATPase [Lactococcus lactis]	76	56	315
4398	1	494	249	gi 987255	Menkes disease gene [Homo sapiens]	76	48	246
4506	1	2	313	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	76	47	312
4546	1	477	247	gi 133950	large subunit of NAM-dependent glucanase synthase [Plectonema boryanum]	76	61	231
4596	1	379	191	gi 560027	cellulose synthase [Acetobacter xylinum]	76	70	189
4	5	5257	4337	gi 882532	ORF_0294 [Escherichia coli]	75	59	921
6	1	164	952	gi 40960	OTCase [Escherichia coli]	75	56	789
12	3	5935	3944	gi 467336	unknown [Bacillus subtilis]	75	57	1992
23	16	18272	17310	gi 1296433	O-acetylserine sulphydrylase B [Alcaligenes eutrophus]	75	55	963
25	3	2356	3393	gi 1502419	Plex [Bacillus subtilis]	75	56	1038
36	8	5765	6037	gi 1256517	unknown [Schizosaccharomyces pombe]	75	45	273
46	13	11186	12058	gi 48972	nitrate transporter [Synechococcus sp.]	75	46	873
51	7	3474	3677	gi 113607	sporulation protein [Bacillus subtilis]	75	61	204
53	16	16850	16590	gi 143402	recombination protein (ttg start codon) [Bacillus subtilis] gi 1303923 RecW [Bacillus subtilis]	75	51	261
74	3	3572	2568	gi 1204847	ornithine carbamoyltransferase [Haemophilus influenzae]	75	61	1005

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
85	3	4628	3930	gi1143368	phosphoribosylformyl glycinamide synthetase I (PUR-L; gtc start odon) (Bacillus subtilis)	75	63	699
85	5	5588	4878	gi1143367	phosphoribosyl aminimidazole succinocarboxamide synthetase (PUR-C; gtc start codon) (Bacillus subtilis)	75	55	713
85	8	6625	7530	gi11303916	yqia (Bacillus subtilis)	75	53	906
87	3	2340	3590	gi11064813	homologous to sp-PHOR_BACSU (Bacillus subtilis)	75	56	1251
87	6	6084	6896	gi11064810	function unknown (Bacillus subtilis)	75	61	813
108	2	1844	1503	gi11001824	hypothetical protein (Synechocystis sp.)	75	53	342
110	3	1748	3727	gi11147593	putative ppGpp synthetase (Streptomyces coelicolor)	75	55	1980
110	7	4151	5252	gi11177251	icld gene product (Bacillus subtilis)	75	75	900
120	14	11266	10649	gi11524394	ORF-2 upstream of gbsAB operon (Bacillus subtilis)	75	55	618
121	5	2050	4221	gi11154632	Mrde (Bacillus subtilis)	75	54	2172
124	1	283	143	gi11405622	unknown (Bacillus subtilis)	75	56	141
128	1	81	1139	gi1143316	igpp gene products (Bacillus megaterium)	75	48	1059
130	8	5760	5903	gi11256654	S4.8A identity with <i>Neisseria gonorrhoeae</i> regulatory protein PilB; putative (Bacillus subtilis)	75	62	144
136	2	4480	3185	gi11467403	seryl-tRNA synthetase (Bacillus subtilis)	75	54	1296
161	10	5439	5798	gi11001195	hypothetical protein (Synechocystis sp.)	75	55	360
172	4	3819	2995	gi11755153	ATP-binding protein (Bacillus subtilis)	75	52	825
179	1	2024	1107	gi1141037	porphobilinogen deaminase (Bacillus subtilis)	75	58	918
195	10	9529	9174	sp1125745 YCFN	HYPOTHETICAL PROTEIN IN PURB 5'-REGION (ORF-15) (FRAGMENT)	75	60	356
200	4	2605	4596	gi1142440	ATP-dependent nuclease (Bacillus subtilis)	75	56	1992
206	3	6900	5820	gi11256135	ybbp (Bacillus subtilis)	75	53	1281
216	2	159	389	gi11052800	unknown (Schizosaccharomyces pombe)	75	58	231
229	1	29	847	gi11205958	branched chain aa transport system II carrier protein (Haemophilus influenzae)	75	49	819
230	2	518	1714	gi11971337	nitrite extrusion protein (Bacillus subtilis)	75	53	1197
231	1	2240	1122	gi11002521	NucB (Bacillus subtilis)	75	54	1119
233	3	1314	1859	gi11467405	unknown (Bacillus subtilis)	75	59	546

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	seqth gene name	% sim	% ident	length (nt)
269	1	325	164	gi11511246	methy1 coenzyme M reductase system, component A2 [Methanococcus jannaschii]	75	50	162
292	1	1389	772	gi11511604	M. jannaschii predicted coding region MJ1651 [Methanococcus jannaschii]	75	46	618
304	4	1773	2261	gi1205328	surfactin [Haemophilus influenzae]	75	55	489
312	3	2437	3387	gi1285621	undefined open reading frame [Bacillus stearothermophilus]	75	62	951
312	5	4622	6403	gi11041097	pyruvate kinase [Bacillus psychrophilus]	75	57	1782
319	1	353	877	gi11212728	YqjH [Bacillus subtilis]	75	54	525
320	5	4321	5031	gi11070361	OMP decarboxylase [Lactococcus lactis]	75	56	711
320	6	5010	5642	gi11433394	OMP-PRPP transferase [Bacillus subtilis]	75	60	633
337	4	1519	2088	gi1487433	citrate synthase II [Bacillus subtilis]	75	58	570
394	2	669	1271	gi1304976	matches P80017: ATP-GTP_A and P800301: EFACROR-GTP; similar to longation factor G, TetH/TetO tetracycline-resistance proteins Escherichia coli	75	51	603
423	1	127	570	gi1183839	unknown [Pseudomonas aeruginosa]	75	59	444
433	2	1603	1929	gi1149211	acetolactate synthase [Klebsiella pneumoniae]	75	63	327
446	2	176	1540	gi1312441	olhydroorotase [Bacillus caldolyticus]	75	62	1365
486	1	494	249	gi1149682	potf gene product [Clostridium perfringens]	75	55	246
496	1	3	794	gi1143582	spoIIIEA protein [Bacillus subtilis]	75	59	792
498	2	824	1504	gi1143328	phop protein (put.): putative [Bacillus subtilis]	75	47	681
499	2	1061	1624	gi11387979	44% identity over 302 residues with hypothetical protein from Synachocystis sp. accession D64006.CD; expression induced by environmental stress; some similarity to glycosyl transferases; two potential membrane-spanning helices [Bacillus subtilis]	75	51	564
568	1	641	453	gi134410JC41 (SC3)	triacylglycerol lipase (EC 3.1.1.3) 2 - Mycoplasma mycoides subsp. mycoides	75	50	189
613	2	430	233	gi1330993	tegument protein [Saimirinae herpesvirus 2]	75	75	198
621	1	1	525	gi1529754	speC [Streptococcus pyogenes]	75	43	525
642	5	1809	2474	gi1176401	Ep10 [Staphylococcus epidermidis]	75	51	666
646	2	454	657	gi1172442	ribonuclease P [Saccharomyces cerevisiae]	75	37	204
657	1	3	347	gi1882541	OMP_o236 [Escherichia coli]	75	47	345
750	1	1662	832	gi146971	ep1P gene product [Staphylococcus epidermidis]	75	57	831

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
754	1	2	481	gi 1101901	yqhP [Bacillus subtilis]	75	57	480
763	2	563	393	gi 1205145	multidrug resistance protein [Haemophilus influenzae]	75	51	171
775	1	961	482	pir 016889 0168	leuA protein, inactive - Lactococcus lactis subsp. lactis (strain IL1403)	75	63	480
793	1	1	180	gi 143316	[gap] gene products [Bacillus megaterium]	75	57	180
800	1	318	160	gi 509411	NFRA protein [Azorhizobium caulinodans]	75	34	159
811	1	1117	560	gi 143434	Rho factor [Bacillus subtilis]	75	60	558
940	1	493	329	gi 1276985	arginase [Bacillus caldovelox]	75	50	165
971	2	37	252	gi 1001373	hypothetical protein [Synecococcus sp.]	75	58	216
1059	1	384	232	gi 1726480	L-glutamine-D-fructose-6-phosphate amidotransferase [Bacillus subtilis]	75	67	153
1109	2	219	374	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis] pir A27650 A27650 regulatory protein phoR - Bacillus subtilis sp P23545 PHOR_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOR (EC 2.7.3.-)	75	53	156
1268	1	271	137	gi 104135	ornithine acetyltransferase [Bacillus stearothermophilus] sp Q07903 ARG1_BACST GLUTAMATE N-ACETYLTRANSFERASE (EC 2.3.1.35) ORNITHINE ACETYLTRANSFERASE (ORNITHINE TRANSACETYLASE) (OATASE) / HIND-ACID ACETYLTRANSFERASE (EC 2.3.1.1) (N-ACETYLGLUTAMATE YNTUA	75	53	135
1500	1	324	163	gi 1203488	lexcinuclease ABC subunit B [Haemophilus influenzae]	75	57	162
1529	1	798	400	gi 1002521	MutL [Bacillus subtilis]	75	54	399
3010	1	770	387	gi 1204435	pyruvate formate-lyase activating enzyme [Haemophilus influenzae]	75	54	384
3105	1	1	180	gi 1041097	pyruvate kinase [Bacillus psychrophilus]	75	57	180
3117	1	45	212	gi 099317	peptide synthetase module [Microcystis aeruginosa] pir S49111 S49111 probable amino acid activating domain - Microcystis aeruginosa (fragment) (SUB 144-528)	75	42	168
3139	2	139	345	gi 145294	adenine phosphoribosyl-transferase [Escherichia coli]	75	66	207
3880	1	618	310	gi 1009166	respiratory nitrate reductase [Bacillus subtilis]	75	58	309
3911	1	48	401	gi 433991	ATP synthase subunit beta [Bacillus subtilis]	75	68	354
3957	1	2	379	pir 016889 0168	3-isopropylmalate dehydratase (EC 4.2.1.33) chain leuC - Lactococcus lactis subsp. lactis (strain IL1403)	75	65	378
4005	1	5	259	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	75	48	255
4080	1	73	333	gi 415855	deoxyribose aldolase [Mycoplama hominis]	75	59	261

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4111	1	1	339	gi149435	putative (Lactococcus lactis)	75	57	339
4136	1	602	303	gi1450688	hard gene of Ecoprr1 gene product (Escherichia coli) p1r[S30437]S30437 hsdM protein - Escherichia coli p1r[S09629]S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	75	56	300
4144	1	668	336	gi148972	nitrate transporter (Synechococcus sp.)	75	49	333
4237	1	664	374	gi1339950	large subunit of NADH-dependent glutamate synthase (Plectonema boryanum)	75	55	291
4306	2	73	318	gi1294260	major surface glycoprotein (Pneumocystis carinii)	75	68	246
4343	1	715	359	gi1204652	methylated-DNA-protein-cysteine methyltransferase (Haemophilus influenzae)	75	52	357
4552	1	620	312	gi1396464	ATPase (Lactococcus lactis)	75	55	309
48	9	5776	6126	gi1443793	NupC (Escherichia coli)	74	50	351
50	8	6910	6221	gi1239988	hypothetical protein (Bacillus subtilis)	74	55	690
56	9	10770	12221	gi13000451	TrpP (Bacillus subtilis)	74	57	1452
64	2	2266	1622	gi141015	aspartate-tRNA ligase (Escherichia coli)	74	57	645
66	6	5063	4848	gi13212729	YohJ (Bacillus subtilis)	74	47	216
67	18	14334	14897	gi1510631	endoglucanase (Methanococcus jannaschii)	74	52	564
102	15	12561	13136	gi1149429	putative (Lactococcus lactis)	74	67	576
102	16	13121	14419	gi1149435	putative (Lactococcus lactis)	74	57	1299
108	4	4873	3902	gi139478	ATP binding protein of transport ATPases (Bacillus firmus) tr[S15486]S15486 ATP-binding protein - Bacillus firmus p1P26946 YATR_BACFI HYPOTHETICAL ATP-BINDING TRANSPORT PROTEIN	74	59	972
116	5	8574	7093	gi1205430	dipeptide transport system permease protein (Haemophilus influenzae)	74	49	1482
120	7	4142	4803	gi1146970	ribonucleoside triphosphate reductase (Escherichia coli) p1r[A47331]A47331 anaerobic ribonucleotide reductase - Escherichia coli	74	58	462
121	7	5961	6581	gi1107528	lctg start (Campylobacter coli)	74	51	621
128	3	2320	3531	gi11443118	phosphoglycerate kinase (Bacillus megaterium)	74	57	1212
130	7	5237	5791	gi1256653	DNA-binding protein (Bacillus subtilis)	74	60	555
136	3	6745	5150	gi1143076	histidase (Bacillus subtilis)	74	58	1596
145	2	664	1368	gi1407733	devA gene product (Anabaena sp.)	74	45	705
152	1	552	277	gi1377833	unknown (Bacillus subtilis)	74	54	276

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	saigh gene name	% sim	% ident	length (nt)
164	10	11064	11375	gi1580900	ORF3 gene product [Bacillus subtilis]	74	52	312
175	2	3109	2624	gi1642656	unknown [Rhizobium meliloti]	74	34	486
175	9	6064	5612	gi1854656	[Na/H antiporter system ORF2 [Bacillus alcalophilus]	74	46	453
195	11	111346	10339	gi11204430	hypothetical protein [SP-25745] [Haemophilus influenzae]	74	55	1008
205	17	9619	9059	gi11044979	ribosomal protein L6 [Bacillus subtilis]	74	64	561
236	7	5574	6710	gi11146207	putative [Bacillus subtilis]	74	63	1137
241	3	4521	3334	gi1694121	malate thiokinase [Methylobacterium extorquens]	74	52	1188
246	6	3105	2799	gi1467374	single strand DNA binding protein [Bacillus subtilis]	74	64	507
249	4	6551	5313	gi11524397	glycine betaine transporter Opud [Bacillus subtilis]	74	55	1239
261	7	4389	4081	gi1809542	CbrB protein [Erwinia chrysanthemi]	74	42	309
278	6	5714	4665	gi11204872	ATP-binding protein [Haemophilus influenzae]	74	54	1050
309	1	1220	666	gi11205579	hypothetical protein [GB:U14003_302] [Haemophilus influenzae]	74	53	555
315	2	1473	862	gi1143198	guinol oxidase [Bacillus subtilis]	74	57	612
320	1	1	1065	gi143389	glutaminase of carbonyl phosphate synthetase [Bacillus subtilis] p1r123945123945 carbamoyl-phosphate synthase glutamine-hydrolysing (EC 6.3.5.5), pyrimidine-repressible, small hain - Bacillus subtilis	74	60	1065
380	2	382	1128	gi1534857	ATPase subunit a [Bacillus stearothermophilus]	74	56	747
405	2	1742	1311	gi1303915	Yqh2 [Bacillus subtilis]	74	65	432
431	5	2503	3270	gi1473902	alpha-acetolactate synthase [Lactococcus lactis]	74	56	768
452	1	1	942	gi1413982	ipa-58r gene product [Bacillus subtilis]	74	52	942
461	1	3	1193	gi1558494	homoserine dehydrogenase [Bacillus subtilis]	74	51	1191
461	2	1174	1407	gi140211	threonine synthase (thrC) [AA 1-352] [Bacillus subtilis] ir[A25364]A25364 p1r104728104728 threonine synthase (EC 4.2.99.2) - Bacillus btills	74	56	234
462	2	402	734	gi1142520	thioredoxin [Bacillus subtilis]	74	62	333
478	1	574	320	gi11699005	glycyl-tRNA synthetase [Methanococcus jannaschii]	74	52	255
501	2	739	1740	gi1217040	acid glycoprotein [Streptococcus pyogenes]	74	58	1002
551	2	4083	2791	gi1143040	glutamate-1-semialdehyde 2,1-aminotransferase [Bacillus subtilis] p1r104728104728 glutamate-1-semialdehyde 2,1-aminotransferase (EC 4.3.8) - Bacillus subtilis	74	51	1293

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
573	1	1	477	gi1006605	hypothetical protein (Synecocystis sp. 1)	74	45	477
596	2	1780	1298	gi13301853	YqgP (Bacillus subtilis)	74	55	483
618	2	2924	1758	gi11146237	21.4% of identity to trans-acting transcription factor of <i>Saccharomyces cerevisiae</i> ; 25% of identity to sucrose synthase of <i>Zea mays</i> ; putative (Bacillus subtilis)	74	55	1167
659	2	1269	1595	gi11072380	ORF3 (Lactococcus lactis)	74	62	327
724	1	373	188	gi1143174	phosphoribosyl glycine synthetase (Pur-D, gtp start codon) (Bacillus subtilis)	74	58	186
743	2	604	1209	gi1151833	ORF1, putative (Streptococcus parasanguis)	74	50	606
836	1	2	259	gi1143458	ORF V (Bacillus subtilis)	74	47	258
989	2	443	724	gi11303994	YqkM (Bacillus subtilis)	74	46	282
1106	1	1	492	gi1146970	epiD gene product (Staphylococcus epidermidis)	74	54	492
1135	2	373	528	gi1141948	ipa-24d gene product (Bacillus subtilis)	74	48	156
1234	1	817	412	gi11495245	rocJ gene product (Erwinia chrysanthemi)	74	36	366
2586	1	2	238	gi11149701	sbcC gene product (Clostridium perfringens)	74	62	237
2959	1	798	400	gi11405454	aconitase (Bacillus subtilis)	74	60	399
2962	1	650	363	gi11450586	3-phosphoglycerate kinase (Thermotoga maritima)	74	58	288
2983	1	3	191	gi11301893	YqkL (Bacillus subtilis)	74	56	189
3018	1	2	223	gi1143040	glutamate-1-semialdehyde 2,1-aminotransferase (Bacillus subtilis) pirD42728[042728] glutamate-1-semialdehyde 2,1-aminotransferase (EC 4.3.8) - Bacillus subtilis	74	56	222
3038	1	510	256	pir1552915[S529]	nitrate reductase alpha chain - Bacillus subtilis (fragment)	74	57	255
3062	1	374	189	gi1107528	ltg start (Campylobacter coli)	74	51	186
4035	1	184	360	gi11022725	unknown (Staphylococcus haemolyticus)	74	64	177
4045	1	607	305	gi11510977	M. jannaschii predicted coding region M10918 (Methanococcus jannaschii)	74	41	303
4283	1	471	304	gi11520844	orf4 (Bacillus subtilis)	74	58	168
4449	1	3	221	gi11580910	peptide synthetase ORF1 (Bacillus subtilis)	74	54	219
4587	1	458	231	gi11370207	orf6 (Lactobacillus sakei)	74	59	228

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4603	1	29	214	gi 146208	glutamate synthase large subunit (EC 2.6.1.5) [Escherichia coli] p1r[A25617/A25617 glutamate synthase (NADPH) (EC 1.4.1.13) large hain - Escherichia coli	74	60	186
4670	1	366	184	gi 1256135	lybB [Bacillus subtilis]	74	61	183
5	10	7953	7162	gi 133727	putative [Bacillus subtilis]	73	42	792
11	2	2454	1372	gi 166338	dihydroorotate dehydrogenase [Agrobacterium tumefaciens]	73	55	1083
14	1	2024	1020	gi 143373	phosphoribosyl aminimidazole carboxy formyl ornyltransferase/inosine monophosphate cyclohydrolyase [Pur-H(2)] [Bacillus subtilis]	73	54	1005
23	5	5426	4635	gi 1468939	meso-2,3-butanediol dehydrogenase (D-acetoin forming) [Klebsiella pneumoniae]	73	58	792
23	17	17379	16360	gi 297060	ornithine cyclodeaminase [Rhizobium meliloti]	73	37	1020
29	2	692	1273	gi 467442	stage V sporulation [Bacillus subtilis]	73	54	582
31	5	6467	4914	gi 414000	lipa-76d gene product [Bacillus subtilis]	73	55	1554
37	8	8658	7402	gi 429259	pept gene product [Bacillus subtilis]	73	59	1257
37	9	7738	7562	gi 168367	alpha-isopropylmalate isomerase (put.1; putative [Rhizomucor ircinellodes])	73	52	177
38	7	3931	4896	gi 405485	yeiN [Escherichia coli]	73	58	966
44	6	5041	4238	gi 580895	unknown [Bacillus subtilis]	73	53	804
44	11	7767	8306	gi 42009	moaB gene product [Escherichia coli]	73	50	540
45	3	2439	3080	gi 1109685	proM [Bacillus subtilis]	73	47	642
54	13	14036	13794	gi 413911	lipa-7d gene product [Bacillus subtilis]	73	61	243
59	4	1430	2248	gi 147923	threonine dehydratase 2 (EC 4.2.1.16) [Escherichia coli]	73	53	819
65	1	1458	730	gi 1677944	AppF [Bacillus subtilis]	73	56	729
80	2	1375	860	gi 580932	murD gene product [Bacillus subtilis]	73	53	516
102	13	10124	11179	gi 580891	3-isopropylmalate dehydrogenase (AA 1 - 365) [Bacillus subtilis] p1r[A2522/A2522 3-isopropylmalate dehydrogenase (EC 1.1.1.85) - acillus subtilis	73	55	1056
109	2	3493	2600	gi 1510849	M. jannaschii predicted coding region M0775 [Methanococcus jannaschii]	73	40	894
120	8	4782	5756	gi 146970	ribonucleoside triphosphate reductase [Escherichia coli] p1r[A47331/A47331 anaerobic ribonucleotide reductase - Escherichia coli	73	56	975
120	9	5726	6223	gi 1204333	anserobic ribonucleoside-triphosphate reductase [Haemophilus influenzae]	73	62	498

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
132	5	4151	4363	gi1871048	HPSM2 - heavy chain potential motor protein (Giardia intestinalis)	73	43	213
140	6	5952	4324	gi1634107	kdpB (Escherichia coli)	73	59	1629
142	6	7060	5919	gi1410125	rlxI gene product (Bacillus subtilis)	73	57	1122
149	4	1866	1717	gi1460892	heparin binding protein-44, HBP-44 (mice, peptide, 160 aa) p1r1JX0281JX0281 heparin-binding protein-44 precursor - mouse gi1220434 ORF Mus musculus (SUB 2-360)	73	53	150
158	1	1	1431	gi1882504	ORF_1560 (Escherichia coli)	73	57	1431
174	6	5352	4525	gi11146240	ketopentolate hydroxymethyltransferase (Bacillus subtilis)	73	55	838
175	8	5537	5178	gi1854657	Na/H antiporter system ORF3 (Bacillus alcalophilus)	73	56	360
186	5	6593	5493	gi1467477	unknown (Bacillus subtilis)	73	48	1101
249	6	6283	5729	gi11524197	glycine betaine transporter OpuD (Bacillus subtilis)	73	56	555
265	4	1873	2280	gi139848	U3 (Bacillus subtilis)	73	41	408
270	1	328	582	gi1780461	220 kDa polypeptide (African swine fever virus)	73	53	255
278	4	4283	3618	gi11204965	hypothetical 23.3 kD protein (Escherichia coli)	73	49	666
279	3	4984	3593	gi11185288	isochloranate synthase (Bacillus subtilis)	73	58	1392
291	4	1207	1575	gi11511440	glutamine-fructose-6-phosphate transaminase (Methanococcus jannaschii)	73	63	369
299	2	735	1166	gi1467437	unknown (Bacillus subtilis)	73	58	432
299	5	2050	3234	gi1467439	temperature sensitive cell division (Bacillus subtilis)	73	53	1185
334	1	1237	728	gi1536655	ORF YBR246w (Saccharomyces cerevisiae)	73	43	510
336	2	1827	1036	gi1790943	urea amidolyase (Bacillus subtilis)	73	51	792
374	3	1389	1874	gi11405451	YneJ (Bacillus subtilis)	73	55	486
433	4	1916	2554	gi1473902	alpha-acetolactate synthase (Lactococcus lactis)	73	54	639
509	2	1795	1028	gi1467483	unknown (Bacillus subtilis)	73	56	768
513	1	1709	918	gi11146220	NAD+ dependent glycerol-3-phosphate dehydrogenase (Bacillus subtilis)	73	56	792
533	2	239	733	gi11510605	hypothetical protein (SP:P42297) (Methanococcus jannaschii)	73	44	495
546	2	1148	2815	gi141748	hcdM protein (AA 1-520) (Escherichia coli)	73	52	1668
549	1	762	382	gi11314847	CinA (Bacillus subtilis)	73	57	381
567	1	1346	675	gi1410137	ORFX13 (Bacillus subtilis)	73	58	672

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% Ident	length (nt)
716	2	654	1112	gi 1256623	exodeoxyribonuclease (Bacillus subtilis)	73	56	459
772	1	3	677	gi 142010	Shows 70-21 similarity and 48-61 identity to the EnvM protein of salmonella typhimurium (Anebaena sp.)	73	57	675
774	1	3	209	gi 409286	hmrU (Bacillus subtilis)	73	52	207
782	1	1	402	gi 141320	[gap] gene products (Bacillus megaterium)	73	56	402
789	2	451	762	gi 1063266	low homology to p14 protein of Hemophilus influenzae and 14.2 kDa protein of Escherichia coli (Bacillus subtilis)	73	56	312
796	1	3	911	gi 853754	ABC transporter (Bacillus subtilis)	73	58	909
806	3	1209	949	gi 143786	tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Bacillus subtilis) pirJTD481 YMB5 tryptophan--tRNA ligase (EC 6.1.1.2) - Bacillus ubtills	73	51	261
816	2	4839	3097	gi 41748	hdm protein (AA 1-520) (Escherichia coli)	73	52	1743
839	1	798	400	gi 886906	argininosuccinate synthetase (Streptomyces clavuligerus) pir S57659 S57659 argininosuccinate synthase (EC 6.3.4.5) - Streptomyces clavuligerus	73	59	399
857	1	3	290	gi 348052	acetoin utilisation protein (Bacillus subtilis)	73	50	288
1008	1	790	398	gi 40100	rodC (tag) polypeptide (AA 1-746) (Bacillus subtilis) L S06049 S06049 rodC protein - Bacillus subtilis p P13685 TAGP-BACSU TECHNIC ACID BIOSYNTHESIS PROTEIN F.	73	41	393
1018	1	1	213	gi 529357	No definition line found (Caenorhabditis elegans) sp P46975 STT3_CABEL OLIGOSACCHARYL TRANSFERASE STT3 SUBUNIT OHOLOG.	73	53	213
1011	1	3	491	gi 142706	comG1 gene product (Bacillus subtilis)	73	51	489
1174	1	395	204	gi 1149513	alpha2 subunit of laminin 5 (Homo sapiens)	73	60	192
1175	1	655	329	gi 473817	ORF (Escherichia coli)	73	57	327
1187	1	3	209	gi 580870	lipa-37d qoxA gene product (Bacillus subtilis)	73	52	207
1206	1	72	245	gi 144816	formyltetrahydrofolate synthetase (FMFS) (ttg start codon) (EC 3.4.3) (Moorella thermocatal)	73	43	174
1454	1	423	241	gi 1213253	unknown (Schizosaccharomyces pombe)	73	53	183
1469	1	517	260	gi 1303787	YqeG (Bacillus subtilis)	73	55	258
1761	1	374	189	gi 9135	Hst26Aa gene product (Drosophila simulans)	73	34	186
1849	1	467	243	gi 162307	DNA topoisomerase II (Trypanosoma cruzi)	73	60	225
2055	1	2	400	gi 159381	p47K protein (Rhodococcus erythropolis)	73	34	399
2556	1	2	244	gi 145925	fecB (Escherichia coli)	73	62	243

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
2947	2	549	400	gi 1184680	[polynucleotide phosphorylase [Bacillus subtilis]]	73	51	150
2956	1	746	375	gi 143397	[guinol oxidase [Bacillus subtilis]]	73	58	372
3017	1	655	329	gi 111091	[acetylactate synthase [Bacillus subtilis]]	73	55	327
3115	1	385	194	gi 1223866	[overlapping out-of-phase protein (Egplant mosaic virus) sp P20139 V70X_EPMV 70 RD PROTEIN.	73	53	192
3603	2	700	527	gi 1419521	[glutaryl-CoA dehydrogenase precursor [Mus musculus]]	73	48	174
3743	1	798	400	gi 450688	[hadM gene of Ecopir1 gene product [Escherichia coli] pir S38437 S38437 hadM protein - Escherichia coli pir S09629 S09629 hypothetical protein A - Escherichia coli (SUS 40-520)]	73	54	399
3752	1	640	359	gi 1524193	[unknown [Mycobacterium tuberculosis]]	73	59	282
3852	1	2	181	gi 216746	[D-lactate dehydrogenase [Lactobacillus plantarum]]	73	68	180
3914	1	475	239	pir S13490 S134	[Hydroxymethylglutaryl-CoA synthase (EC 4.1.3.5) - Chicken (fragment)]	73	53	237
3914	2	570	343	gi 528991	[unknown [Bacillus subtilis]]	73	30	228
4069	1	2	316	gi 40003	[oxoglutarate dehydrogenase (HADP+)] [Bacillus subtilis] n P23129 OD01_BACSU 2-OXOGLUTARATE DEHYDROGENASE. EI COMPONENT (EC 2.4.2) (ALPHA- KETOGLUTARATE DEHYDROGENASE)]	73	55	315
4165	1	715	365	gi 1419521	[glutaryl-CoA dehydrogenase precursor [Mus musculus]]	73	48	351
4196	1	1	177	gi 809660	[deoxyribose-phosphate aldolase [Bacillus subtilis] pir S49455 S49455 deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis]	73	60	177
4202	1	572	378	gi 528991	[unknown [Bacillus subtilis]]	73	38	195
4314	1	2	193	gi 416797	[N-acyl-L-amino acid amidohydrolase [Bacillus stearothermophilus] sp P37112 AHA_BACST N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 5.1.14) (AMINOACYLASE)]	73	47	192
4393	1	3	263	gi 216267	[ORF2 [Bacillus megaterium]]	73	47	261
35	2	903	1973	gi 1146196	[phosphoglycerate dehydrogenase [Bacillus subtilis]]	72	53	1071
38	22	19094	17877	gi 602031	[similar to trimethylamine dihydrogenase (EC 5.99.7) - Mycoplasma capricolus. probable trimethylamine dehydrogenase (EC 5.99.7) - Mycoplasma capricolus. (SGC3) (fragment)]	72	54	1218
38	23	18134	19162	gi 413968	[lpa-44d gene product [Bacillus subtilis]]	72	54	1029
44	19	11895	12953	gi 1516272	[unknown [Bacillus subtilis]]	72	49	1059
48	7	6248	7117	gi 414699	[pyruvate synthase [Halobacterium halobium]]	72	49	870
50	7	6563	5691	gi 11205399	[proton glutamate symport protein [Haemophilus influenzae]]	72	53	873

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
53	9	10521	9259	gi 1303956	YqjE [Bacillus subtilis]	72	52	1263
56	23	29549	29995	gi 467471	unknown [Bacillus subtilis]	72	47	447
69	4	5298	4123	gi 1354775	pfoS/R [Treponema pallidum]	72	46	1176
69	5	4377	4982	gi 904198	hypothetical protein [Bacillus subtilis]	72	43	606
73	1	2	856	gi 142997	glycerol uptake facilitator [Bacillus subtilis]	72	59	855
98	13	9371	10258	gi 467435	unknown [Bacillus subtilis]	72	50	888
127	1	1	1593	gi 217144	alanine carrier protein [thermophilic bacterium PS3] pir[A45111][A45111]	72	56	1593
131	1	5197	2600	gi 153952	alanine transport protein - thermophilic acterium PS-3	72	53	2598
141	4	1040	1978	gi 1405446	polymerase III polymerase subunit (dnaB) [Salmonella typhimurium]	72	54	939
149	8	2819	2535	gi 606234	pir[A45915][A45915 DNA-directed DNA polymerase (EC 2.7.7.7)] III lpha chain - Salmonella typhimurium	72	44	285
149	17	5472	5245	gi 1304472	transketolase [Bacillus subtilis]	72	55	228
154	1	1	210	gi 1205620	secY [Escherichia coli]	72	40	210
155	1	2207	1320	gi 391610	DNA polymerase (Unidentified phycodnavirus clone OT04)	72	57	888
180	1	2	328	gi 133630	ferritin like protein [Haemophilus influenzae]	72	62	327
184	3	1145	3553	gi 1205110	[farnesyl diphosphate synthase [Bacillus stearothermophilus] pir[JX0257][JX0257 geranyltransferase (EC 2.5.1.10) - Bacillus stearothermophilus]	72	49	2409
195	2	1923	1279	gi 1001730	virulence associated protein homolog [Haemophilus influenzae]	72	45	645
206	13	14646	15969	gi 1064807	hypothetical protein [Synecocystis sp.]	72	50	1224
209	2	462	932	gi 1204666	hypothetical protein [Bacillus subtilis]	72	60	471
215	2	764	522	gi 881513	ORTHINE AMINOTRANSFERASE [Bacillus subtilis]	72	63	243
224	1	2	790	gi 949974	hypothetical protein (GB:X73124_53) [Haemophilus influenzae]	72	54	789
233	1	1526	765	gi 1408493	insulin receptor homolog [Staphylococcus xylosus]	72	52	762
240	1	220	1485	gi 537049	homologous to SwissProt:VIDA_EC01 hypothetical protein [Bacillus subtilis]	72	52	1266
245	1	3	1340	gi 1204578	[ORF_0470 [Escherichia coli]	72	46	1338
					hypothetical protein (GB:U06949_1) [Haemophilus influenzae]	72		

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
259	2	2108	1245	gi 1340128	ORF1 [Staphylococcus aureus]	72	59	864
304	2	285	1094	gi 1205330	glutamine-binding periplasmic protein (Haemophilus influenzae)	72	52	810
307	110	5326	5039	gi 1070015	protein-dependent (Bacillus subtilis)	72	53	288
315	1	517	260	gi 143399	quinol oxidase (Bacillus subtilis)	72	55	258
316	111	9622	9308	gi 1204445	hypothetical protein (SP27857) (Haemophilus influenzae)	72	58	315
317	3	926	1609	gi 487433	citrate synthase II (Bacillus subtilis)	72	59	684
364	7	12538	10493	gi 1510643	ferrous iron transport protein B (Methanococcus jannaschii)	72	53	2046
409	2	340	1263	gi 1402944	orfH1 gene product (Bacillus subtilis)	72	49	924
441	3	2177	1590	gi 312379	highly conserved among subacteria (Clostridium acetobutylicum) pif[S34312]S34312 hypothetical protein V - Clostridium acetobutylicum	72	48	588
453	6	2654	2505	pir[S00601]BXSA	antibacterial protein 3 - Staphylococcus haemolyticus	72	70	150
460	1	2	625	gi 1016162	ABC transporter subunit (Cyanophora paradoxa)	72	51	624
463	1	3253	1628	gi 666014	The polymorphism (RFLP) of this gene is associated with susceptibility to essential hypertension. The SA gene product has light homology to acetyl-CoA synthetase (Homo sapiens)	72	60	1626
480	4	3047	3466	gi 433992	ATP synthase subunit epsilon (Bacillus subtilis)	72	53	420
502	1	1086	586	gi 310859	ORF2 [Synthococcus sp.]	72	50	501
519	1	81	1184	gi 1303704	YrkE (Bacillus subtilis)	72	54	1104
559	1	3	746	gi 1107530	icaUD gene product (Campylobacter coli)	72	56	744
575	1	1142	573	gi 1303866	YqgS (Bacillus subtilis)	72	56	570
671	1	2	592	gi 1204897	protein-export membrane protein (Haemophilus influenzae)	72	44	591
679	2	295	1251	gi 563258	virulence-associated protein E (Dichelobacter nodosus)	72	52	957
687	2	295	957	gi 1146214	44% identical amino acids with the Escherichia coli smba suprase; putative (Bacillus subtilis)	72	49	663
837	1	1	435	gi 1146183	putative (Bacillus subtilis)	72	54	435
868	1	150	788	gi 1377842	unknown (Bacillus subtilis)	72	55	639
922	1	130	432	gi 1088269	unknown protein (Azotobacter vinelandii)	72	58	303
941	1	2	238	gi 153929	[NADPH-sulfite reductase flavoprotein component (Salmonella typhimurium)]	72	49	237
980	1	840	421	gi 853767	[UDP-N-acetylglucosamine 1-carboxyvinyltransferase (Bacillus subtilis)]	72	59	420

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1209	1	381	213	gi1144735	neurotoxin type B (Clostridium botulinum)	72	44	171
1469	2	671	474	gi11205458	hypothetical protein (CB.D2562_47) (Haemophilus influenzae)	72	63	198
1956	1	727	365	gi1154409	hexosephosphate transport protein (Salmonella typhimurium) p1r181853 hexose phosphate transport system regulatory protein uhpa - Salmonella typhimurium	72	44	363
2101	1	3	401	gi11303950	VqIV (Bacillus subtilis)	72	50	399
2503	1	569	399	gi1149713	formate dehydrogenase (Methanobacterium formicicum) p1r142712 formate dehydrogenase (EC 1.2.1.2) - ethanobacterium formicicum	72	56	171
2967	1	3	155	gi11212729	VqH3 (Bacillus subtilis)	72	46	153
3004	1	367	185	gi11665999	hypothetical protein (Bacillus subtilis)	72	55	183
3109	1	278	141	gi11413968	lpa-44d gene product (Bacillus subtilis)	72	45	138
3171	1	3	287	gi11515938	glutamate synthase (ferredoxin) (Synecocystis sp.) p1r1546957 glutamate synthase (ferredoxin) (EC 1.4.1.1) - synecocystis sp.	72	52	285
3771	1	26	367	gi11408501	homologous to N-acetyl-L-amino acid amidohydrolase of Bacillus stearothermophilus (Bacillus subtilis)	72	63	342
3951	1	1	222	gi11500409	M. jannaschii predicted coding region MJ1519 (Methanococcus jannaschii)	72	38	222
4190	1	721	362	gi1139956	IIIGC (Bacillus subtilis)	72	57	360
4444	1	3	347	gi11009366	respiratory nitrate reductase (Bacillus subtilis)	72	55	345
6	2	911	1200	gi11537095	amylolysin carboxymethyltransferase (Escherichia coli)	71	54	270
11	15	11350	10859	gi11533309	25 kDa protein (Escherichia coli)	71	47	492
19	2	1248	2435	gi11244574	D-alanine:D-alanine ligase (Enterococcus hirae)	71	52	1188
21	2	898	1488	gi1149629	anthranilate synthase component 2 (Leptospira biflexa) p1r132840 anthranilate synthase (EC 4.1.3.27) component II Leptospira biflexa	71	45	591
34	1	1	567	gi11303983	VqKF (Bacillus subtilis)	71	59	567
37	3	3192	2806	gi11209681	glutamate-rich protein (Bacillus firmus)	71	50	387
38	18	12250	12462	gi11927645	arginyl endopeptidase (Porphyromonas gingivalis)	71	50	213
39	3	1246	4431	gi11509411 S094	spoIIIE protein - Bacillus subtilis	71	49	3186
53	14	15770	14760	gi1142611	branched chain alpha-keto acid dehydrogenase E1-alpha (Bacillus subtilis)	71	58	1011
54	11	13461	12625	gi1143014	lgt repressor (Bacillus subtilis)	71	46	837

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
57	7	7152	5860	gi 508175	ELIC domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	71	48	1293
57	18	13897	14334	gi 1063247	high homology to flavohemoprotein (haemoglobin-like protein) of Alcaligenes eutrophus and Saccharomyces cerevisiae [Bacillus subtilis]	71	56	438
62	16	9831	10955	gi 1303926	YqjO [Bacillus subtilis]	71	54	1125
70	12	8505	8966	gi 147198	lpnE protein [Escherichia coli]	71	38	462
86	5	2394	2089	gi 904205	hypothetical protein [Bacillus subtilis]	71	51	306
96	7	7601	8269	gi 709991	hypothetical protein [Bacillus subtilis]	71	49	669
100	6	4822	5931	gi 1060848	Opine dehydrogenase [Arthrobacter sp.]	71	45	1110
103	1	1062	532	gi 143089	lep protein [Bacillus subtilis]	71	41	531
109	18	15312	15695	gi 413985	tpa-61d gene product [Bacillus subtilis]	71	57	384
113	1	630	316	gi 563254	probable protein kinase [Saccharomyces cerevisiae]	71	57	315
114	5	6598	5603	gi 143156	membrane bound protein [Bacillus subtilis]	71	40	996
133	2	3087	1723	gi 1303913	YqjX [Bacillus subtilis]	71	53	1365
149	19	6335	5895	gi 529650	GADP [Bacteriophage SP21]	71	51	441
154	5	3635	3087	gi 425488	repressor protein [Streptococcus sobrinus]	71	47	549
164	11	11354	11609	gi 49318	ORF4 gene product [Bacillus subtilis]	71	52	316
169	5	1936	2745	gi 1403403	unknown [Mycobacterium tuberculosis]	71	56	810
193	2	272	1234	gi 1303788	YqjH [Bacillus subtilis]	71	49	963
205	1	1743	895	gi 1215694	GlnQ [Mycoplasma pneumoniae]	71	46	849
233	4	1849	2022	gi 633732	ORF1 [Campylobacter jejuni]	71	50	174
237	7	4501	5169	gi 149384	HisE [Lactococcus lactis]	71	54	669
272	4	2848	2273	gi 709993	hypothetical protein [Bacillus subtilis]	71	48	576
274	2	618	1496	gi 143035	NAD(P)H:glutaryl-transfer RNA reductase [Bacillus subtilis] pir(A)5252(A)5252 5-aminolevulinate synthase (EC 2.3.1.37) - acillus subtilis	71	53	879
276	5	3349	2720	gi 303562	ORF210 [Escherichia coli]	71	50	630
287	1	116	660	gi 110634	20 kDa protein [Streptococcus gordonii]	71	53	525
288	6	3322	2771	gi 1256625	putative [Bacillus subtilis]	71	47	552

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
301	6	3492	2461	gi1467417	similar to lysine decarboxylase [Bacillus subtilis]	71	57	1032
306	4	6607	5222	gi1256618	transport protein [Bacillus subtilis]	71	56	1386
307	2	1536	925	gi1602683	lorC [Mycoplasma capricolum]	71	45	612
310	5	5793	5146	gi1348052	acetoin utilization protein [Bacillus subtilis]	71	51	648
322	1	2	1303	gi11001819	hypothetical protein [Synchocystis sp.]	71	46	1302
333	4	4171	3995	gi1467473	unknown [Bacillus subtilis]	71	57	177
350	2	548	922	gi1551879	ORF 1 [Lactococcus lactis]	71	55	375
375	4	1860	3071	gi1467447	unknown [Bacillus subtilis]	71	57	1212
380	5	1560	2102	gi1142557	ATP synthase b subunit [Bacillus megaterium]	71	43	543
414	2	251	637	gi1580904	homologous to E.coli rnpA [Bacillus subtilis]	71	49	387
424	1	335	1354	gi1581305	L-lactate dehydrogenase [Lactobacillus plantarum]	71	57	1020
436	4	3701	3270	gi1580501 PNO5	phosphoribosylanthranilate isomerase (EC 5.3.1.24) - Bacillus subtilis (fragment)	71	66	432
482	1	3	1280	gi1410142	ORFX18 [Bacillus subtilis]	71	49	1278
525	3	2272	1844	gi1143370	phosphoribosylpyrophosphate amidotransferase (Pur-P; EC 2.4.2.14) Bacillus subtilis	71	56	429
529	4	2739	2047	gi1606150	ORF_1309 [Escherichia coli]	71	41	693
563	1	22	969	gi11237015	ORF4 [Bacillus subtilis]	71	53	948
581	1	506	255	gi11301730	T2503.2 [Caenorhabditis elegans]	71	47	252
612	2	1088	913	gi1153968	fimbriae 2 [Salmonella typhimurium]	71	55	156
613	1	1	654	gi1466778	lysine specific permease [Escherichia coli]	71	50	654
618	1	1243	623	gi1146238	poly(A) polymerase [Bacillus subtilis]	71	52	621
630	1	1170	586	gi11486243	unknown [Bacillus subtilis]	71	53	585
691	1	1126	641	gi1289260	comE ORF1 [Bacillus subtilis]	71	51	486
694	2	149	427	gi112971	NADH dehydrogenase subunit V (AA 1-605) [Gallus gallus] ir S10197 S10197 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain - chicken mitochondrion (LOC1)	71	47	279
715	2	169	777	gi11303830	rgfL [Bacillus subtilis]	71	53	609
746	2	1473	970	gi11377843	unknown [Bacillus subtilis]	71	52	504

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
748	1	1437	802	gi11405459	[YneS [Bacillus subtilis]	71	49	636
753	1	1018	524	gi11510389	[H. jannaschii predicted coding region M0296 [Methanococcus jannaschii]	71	53	495
761	1	3	215	gi1475972	[pentafunctional enzyme [Pneumocystis carinii]	71	47	213
783	1	1203	703	gi1536655	[ORF YBR244w [Saccharomyces cerevisiae]	71	52	501
800	3	1292	987	gi11204326	[tRNA delta(2)-isopentenylpyrophosphate transferase [Haemophilus influenzae]	71	48	306
806	1	116	286	gi11419075	[cbiM gene product [Methanobacterium thermoautotrophicum]	71	50	171
931	1	973	488	gi1893358	[PgsA [Bacillus subtilis]	71	56	485
1041	1	2	262	gi11408507	[pyrimidine nucleoside transport protein [Bacillus subtilis]	71	45	261
1070	1	2	172	gi1709993	[hypothetical protein [Bacillus subtilis]	71	46	171
1176	1	57	365	gi1151259	[HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mvaloni] p1r[A44756[A44756 Pseudomonas sp. hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88)]	71	49	309
1181	1	366	184	gi146971	[epiP gene product [Staphylococcus epidermidis]	71	50	183
1281	1	3	290	gi1153016	[ORF 419 protein [Staphylococcus aureus]	71	50	289
1148	1	456	229	gi1602683	[orfC [Mycoplasma capricolum]	71	48	228
2002	1	756	379	gi11008177	[ORF YJL046w [Saccharomyces cerevisiae]	71	48	378
2119	1	2	217	gi11046088	[arginyl-tRNA synthetase [Mycoplasma genitalium]	71	50	216
2418	1	3	320	gi11499771	[H. jannaschii predicted coding region M0936 [Methanococcus jannaschii]	71	57	318
2961	1	2	187	gi1312443	[carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus aldyticus]	71	57	186
2999	2	67	306	gi1710020	[nitrite reductase (nirB) [Bacillus subtilis]	71	43	240
3033	1	2	184	gi11262335	[YnaA [Bacillus subtilis]	71	57	183
3584	1	3	338	gi1401716	[beta-isopropylmalate dehydrogenase [Neurospora crassa]	71	55	336
3715	2	743	399	gi1563952	[gluconate permease [Bacillus licheniformis]	71	59	345
3785	1	770	387	gi147382	[acyl-CoA-dehydrogenase [Streptomyces purpurascens]	71	57	384
3875	1	541	272	gi11001541	[hypothetical protein [Synecococcus sp.]	71	38	270
4135	1	637	320	gi1142695	[S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase [Bacillus megaterium]	71	52	318
4249	1	63	239	gi11205363	[deoxyribose aldolase [Haemophilus influenzae]	71	63	177
4508	1	530	267	gi11197667	[vitellogenin [Anolis pulchellus]	71	46	264

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
6	3	1237	2721	gi1321788	arginine ornithine antiporter [Clostridium perfringens]	70	54	1485
11	11	6572	7486	gi1216854	p47K [Pseudomonas chlororaphis]	70	41	915
12	1	2890	1481	gi1467330	replicative DNA helicase [Bacillus subtilis]	70	49	1410
15	1	1756	893	gi1451216	mannosephosphate isomerase [Streptococcus mutans]	70	46	864
15	2	1277	1050	gi1476092	unknown [Bacillus subtilis]	70	50	228
17	2	2132	1350	gi1145402	choline dehydrogenase [Escherichia coli]	70	52	783
21	1	2	925	gi1149516	anthranilate synthase alpha subunit [Lactococcus lactis] p1r/S3124 anthranilate synthase (EC 4.1.3.27) alpha chain - lactococcus lactis subsp. lactis	70	50	924
25	7	5580	6251	gi1389549	ORF3 [Bacillus subtilis]	70	52	672
33	6	6071	7423	gi1301875	YqhB [Bacillus subtilis]	70	51	1353
36	2	959	1594	gi1500755	methyl purine glycosylase [Mus musculus]	70	47	636
38	8	4901	5860	gi11408507	pyrimidine nucleoside transport protein [Bacillus subtilis]	70	44	960
44	8	5312	5989	gi11006620	hypothetical protein [Synecocystis sp.]	70	49	678
46	10	8950	10020	gi1403126	cscD gene product [Alcaligenes eutrophus]	70	45	1071
52	2	2727	1900	gi11486247	unknown [Bacillus subtilis]	70	53	828
52	6	4048	4656	gi1244501	esterase II-carboxylesterase (EC 3.1.1.1) [Pseudomonas fluorescens, eptido, 218 aa]	70	50	609
56	8	8460	9962	gi1333951	small subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	70	51	1503
62	1	48	290	gi1142702	A competence protein 2 [Bacillus subtilis]	70	47	243
64	1	1080	541	gi13204377	molybdopterin biosynthesis protein [Haemophilus influenzae]	70	47	540
70	5	5139	3595	gi13204834	2',3'-cyclic-nucleotide 2'-phosphodiesterase [Haemophilus influenzae]	70	47	1545
91	4	7793	5466	gi1886471	methionine synthase [Catharantus roseus]	70	56	2328
96	5	8754	7255	gi1390956	alkaline phosphatase (EC 3.1.3.1) III precursor - Bacillus subtilis	70	54	1500
110	2	767	1300	gi1145294	adenine phosphoribosyl-transferase [Escherichia coli]	70	51	534
116	6	7026	7976	gi1143607	sporulation protein [Bacillus subtilis]	70	50	951
121	8	6401	6988	gi1107528	ttg start [Campylobacter coli]	70	45	588
131	8	6842	7936	gi11150454	proliferase PcpQ [Lactobacillus delbrueckii]	70	48	1095

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
135	1	2	1489	gi 311309	putative membrane-bound protein with four times repetition of ro-Ser-Ala at the N-terminus; function unknown (Alcaligenes utrophus)	70	49	1488
138	3	418	714	gi 504181	hypothetical protein (Bacillus subtilis)	70	46	297
164	8	9344	9874	gi 49315	ORF1 gene product (Bacillus subtilis)	70	47	531
164	16	15826	16618	gi 1205212	hypothetical protein (GB:010483_18) (Haemophilus influenzae)	70	50	993
205	2	2735	1803	gi 1215695	peptide transport system protein SapF homolog; SapF homolog (Mycoplasma pneumoniae)	70	47	933
209	3	910	1386	gi 1204665	hypothetical protein (GB:X73124_26) (Haemophilus influenzae)	70	48	477
246	3	340	756	gi 215098	excisionase (Bacteriophage 154a)	70	46	417
263	7	7876	6749	gi 142540	aspartokinase II (Bacillus sp.)	70	51	1128
268	3	3212	4117	gi 1340128	ORF1 (Staphylococcus aureus)	70	50	906
302	6	3201	3827	gi 147782	ruvA protein (atg start) (Escherichia coli)	70	46	627
302	10	5879	7051	pir C18530 C385	queuine tRNA-ribosyltransferase (EC 2.4.2.29) - Escherichia coli	70	55	1173
313	1	2520	1414	gi 1205934	aminopeptidase a/i (Haemophilus influenzae)	70	46	1107
355	2	379	669	gi 1070013	protein-dependent (Bacillus subtilis)	70	48	291
403	1	1255	629	gi 733147	GuaF (Xanthomonas campestris)	70	31	627
444	10	8770	9273	gi 1204752	high affinity ribose transport protein (Haemophilus influenzae)	70	52	504
449	1	2	1243	gi 619724	HgtG (Bacillus firmus)	70	44	1242
472	1	637	320	gi 727145	open reading frame; putative (Bacillus amyloliquefaciens) pir[329091]329091 hypothetical protein (bglA region) - Bacillus myoliquefaciens (fragment)	70	41	318
480	2	727	1608	gi 142560	ATP synthase gamma subunit (Bacillus megaterium)	70	44	882
524	1	2	307	gi 602392	RCN2 protein (Brassica napus)	70	45	306
525	1	823	413	gi 143372	phosphoribosyl glycinate formyltransferase (PUR-N) (Bacillus subtilis)	70	52	411
565	4	1625	2552	gi 881434	ORFP (Bacillus subtilis)	70	51	1074
607	4	829	1284	gi 151524	hypothetical protein (SP:P37002) (Methanococcus jannaschii)	70	50	456
633	1	1383	703	gi 431231	lucell permease (Bacillus caldolyticus)	70	53	681
646	3	1683	1309	gi 467340	unknown (Bacillus subtilis)	70	49	375
663	1	830	417	gi 1303873	Y-172 (Bacillus subtilis)	70	40	414

TABLE 2

E. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
681	1	1488	781	gi 1001678	hypothetical protein (Synchocystis sp.)	70	53	708
708	1	2	448	sp P33940 VOXH_	HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION.	70	51	447
725	1	51	722	gi 1001644	hypothetical protein (Synchocystis sp.)	70	48	672
776	1	1371	787	gi 145165	putative [Escherichia coli]	70	47	585
834	1	250	783	gi 552971	NADH dehydrogenase (ndhF) [Vicia faba]	70	47	534
865	2	1585	1179	gi 1204636	ATP-dependent helicase [Haemophilus influenzae]	70	45	207
894	1	535	269	gi 467364	DNA binding protein (probae) [Bacillus subtilis]	70	41	267
919	1	3	317	gi 1314847	ClnA [Bacillus subtilis]	70	40	315
944	1	3	572	gi 709991	hypothetical protein [Bacillus subtilis]	70	44	570
988	2	772	605	gi 122441	ORF 3; putative [Bacillus subtilis]	70	50	168
1055	1	3	335	gi 529755	lepac [Streptococcus pyogenes]	70	37	333
1093	1	2	904	gi 853754	ABC transporter [Bacillus subtilis]	70	49	903
1109	1	2	310	gi 1001827	hypothetical protein [Synchocystis sp.]	70	42	309
1220	1	468	235	pir s23416 s234	epib protein - Staphylococcus epidermidis	70	40	234
1279	1	73	348	gi 153015	FemA protein [Staphylococcus aureus]	70	47	276
1336	1	195	542	sp P31776 PBPA_	PENICILLIN-BINDING PROTEIN 1A (PBP-1A) [PENICILLIN-BINDING PROTEIN A]	70	50	348
1537	2	232	402	gi 1146181	putative [Bacillus subtilis]	70	50	171
1574	1	451	272	gi 219630	endothelin-A receptor [Homo sapiens]	70	47	180
1640	1	690	346	gi 1146243	22.4k identity with Escherichia coli DNA-damage inducible protein ...	70	46	345
2504	1	2	286	gi 1495179	transmembrane protein [Lactococcus lactis]	70	51	285
3061	1	564	301	gi 508175	EIC domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	70	44	284
3128	1	2	199	gi 1340096	unknown [Mycobacterium tuberculosis]	70	51	198
3218	1	3	488	gi 1515938	glutamate synthase (ferredoxin) [Synchocystis sp.] pir s46957 s46957	70	50	486
3323	1	794	399	gi 1154891	glutamate synthase (ferredoxin) [EC 1.4.7.1] - synchocystis sp.	70	52	396
3679	1	599	399	gi 529385	ATP binding protein [Phormidium laminosum]	70	30	201
					chromosome condensation protein [Caenorhabditis elegans]			

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
3841	1	706	398	gi 1208965	hypothetical 23.3 kd protein [Escherichia coli]	70	47	309
3929	1	3	401	gi 149435	putative [Lactococcus lactis]	70	49	399
4044	1	595	374	gi 602031	similar to trimethylamine DH [Mycoplasma capricolum] p1r 549950 549950 probable trimethylamine dehydrogenase [EC 5.99.7] - Mycoplasma capricolum (SGC3) (fragment)	70	40	222
4129	1	558	280	gi 1339951	small subunit of NADH-dependent glutamate synthase [Electronema boryanum]	70	49	279
4422	1	576	289	gi 296464	ATPase [Lactococcus lactis]	70	57	288
4647	1	361	200	gi 166412	NADH-glutamate synthase [Medicago sativa]	70	59	162
16	8	7571	9031	gi 1499620	M. Jannaschii predicted coding region M30798 [Methanococcus jannaschii]	69	44	1461
16	9	9080	10033	gi 1353197	thioredoxin reductase [Eubacterium acidaminophilum]	69	54	954
30	1	1452	727	gi 1204910	hypothetical protein (G8:U14003_302) [Haemophilus influenzae]	69	52	726
38	4	1023	1298	gi 407773	dva gene product [Anabaena sp.]	69	41	276
44	9	5987	6595	gi 1205920	molybdate uptake system hydrophilic membrane-bound protein [Haemophilus influenzae]	69	45	609
62	15	9104	9475	gi 385178	unknown [Bacillus subtilis]	69	44	372
66	4	2402	2803	gi 1303893	YqH [Bacillus subtilis]	69	51	402
67	15	114124	11627	gi 149647	ORF2 [Listeria monocytogenes]	69	37	498
67	17	14053	14382	gi 305002	ORF_335 [Escherichia coli]	69	49	330
67	19	15130	15807	gi 1109684	ProV [Bacillus subtilis]	69	45	678
78	3	1447	2124	gi 1256633	putative [Bacillus subtilis]	69	53	678
78	4	4513	3725	gi 1103958	YqJG [Bacillus subtilis]	69	32	789
85	4	4521	4213	pir E29326 E293	hypothetical protein (pur operon) - Bacillus subtilis	69	32	309
86	6	3253	2654	gi 973332	OrfC [Bacillus subtilis]	69	50	600
95	1	96	710	gi 766468	4A11 antigen, sperm tail membrane antigen-putative sucrose-specific phosphotransferase enzyme II homolog [mice, testis, Peptide Partial, 72 aa]	69	43	615
100	7	6023	7426	gi 1205355	[Na+/H+ antiporter [Haemophilus influenzae]	69	39	1404
102	2	2678	1650	gi 561690	Leleoglycoprotease [Pasteurella haemolytica]	69	47	1029
103	8	12241	8537	gi 1009366	Respiratory nitrate reductase [Bacillus subtilis]	69	54	3705
103	11	14987	12552	gi 710020	Nitrite reductase (nirB) [Bacillus subtilis]	69	51	2436

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
112	11	8708	10168	gi 154111	hexosephosphate transport protein [Salmonella typhimurium] pfr 01853 01853 hexose phosphate transport system protein uhpT - salmonella typhimurium	69	51	1461
112	16	16644	17414	gi 1204435	pyruvate formate-lyase activating enzyme [Haemophilus influenzae]	69	50	771
113	2	33	953	gi 1290509	ec07 [Escherichia coli]	69	43	921
114	2	1537	1058	gi 142771 1427	reticulocyte-binding protein 1 - Plasmodium vivax	69	39	480
121	6	4309	5310	gi 154633	hrdF [Bacillus subtilis]	69	53	1002
125	2	267	854	gi 113931	lpa-7d gene product [Bacillus subtilis]	69	43	588
149	27	10666	10400	gi 528089 5280	hypothetical protein A - yeast [Zygosaccharomyces bisporus] plasmid pSD1	69	39	267
161	1	1598	813	gi 1205538	hypothetical protein (GB:U14003_302) [Haemophilus influenzae]	69	47	786
165	4	2222	4633	gi 140054	phenylalanyl-tRNA synthetase beta subunit (AA 1-804) [Bacillus btillia]	69	52	2412
169	3	1210	1761	gi 1296031	elongation factor Ts [Spirulina platensis]	69	45	552
175	12	8686	8339	gi 732682	Fine protein [Escherichia coli]	69	69	348
190	2	484	1671	gi 17731 HIS8	HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (HIDAZOLE ACETOL-PHOSPHATE TRANSAMINASE)	69	48	1188
206	1	5551	2777	gi 41750	hadr protein (AA 1-1033) [Escherichia coli]	69	49	2775
206	4	6038	5796	gi 1256135	YbbF [Bacillus subtilis]	69	48	243
249	1	636	319	gi 1405456	YnuP [Bacillus subtilis]	69	50	118
302	8	4820	5776	gi 1001768	hypothetical protein [Synechocystis sp.]	69	48	957
324	2	7384	3893	gi 1256798	pyruvate carboxylase [Rhizobium etli]	69	53	3492
351	3	2098	1808	gi 1491664	T04H1.4 [Caenorhabditis elegans]	69	30	291
369	3	2075	2305	gi 136458	ORF [Balaenoptera acutorostrata]	69	61	231
392	3	1999	2424	gi 1556015	ORF1 [Bacillus subtilis]	69	45	426
410	1	87	779	gi 155611	phosphoglyceromutase [Zymomonas mobilis]	69	58	693
421	1	2085	1129	gi 1276985	arginase [Bacillus caldovelox]	69	54	957
444	8	6713	7741	gi 1221782	purine synthesis repressor [Haemophilus influenzae]	69	40	1029
453	1	828	415	gi 1122758	unknown [Bacillus subtilis]	69	57	414
469	2	3286	2246	gi 1458228	mutY homolog [Homo sapiens]	69	44	1041

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
509	3	1730	1371	gi 49224	URF 4 [Synachococcus sp.]	69	39	360
520	5	3023	2823	gi 726427	similar to D. melanogaster NS101-2 protein (P18:S34154) Caenorhabditis elegans	69	39	201
531	1	26	760	gi 509672	repressor protein (Bacteriophage Tuc2009)	69	33	735
589	1	107	253	gi 169101	17.9 kDa heat shock protein (hsp17.9) [Plum sativum]	69	52	147
594	2	597	1391	gi 142783	DNA photolyase [Bacillus firmus]	69	48	795
604	4	2476	2114	gi 411910	lpa-6d gene product [Bacillus subtilis]	69	45	361
607	1	2	313	gi 1236103	W0802.3 [Caenorhabditis elegans]	69	47	312
607	2	590	312	gi 536715	ORF Y8275c [Saccharomyces cerevisiae]	69	39	279
734	1	864	433	gi 467327	unknown [Bacillus subtilis]	69	44	432
759	1	3	338	gi 1009367	Respiratory nitrate reductase [Bacillus subtilis]	69	50	336
761	2	392	586	gi 1508	Leucyl-tRNA synthetase (cytoplasmic) [Saccharomyces cerevisiae] J370340 ORF YPL160w [Saccharomyces cerevisiae]	69	46	195
802	1	72	1013	gi 143044	ferrochelatase [Bacillus subtilis]	69	55	942
816	1	2573	1368	gi 1510268	restriction modification system S subunit [Methanococcus jannaschii]	69	45	1206
838	2	133	387	gi 1255371	coded for by C. elegans cDNA YK349.5; coded for by C. elegans cDNA YK349.1; Similar to guanylate kinase [Caenorhabditis elegans]	69	46	255
851	2	745	1005	gi 188998	sacA gene product [Antihannion sp.]	69	39	261
867	1	535	269	gi 1070014	protein-dependent [Bacillus subtilis]	69	47	267
995	1	934	478	gi 1205569	transcription elongation factor [Haemophilus influenzae]	69	53	477
999	1	1009	506	gi 899254	predicted trithorax protein [Drosophila virilis]	69	21	504
1127	1	1315	659	gi 1205434	[H. influenzae predicted coding region H1191 [Haemophilus influenzae]	69	56	657
1138	1	248	460	gi 1510646	[M. jannaschii predicted coding region M20568 [Methanococcus jannaschii]	69	48	213
2928	1	3	401	gi 120503	glutamate permease [Escherichia coli]	69	41	399
3090	1	444	223	gi 1204987	DNA polymerase III, alpha chain [Haemophilus influenzae]	69	36	222
3817	1	2	400	gi 1483199	peptide-synthetase [Mycobacterium mediterranei]	69	45	399
3833	1	667	335	gi 1524193	unknown [Mycobacterium tuberculosis]	69	46	333

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
4079	1	747	400	gi 546918	orfy 3' of comK [Bacillus subtilis, E26, Peptide Partial, 140 aa] p1r[S3612]S3612 hypothetical protein Y - Bacillus subtilis sp P40398 YHND_BACSU HYPOTHETICAL PROTEIN IN COMK 3' REGION (ORFY) FRAGMENT)	69	64	348
4115	2	215	400	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen [Streptococcus pyogenes]	69	59	186
4139	1	1	333	gi 1208451	hypothetical protein [Synechocystis sp.]	69	36	333
4258	1	457	230	gi 496158	restriction-modification enzyme subunit M1 [Mycoplasma pulmonis] p1r[S49395]S49395 HsdM1 protein - Mycoplasma pulmonis (SGC3)	69	43	228
4317	1	90	374	gi 413967	lpa-43D gene product [Bacillus subtilis]	69	44	285
4465	1	3	293	gi 396296	similar to phosphotransferase system enzyme II [Escherichia coli] sp P32672 PTWC_ECOLI PTS SYSTEM, FRUCTOSE-LIKE-2 IIC COMPONENT PHOSPHOTRANSFERASE ENZYME II, C COMPONENT)	69	49	291
3	1	2302	1193	gi 1109685	ProW [Bacillus subtilis]	68	46	1110
15	4	2592	2074	gi 807973	unknown [Saccharomyces cerevisiae]	68	45	519
31	8	6328	8772	gi 290642	ATPase [Enterococcus hirae]	68	48	2445
40	2	1115	750	gi 606342	ORF_0522; reading frame open far upstream of start; possible ranseshift. linking to previous ORF [Escherichia coli]	68	55	166
46	9	6886	8415	gi 155276	aldehyde dehydrogenase [Vibrio cholerae]	68	44	1530
48	3	3643	3404	gi 285608	241k polypeptide (Apple stem grooving virus)	68	47	240
48	4	3536	4132	gi 1045937	M. genitalium predicted coding region M0246 [Mycoplasma genitalium]	68	39	597
53	10	11671	10685	gi 1303952	YqJA [Bacillus subtilis]	68	46	987
70	9	7346	8155	gi 147198	phnE protein [Escherichia coli]	68	40	810
89	4	1899	2966	gi 145173	35 kDa protein [Escherichia coli]	68	43	1068
108	1	2187	1150	gi 38722	precursor (aa -20 to 381) [Acinetobacter calcoaceticus] ir A29277 A29277 aldose 1-epimerase (EC 5.1.3.3) - Acinetobacter calcoaceticus	68	57	1038
112	5	2666	3622	gi 153724	MalC [Streptococcus pneumoniae]	68	55	957
116	7	7865	8638	gi 143608	sporulation protein [Bacillus subtilis]	68	48	774
118	3	2484	3698	gi 1303805	YqeR [Bacillus subtilis]	68	46	1215
120	2	1424	1594	sp P38038 CVSJ_	SULFITE REDUCTASE (NADPH) FLAVOPROTEIN ALPHA-COMPONENT (EC 1.8.1.2) (SIR- FP)	68	45	171
129	1	1	1011	gi 396307	argininosuccinate lyase [Escherichia coli]	68	50	1011

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
132	3	1857	2739	gi 216267	ORF2 (Bacillus megaterium)	68	48	873
134	2	848	1012	gi 147545	DNA recombinase (Escherichia coli)	68	50	165
141	2	372	614	gi 1872116	eti (stress inducible protein) (Glycine max)	68	36	243
149	7	2454	2260	gi 145774	hsp70 protein (dnaK gene) (Escherichia coli)	68	48	195
155	2	1776	1534	gi 216583	ORF1 (Escherichia coli)	68	36	243
158	3	1826	3289	ap P31940 YOHJ	HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION.	68	51	1464
169	6	2749	3318	gi 1401402	unknown (Mycobacterium tuberculosis)	68	46	570
175	10	9158	7365	gi 1072395	phaA gene product (Rhizobium meliloti)	68	51	1794
188	7	4184	5434	gi 1173843	3-ketoacyl-ACP synthase II (Vibrio Harveyi)	68	48	1251
189	3	907	1665	gi 467393	DNA binding protein (probable) (Bacillus subtilis)	68	55	759
206	5	7683	6709	gi 1256138	YbbI (Bacillus subtilis)	68	48	975
206	8	10425	12176	gi 452687	pyruvate decarboxylase (Saccharomyces cerevisiae)	68	48	1752
212	8	3421	3648	gi 1369941	lci gene product (Bacteriophage B1)	68	39	228
214	8	5457	6482	gi 1420467	ORF YOR196c (Saccharomyces cerevisiae)	68	45	1026
237	4	2507	3088	gi 149381	HlsH (Lactococcus lactis)	68	46	582
243	5	5540	4542	gi 1235684	mevalonate pyrophosphate decarboxylase (Saccharomyces cerevisiae)	68	47	999
262	1	3	164	gi 150974	4-oxalocrotonate tautomerase (Pseudomonas putida)	68	42	162
262	2	1984	1116	gi 1147744	PSR (Enterococcus hirae)	68	49	867
276	6	3702	3139	ap P30750 ABC_E	ATP-BINDING PROTEIN ABC (FRAGMENT)	68	50	564
306	6	6345	5725	gi 1256617	adenine phosphoribosyltransferase (Bacillus subtilis)	68	53	621
333	3	4599	3850	gi 467473	unknown (Bacillus subtilis)	68	45	750
365	6	5017	4838	gi 1130643	T22B3.3 (Caenorhabditis elegans)	68	45	180
376	2	549	1646	gi 1277026	DAPA aminotransferase (Bacillus subtilis)	68	51	1098
405	1	1741	872	gi 1303917	YqjB (Bacillus subtilis)	68	47	870
406	2	853	539	gi 1511513	ABC transporter, probable ATP-binding subunit (Methanococcus jannaschii)	68	44	315
426	6	3558	3391	gi 624632	GltL (Escherichia coli)	68	48	168
438	1	108	329	gi 146923	nitrogenase reductase (Escherichia coli)	68	43	222

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
443	1	476	240	gi 535810	hippuricase (Campylobacter jejuni)	68	42	237
443	2	518	1015	gi 1204742	H. influenzae predicted coding region H10491 (Haemophilus influenzae)	68	48	498
443	5	4447	3779	gi 80660	deoxyribose-phosphate aldolase (Bacillus subtilis) [S49455]S49455 deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis	68	55	669
476	2	240	1184	gi 571345	unknown, similar to E. coli cardiolipin synthase (Bacillus subtilis) sp P45860 YWI_E_BACSU HYPOTHETICAL 58.2 PROTEIN IN NARI-ACDA INTERGENIC REGION.	68	45	945
486	2	1876	1046	gi 147328	transport protein (Escherichia coli)	68	41	831
517	3	1764	2084	gi 523809	orf2 (Bacteriophage A2)	68	64	321
572	1	2	571	sp P19237 Y05L	HYPOTHETICAL 6.8 KD PROTEIN IN NRDC-TK INTERGENIC REGION.	68	47	570
646	1	914	459	gi 413982	lta-58r gene product (Bacillus subtilis)	68	52	456
659	3	1668	1901	gi 1107541	C3J09.8 (Caenorhabditis elegans)	68	36	234
864	5	1510	1716	gi 145774	hsp70 protein (dnaK gene) (Escherichia coli)	68	48	207
920	1	860	432	gi 1140416	hypothetical protein (SP:P31466) (Methanococcus jannaschii)	68	54	429
952	1	1096	611	gi C33456	reductase (Leishmania major)	68	46	486
970	1	91	402	gi 1354775	pfoS/R (Treponema pallidum)	68	46	312
1028	1	1064	534	gi 410117	diaminopimelate decarboxylase (Bacillus subtilis)	68	47	531
1029	1	428	216	gi 1335714	Plasmodium falciparum mRNA for asparagine-rich antigen (clone 17C1) [Plasmodium falciparum]	68	31	213
1058	1	692	348	gi 181649	epic gene product (Staphylococcus epidermidis)	68	46	345
1096	2	665	465	gi 143434	Rho Factor (Bacillus subtilis)	68	43	201
1308	1	2	694	gi 116939	Group B oligopeptidase.PepB (Streptococcus agalactiae)	68	50	693
1679	1	2	238	gi 117205	167 kDa Myosin-crossreactive streptococcal antigen (Streptococcus yogenes)	68	53	237
2039	1	3	383	gi 153898	transport protein (Salmonella typhimurium)	68	51	381
2077	1	3	326	pir C33496 C334	hlaC homolog - Bacillus subtilis	68	47	324
2112	1	613	374	gi 64884	flavin LII (Xenopus laevis)	68	50	240
2273	1	793	398	gi 581648	lepB gene product (Staphylococcus epidermidis)	68	45	396
2948	1	2	385	gi 126869	branched-chain amino acid transport carrier (Pseudomonas aeruginosa) pir A38514 A38514 branched-chain amino acid transport protein braz Pseudomonas aeruginosa	68	41	384

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
2955	1	768	gi 1904179	hypothetical protein [Bacillus subtilis]	68	49	369
2981	1	572	gi 508979	CRP-binding protein [Bacillus subtilis]	68	48	285
3014	1	584	gi 1324194	ORF-2 upstream of gbaB operon [Bacillus subtilis]	68	45	291
3082	1	336	gi 1204696	fructose-permease IIC component [Haemophilus influenzae]	68	53	168
3108	1	103	gi 217855	heat-shock protein [Arabidopsis thaliana]	68	48	156
3639	1	919	gi 1510490	nitrate transport permease protein [Methanococcus jannaschii]	68	47	459
3657	1	330	gi 155369	PTS enzyme-II fructose [Xanthomonas campestris]	68	48	310
3823	1	780	gi 603768	HutI protein, imidazolone-5-propionate hydrolase [Bacillus subtilis]	68	54	390
3982	1	277	gi 149435	putative [Lactococcus lactis]	68	47	276
4051	1	342	gi 450688	hadM gene of Ecoprr1 gene product [Escherichia coli] pfr[S38437/S38437] hsdM protein - Escherichia coli pfr[S09629/S09629] hypothetical protein A - Escherichia coli (SUB 40-520)	68	48	342
4089	1	12	gi 1353678	heavy-metal transporting P-type ATPase [Proteus mirabilis]	68	47	198
4143	1	47	gi 603769	HutU protein, urocanase [Bacillus subtilis]	68	55	141
4148	1	2	gi 450688	hadM gene of Ecoprr1 gene product [Escherichia coli] pfr[S38437/S38437] hsdM protein - Escherichia coli pfr[S09629/S09629] hypothetical protein A - Escherichia coli (SUB 40-520)	68	51	351
4173	1	2	gi 1041097	Pyruvate Kinase [Bacillus psychrophilus]	68	48	381
4182	1	498	gi 413968	lipa-44d gene product [Bacillus subtilis]	68	50	249
4362	2	148	gi 450688	hadM gene of Ecoprr1 gene product [Escherichia coli] pfr[S38437/S38437] hsdM protein - Escherichia coli pfr[S09629/S09629] hypothetical protein A - Escherichia coli (SUB 40-520)	68	44	171
5	11	9493	gi 183727	putative [Bacillus subtilis]	67	46	1194
31	11	10318	gi 216746	D-lactate dehydrogenase [Lactobacillus plantarum]	67	41	486
32	3	1560	gi 1098557	renal sodium/dicarboxylate cotransporter [Homo sapiens]	67	46	1596
32	5	4945	gi 1510720	prephenate dehydratase [Methanococcus jannaschii]	67	51	801
36	5	5350	gi 146216	45% identity with the product of the ORF6 gene from the Erwinia herbicola carotenoid biosynthesis cluster; putative [Bacillus subtilis]	67	58	1083
44	7	4492	gi 1006621	hypothetical protein [Synechocystis sp.]	67	43	813

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
56	7	3943	8481	gi 394131	glutamate synthase large subunit precursor [Azospirillum brasilense] pIR 946602 946602 glutamate synthase (NADPH) (EC 1.4.1.13) alpha chain - Azospirillum brasilense	67	52	4539
56	12	13923	14678	gi 1000453	TrkB [Bacillus subtilis]	67	48	756
62	8	5092	4757	gi 1113949	orf3 [Bacillus, C-125, altali-sensitive mutant 18224, Peptide Mutant, 112 aa]	67	45	336
62	10	7570	6338	gi 1054655	Na/H antiporter system [Bacillus alcalophilus]	67	49	1233
99	3	2119	3321	gi 1204349	hypothetical protein [GB:D90212_3] [Haemophilus influenzae]	67	50	1203
102	9	5695	7176	gi 149432	putative [Lactococcus lactis]	67	51	1482
103	13	14549	14049	gi 1408497	LP9D gene product [Bacillus subtilis]	67	48	501
109	15	14821	13982	gi 413976	lpa-52r gene product [Bacillus subtilis]	67	49	840
109	17	14811	15194	gi 413983	lpa-59d gene product [Bacillus subtilis]	67	29	384
121	4	1713	2153	gi 1262335	YneA [Bacillus subtilis]	67	54	441
122	1	1	1149	gi 143047	ORFB [Bacillus subtilis]	67	35	1149
124	5	4060	3518	gi 556885	Unknown [Bacillus subtilis]	67	47	543
131	2	4584	3589	gi 1046081	hypothetical protein [GB:D6189_10] [Mycoplasma genitalium]	67	30	996
140	3	2899	2297	gi 146549	kdpC [Escherichia coli]	67	45	603
142	4	5409	4198	gi 1212775	GTP cyclohydrolase 11 [Bacillus amyloliquefaciens]	67	55	1212
147	5	2913	2374	gi 1303709	YrkJ [Bacillus subtilis]	67	44	540
152	8	6341	6673	gi 1377841	Unknown [Bacillus subtilis]	67	48	333
161	4	2720	3763	gi 496319	SphX [Synechococcus sp.]	67	67	1044
163	6	1989	3428	gi 595681	2-oxoglutarate/malate translocator [Spinacia oleracea]	67	47	1440
193	3	1351	1626	gi 1511101	shikimate 5-dehydrogenase [Methanococcus jannaschii]	67	53	276
200	2	917	2179	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	67	48	1263
206	10	12445	12801	sp P37347 VECD	HYPOTHETICAL 21.8 KD PROTEIN IN ASPS 5' REGION	67	47	357
206	11	13047	14432	gi 732813	branched-chain amino acid carrier [Lactobacillus delbrueckii]	67	46	1386
208	2	1321	809	gi 1033037	100 kDa heat shock protein (Hsp100) [Leishmania major]	67	36	513
238	3	1039	2052	gi 809542	CbrB protein [Erwinia chrysanthemi]	67	42	1014

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
246	2	176	367	gi 1215098	excisionase [Bacteriophage 154a]	67	37	192
276	2	2260	1412	gi 1303560	ORF271 [Escherichia coli]	67	50	849
297	6	2223	3056	gi 142784	CtaA protein [Bacillus firmus]	67	46	834
307	7	5220	4186	gi 1070013	protein-dependent [Bacillus subtilis]	67	43	1035
316	1	36	1028	gi 1161061	dioxigenase [Methylobacterium extorquens]	67	57	993
324	3	5650	5030	gi 1469784	putative cell division protein ftsW [Enterococcus hirae]	67	49	621
336	1	524	284	gi 173122	urea amidolyase [Saccharomyces cerevisiae]	67	45	261
360	1	108	1394	gi 130053	SYN_S [HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.21) (HISTIDINE--THNA LIGASE) (HISRS)]	67	47	1287
364	3	4890	3592	gi 151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir[A44756/A44756 hydroxymethylglucaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.]	67	46	1299
365	3	2940	2113	gi 1296823	orf2 gene product [Lactobacillus helveticus]	67	47	828
367	2	325	918	gi 1039479	ORFU [Lactococcus lactis]	67	47	594
395	3	666	1271	gi 1204516	hypothetical protein (GB:000014_4) [Haemophilus influenzae]	67	55	606
415	1	1800	901	gi 1482579	[CG Site No. 29739 [Escherichia coli]	67	46	900
419	1	1799	903	gi 520752	putative [Bacillus subtilis]	67	48	897
474	1	2	796	gi 1486906	argininosuccinate synthetase [Streptomyces clavuligerus] pir[S57659/S57659 argininosuccinate synthase (EC 6.3.4.5) - treptomyces clavuligerus]	67	49	795
485	2	1921	2226	gi 143434	Rho Factor [Bacillus subtilis]	67	43	306
596	1	1728	865	gi 1303853	[YggF [Bacillus subtilis]	67	47	864
700	1	433	218	gi 1204628	hypothetical protein (SP:P21498) [Haemophilus influenzae]	67	47	216
806	2	249	647	gi 1677947	AppC [Bacillus subtilis]	67	51	399
828	2	340	900	gi 1777761	[IrrA [Synecococcus sp.]	67	37	561
833	1	1407	916	gi 142996	regulatory protein [Bacillus subtilis]	67	41	492
856	1	1555	779	gi 1780224	[K970.2 [Caenorhabditis elegans]	67	38	777
888	1	1614	850	gi 1437315	[TTC start codon [Bacillus licheniformis]	67	40	765
1034	1	1190	597	gi 11205113	hypothetical protein (GB:L19201_15) [Haemophilus influenzae]	67	45	594
1062	1	636	319	gi 1303850	[YggC [Bacillus subtilis]	67	41	318
1067	1	918	460	gi 1329501A329	probable reductase protein - Leishmania major	67	54	459

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1358	1	3	293	gi 1001369	hypothetical protein (Synecococcus sp.)	67	64	291
2181	1	3	302	gi 1510416	hypothetical protein (SP:P11466) (Methanococcus jannaschii)	67	48	300
3000	1	1	507	gi 517205	67 kDa Myosin-crossreactive streptococcal antigen (Streptococcus yogenes)	67	56	507
3066	1	464	234	gi 308861	GTD start codon (Lactococcus lactis)	67	46	231
3087	1	454	251	gi 1205366	oligopeptide transport ATP-binding protein (Haemophilus influenzae)	67	44	204
3101	1	2	256	gi 1531541	uroporphyrinogen III methyltransferase (Zea mays)	67	55	255
3598	1	728	393	gi 151259	HMG-CoA reductase (EC 1.1.1.88) (Pseudomonas mevalonii) p1r[A44756]A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	67	56	336
3765	2	584	366	gi 557489	menD (Bacillus subtilis)	67	45	219
3788	1	658	398	pir S52915 S529	nitrate reductase alpha chain - Bacillus subtilis (fragment)	67	45	261
3883	1	2	265	gi 704397	cystathionine beta-lyase (Arabidopsis thaliana)	67	46	264
3926	1	2	340	gi 1481199	peptide-synthetase (Aerobacterales mediterranei)	67	44	339
4417	1	82	396	gi 1205337	ribonucleotide transport ATP-binding protein (Haemophilus influenzae)	67	46	315
2	3	3075	3989	gi 535348	CodV (Bacillus subtilis)	66	42	915
15	6	2273	2542	gi 46491	SatB (Synecococcus PCC7942)	66	37	270
31	9	8059	7826	gi 292046	mucin (Homo sapiens)	66	44	234
31	10	9034	9258	gi 1204545	mercury scavenger protein (Haemophilus influenzae)	66	48	225
32	6	6347	5253	gi 998342	inducible nitric oxide synthase (Gallus gallus)	66	47	1095
44	13	8856	10124	gi 1510751	molybdenum cofactor biosynthesis moaA protein (Methanococcus jannaschii)	66	46	1269
48	2	1276	2868	gi 150209	ORF 1 (Mycoplasma mycoides)	66	40	1593
58	8	7178	8428	gi 665999	hypothetical protein (Bacillus subtilis)	66	47	1251
62	7	5143	4370	gi 1072398	phd gene product (Rhizobium meliloti)	66	40	774
70	14	11693	10998	gi 809660	deoxyribose-phosphate aldolase (Bacillus subtilis) p1r[S49455]S49455 deoxyribose-phosphate aldolase (EC 4.1.2.4) - acillus subtilis	66	55	696
76	1	1	1305	gi 142440	ATP-dependent nuclease (Bacillus subtilis)	66	42	1305
91	6	9236	8205	gi 704397	cystathionine beta-lyase (Arabidopsis thaliana)	66	43	1032
102	5	3810	3785	gi 1204223	hypothetical protein (SP:P11805) (Haemophilus influenzae)	66	41	546

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Accession name	% sim	% ident	length (nt)
103	4	3418	2732	gi 971344	nitrate reductase gamma subunit [Bacillus subtilis] sp P42177 NARI_BACSU NITRATE REDUCTASE GAMMA CHAIN (EC 1.7.99.4) - gi 1009369 Respiratory nitrate reductase [Bacillus subtilis] (SUB-160)	66	48	687
109	6	4243	4674	gi 170886	glucosamine-6-phosphate deaminase [Candida albicans] pir A46692 A46692 glucosamine-6-phosphate isomerase (EC 5.3.1.10) - east [Candida albicans]	66	45	432
112	17	17491	17712	gi 1333179	ORF YGR111w [Saccharomyces cerevisiae]	66	33	222
116	2	4667	2637	gi 1491813	gamma-glutamyltranspeptidase [Bacillus subtilis]	66	43	2031
150	5	3189	2989	gi 1146224	putative [Bacillus subtilis]	66	30	201
172	5	3264	3662	gi 755152	highly hydrophobic integral membrane protein [Bacillus subtilis] sp P42953 TAGC_BACSU TEICHOIC ACID TRANSLATION PERMEASE PROTEIN AGG.	66	41	399
174	5	4592	3723	gi 1146243	pantothenate synthetase [Bacillus subtilis]	66	49	870
175	4	3209	2880	gi 642655	unknown [Rhizobium melliottii]	66	29	330
175	11	8743	7994	gi 854655	Na/H antiporter system [Bacillus alcalophilus]	66	43	750
190	5	7079	5727	gi 451072	di-tripeptide transporter [Lactococcus lactis]	66	40	1353
195	15	13919	13713	gi 1322411	unknown [Mycobacterium tuberculosis]	66	42	207
217	3	2822	2595	gi 1143542	alternative stop codon [Rattus norvegicus]	66	36	228
231	9	7133	6135	gi 1458327	P08F3.4 gene product [Caenorhabditis elegans]	66	47	999
238	1	41	1043	gi 809541	CbrA protein [Arabidopsis thaliana]	66	42	999
241	1	2102	1053	gi 153067	peptidoglycan hydrolase [Staphylococcus aureus]	66	53	1050
261	1	1178	640	gi 1510859	M. Jannaschii predicted coding region M30790 [Methanococcus jannaschii]	66	40	531
263	3	3731	2973	gi 1205865	tetrahydrodipicolinate N-succinyltransferase [Haemophilus influenzae]	66	47	759
272	8	6548	5484	gi 882101	high affinity nickel transporter [Alcaligenes eutrophus] sp P23516 HOKW_ALCEU HIGH-AFFINITY NICKEL TRANSPORT PROTEIN.	66	44	1065
276	3	2805	2104	gi 1208965	hypothetical 23.3 kD protein [Escherichia coli]	66	47	702
278	2	2830	1784	gi 1488662	phosphatase-associated protein [Bacillus subtilis]	66	48	1047
278	3	3430	2952	gi 303560	ORF271 [Escherichia coli]	66	45	879
279	2	3894	2218	gi 1185289	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase [Bacillus subtilis]	66	48	1677
288	4	2515	2275	gi 1256625	putative [Bacillus subtilis]	66	42	261
292	2	1133	942	gi 1511604	M. Jannaschii predicted coding region M31651 [Methanococcus jannaschii]	66	30	192

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
294	1	1116	559	gi 216314	esterase [Bacillus stearothermophilus]	66	45	558
297	4	2913	1978	gi 1994794	cytochrome a assembly facto [Bacillus subtilis] sp P24009 COXK_BACSU PROBABLE CYTOCHROME C OXIDASE ASSEMBLY FACTOR	66	45	936
316	4	2053	2682	gi 1107839	alginate lyase [Pseudomonas aeruginosa]	66	40	630
338	4	2460	2302	gi 520750	biotin synthetase [Bacillus sphaericus]	66	59	159
339	1	1214	735	gi 467468	7, 8-dihydro-6-hydroxymethylpterin-pyrophosphokinase [Bacillus subtilis]	66	52	480
363	1	3	863	gi 581649	epiC gene product [Staphylococcus epidermidis]	66	47	863
366	2	232	483	gi 1103505	unknown [Schizosaccharomyces pombe]	66	53	252
367	4	2468	1845	sp P20692 TYRA_	PREPHENATE DEHYDROGENASE (EC 1.3.1.12) (PDH)	66	50	624
372	3	2150	1599	gi 467416	unknown [Bacillus subtilis]	66	38	552
378	1	212	1009	gi 147309	purine nucleoside phosphorylase [Escherichia coli]	66	50	798
401	1	1	462	gi 388263	p-aminobenzoic acid synthase [Streptomyces griseus] p JN0531 JN0531 p-aminobenzoic acid synthase - Streptomyces riveus	66	46	462
404	7	4826	5254	gi 606744	cytidine deaminase [Bacillus subtilis]	66	51	429
411	2	1738	1103	gi 1460081	unknown [Mycobacterium tuberculosis]	66	44	636
420	1	2	541	gi 1046024	Na+ ATPase subunit 3 [Mycoplasma genitalium]	66	49	540
431	1	1	858	gi 1500008	M. jannaschii predicted coding region M31154 [Methanococcus jannaschii]	66	50	858
443	7	5679	5299	gi 852076	Mga [Bacillus subtilis]	66	46	381
444	3	3405	2413	gi 153047	lysoestaphin (tbp start codon) [Staphylococcus simulans] p P25083 P25083 lysoestaphin precursor - Staphylococcus simulans sp P10547 STP_STASI LYSOESTAPHIN PRECURSOR (EC 3.5.1.-)	66	51	993
561	1	956	480	gi 1204905	DNA-3-methyladenine glycosidase I [Haemophilus influenzae]	66	45	477
562	3	1066	1383	gi 1046082	M. genitalium predicted coding region M0372 [Mycoplasma genitalium]	66	52	318
576	1	11	724	gi 305014	ORF_0214 [Escherichia coli]	66	43	714
577	3	1190	903	gi 1001353	hypothetical protein [Synachocystis sp.]	66	52	288
584	1	2	331	sp P24204 YEB_A_	HYPOTHETICAL 46.7 KD PROTEIN IN HSB8-RV08 INTERGENIC REGION (ORFU)	66	48	330
592	1	1410	706	gi 928839	ORF266; putative [Lactococcus lactis phage BK5-T]	66	51	705
601	1	1433	720	gi 1488695	novel antigen: orf-2 [Staphylococcus aureus]	66	55	714

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
619	3	468	845	gi 746573	similar to H. musculus transport system membrane protein, N-ramp PIR.A40739 and S. cerevisiae SHP1 protein (PIR.A45154) Caenorhabditis elegans	66	45	378
706	2	561	355	gi 804808	unknown protein (Rattus norvegicus)	66	46	207
734	2	673	512	gi 1519085	phosphatidylcholine binding immunoglobulin heavy chain IgM variable region (Mus musculus)	66	60	162
740	1	3	317	gi 1209272	argininosuccinate-lyase (Campylobacter jejuni)	66	47	315
764	1	310	747	gi 435296	alkaline phosphatase like protein (Lactococcus lactis) pIR[S39339]S39339 alkaline phosphatase-like protein - Lactococcus actis	66	42	438
852	1	338	171	gi 516955	CG Site No. 361 (Escherichia coli)	66	43	168
886	1	3	158	gi 289272	ferrichrome-binding protein (Bacillus subtilis)	66	44	156
889	1	462	232	gi 833061	HCMVUL77 (AA 1-642) (Human cytomegalovirus)	66	66	231
893	1	2	247	gi 149008	putative Helicobacter pylori	66	45	246
900	1	1425	733	gi 580842	F3 (Bacillus subtilis)	66	51	693
906	2	2300	1473	gi 790945	aryl-alcohol dehydrogenase (Bacillus subtilis)	66	53	828
947	1	79	549	gi 410117	dlaminopimelate decarboxylase (Bacillus subtilis)	66	47	471
950	1	1100	552	gi 48713	orf145 (Staphylococcus aureus)	66	35	549
955	2	89	475	gi 1204390	uridine kinase (uridine monophosphokinase) (Haemophilus influenzae)	66	50	387
981	2	1308	997	gi 457146	rhostry protein (Plasmodium yoelii)	66	38	312
986	1	25	315	gi 305002	ORF_4356 (Escherichia coli)	66	31	291
1057	1	3	203	gi 1303853	YggF (Bacillus subtilis)	66	40	201
1087	1	1	294	gi 575913	unknown (Saccharomyces cerevisiae)	66	53	294
1105	1	1	231	gi 1045799	methylgalactoside permease ATP-binding protein (Mycoplasma genitalium)	66	46	231
1128	1	2	574	gi 1001493	hypothetical protein (Synecocystis sp.)	66	46	573
1150	1	498	250	gi 1499034	M. jamaeensis predicted coding region M0255 (Methanococcus jamaeensis)	66	40	249
1180	2	707	453	gi 215908	DNA polymerase (g43) (Bacteriophage T4)	66	46	255
1208	1	1123	587	gi 1256653	DNA-binding protein (Bacillus subtilis)	66	58	537
1342	1	1	402	gi 1208474	hypothetical protein (Synecocystis sp.)	66	53	402
1761	2	589	398	gi 215811	tail fiber protein (Bacteriophage T3)	66	50	192

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1983	1	499	251	gi1045935	DNA helicase II (Mycoplasma genitalium)	66	40	249
2103	2	176	400	gi1929798	precursor for the major merozoite surface antigen (Plasmodium alciaparum)	66	46	225
2141	1	373	188	gi1256623	exodeoxyribonuclease (Bacillus subtilis)	66	38	186
2458	1	325	164	gi1019410	unknown (Schizosaccharomyces pombe)	66	47	162
2505	1	468	235	gi1510394	putative transcriptional regulator (Methanococcus jannaschii)	66	39	234
2525	1	558	280	gi1000695	cytotoxin L (Clostridium sordellii)	66	44	279
2935	1	3	275	gi1765073	autolysin (Staphylococcus aureus)	66	47	273
3005	1	114	305	gi1205784	heterocyst maturation protein (Haemophilus influenzae)	66	46	192
3088	1	80	277	gi1303813	Yqew (Bacillus subtilis)	66	42	198
3071	1	1	189	gi1070014	protein-dependent (Bacillus subtilis)	66	41	189
3081	1	404	225	gi1984212	unknown (Schizosaccharomyces pombe)	66	44	180
3090	2	580	386	gi1204987	DNA polymerase III, alpha chain (Haemophilus influenzae)	66	48	195
3118	1	1	387	gi1009366	respiratory nitrate reductase (Bacillus subtilis)	66	49	387
3739	1	798	400	gi1109684	ProV (Bacillus subtilis)	66	47	399
3796	1	402	202	gi1853760	acyl-CoA dehydrogenase (Bacillus subtilis)	66	60	201
1924	1	525	347	gi1563952	gluconate permease (Bacillus licheniformis)	66	46	249
4240	1	3	350	gi1151259	HMG-CoA reductase (EC 1.1.1.88) (Pseudomonas mevalonii pIR1A44756) hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) (Pseudomonas sp.)	66	51	348
4604	1	7	234	pir1A26713 BHHC	hemocyanin subunit II - Atlantic horseshoe crab	66	46	228
4	9	8845	9750	gi1145646	cynR (Escherichia coli)	65	35	906
6	5	2708	3565	gi1887824	ORF_0310 (Escherichia coli)	65	47	858
13	1	1993	998	gi143402	recombination protein (ttg start codon) (Bacillus subtilis) gi1303923 RecN (Bacillus subtilis)	65	44	996
15	7	2493	3524	gi11403126	cacD gene product (Alcaligenes eutrophus)	65	38	1032
18	3	1908	1372	gi1349187	acyltransferase (Saccharomyces cerevisiae)	65	50	537
21	3	1467	2492	gi1149518	phosphoribosyl anthranilate transferase (Bacillus subtilis) pir1S35126 S35126 anthranilate phosphoribosyltransferase (EC 4.2.18) - Lactococcus lactis subsp. lactis	65	52	1026
25	4	3374	4312	gi11502420	malonyl-CoA:Acyl carrier protein transacylase (Bacillus subtilis)	65	44	939

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
27	2	390	626	gi 1212729	[yqhJ (Bacillus subtilis)]	65	45	237
31	12	11040	10387	gi 509245	[D-hydroxylase dehydrogenase (Lactobacillus delbrueckii)]	65	41	654
38	24	19172	19528	gi 547...	[W-protein (Flavobacterium crovis)]	65	41	357
44	2	790	1746	gi 405882	[yolK (Escherichia coli)]	65	46	957
44	12	9356	8832	gi 1205905	[molybdenum cofactor biosynthesis protein (Haemophilus influenzae)]	65	50	525
45	8	6635	7588	gi 493074	[ApoA protein (Salmonella typhimurium)]	65	46	954
51	2	580	1503	gi 580897	[OppB gene product (Bacillus subtilis)]	65	45	924
52	1	225	953	gi 1205518	[NAD(P)H:flavin oxidoreductase (Haemophilus influenzae)]	65	45	729
55	4	1339	1058	pir A44459 A444	[tropomyosin T beta TmT-5 - rabbit]	65	41	282
67	9	7421	8272	gi 143607	[sporulation protein (Bacillus subtilis)]	65	42	852
73	5	4446	5375	gi 1204896	[lysophospholipase L2 (Haemophilus influenzae)]	65	37	910
74	1	954	478	gi 1204844	[H. influenzae predicted coding region H10594 (Haemophilus influenzae)]	65	50	477
77	1	2	757	gi 1046082	[M. genitalium predicted coding region M0372 (Mycoplasma genitalium)]	65	46	756
77	2	795	1433	gi 1222116	[permease (Haemophilus influenzae)]	65	37	639
81	3	4728	3454	gi 1001708	[hypothetical protein (Synecococcus sp.)]	65	49	1275
91	7	8548	8357	gi 1399263	[cystathionine beta-lyase (Enterococcus nidulans)]	65	40	192
98	3	1608	1906	gi 467423	[unknown (Bacillus subtilis)]	65	38	381
98	4	2250	2987	gi 467424	[unknown (Bacillus subtilis)]	65	45	738
102	3	2598	2119	gi 1511532	[N-terminal acetyltransferase complex, subunit ARD1 (Methanococcus jannaschii)]	65	39	480
102	4	3647	2862	gi 1204637	[H. influenzae predicted coding region H10388 (Haemophilus influenzae)]	65	32	786
103	9	10851	9841	gi 142695	[S-adenosyl-L-methionine:uroporphyrinogen III methyltransferase Bacillus megaterium]	65	47	1011
103	10	10439	10119	gi 710021	[nitrite reductase (nirD) (Bacillus subtilis)]	65	51	321
106	2	262	1140	gi 39881	[ORF 311 (AA 1-311) (Bacillus subtilis)]	65	44	879
109	5	3909	4268	gi 1204399	[glucosamine-6-phosphate deaminase protein (Haemophilus influenzae)]	65	44	360
109	10	7165	8595	gi 536955	[CG Site No. 361 (Escherichia coli)]	65	41	1431

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim.	% ident	length (nt)
110	4	3688	3915	gi 407881	stringent response-like protein [Streptococcus equisimilis] p1r[S39975]S39975 stringent response-like protein - Streptococcus equisimilis	65	45	228
110	5	3882	4295	gi 407880	ORF1 [Streptococcus equisimilis]	65	50	414
110	6	4231	4380	gi 1139574	orf2 [Streptomyces griseus]	65	56	150
112	110	9218	8640	gi 1204571	H. influenzae predicted coding region H0318 [Haemophilus influenzae]	65	52	579
112	112	112049	11288	gi 710496	transcriptional activator protein [Bacillus brevis]	65	32	762
125	1	2	202	gi 1151158	repeat organellar protein [Plasmodium chabaudi]	65	39	201
126	1	3	422	gi 137589	intecursor [Homo sapiens]	65	46	420
127	11	10733	12658	gi 1064809	homologous to sp:HTRA_ECOU1 [Bacillus subtilis]	65	41	1926
143	8	7543	7004	gi 216513	mutator mutT (AT-GC transversion) [Escherichia coli]	65	56	540
145	5	3587	3838	gi 1209768	p02_orf569 [Mycoplasma pneumoniae]	65	27	252
150	4	3482	2841	gi 1146225	putative [Bacillus subtilis]	65	37	642
166	1	3858	1948	gi 148304	Beta-1,4-N-acetylmuramoylhydrolase [Enterococcus hirae] p1r[JA2296]JA2296 lysosyme 2 [EC 3.2.1.-] precursor - Enterococcus faec (ATCC 9790)	65	50	1911
188	6	3195	4178	gi 151943	ORF3, putative [Rhodospirillum rubrum]	65	46	984
189	9	4982	4785	gi 58812	ORF IV (AA 1-489) [Fimbrin mosaic virus]	65	40	1798
195	6	7908	5272	gi 145220	alanyl-tRNA synthetase [Escherichia coli]	65	49	2637
195	7	10599	8104	gi 882711	endonuclease V alpha-subunit [Escherichia coli]	65	38	2496
206	16	11686	18191	gi 408115	ornithine acetyltransferase [Bacillus subtilis]	65	53	1296
217	4	3844	3215	gi 1205974	5'-guanylate kinase [Haemophilus influenzae]	65	41	630
220	4	5265	3751	gi 580920	rodin (gral) polypeptide (AA 1-673) [Bacillus subtilis] p1r[S06048]S06048 probable rodin protein - Bacillus subtilis sp[1348]TAGIR, BACSU PROBABLE POLYGLYCEROL-PHOSPHATE) LIPID-GLUCOSYLTRANSFERASE [EC 2.4.1.52] (TECHIOIC ACID BIOSYNTHESIS ROUTE IN E.)	65	40	1515
236	5	2327	3709	gi 1146200	DNA or RNA helicase, RNA-dependent ATPase [Bacillus subtilis]	65	46	1381
237	3	1902	2513	gi 149379	HlsA [Lactococcus lactis]	65	46	612
241	4	4968	4195	gi 1205308	ribonuclease H1 [EC 3.1.264] (RNASE H1) [Haemophilus influenzae]	65	50	774
252	1	1278	940	gi 11204989	hypothetical protein (CB-U00022.9) [Haemophilus influenzae]	65	40	339
261	5	4780	3794	gi 145927	fecD [Escherichia coli]	65	43	987

TABLE 2

S aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
274	1	3	278	gi1496558	orfX (Bacillus subtilis)	65	42	276
301	2	982	815	gi147418	unknown (Bacillus subtilis)	65	45	168
307	4	3586	2864	gi1070014	protein-dependent (Bacillus subtilis)	65	40	723
335	2	2286	1399	gi146913	N-acetylglucosamine transport protein (Escherichia coli) p1r189895[MQEC2N phosphotransferase system enzyme II (EC 7.1.69), N-acetylglucosamine-specific - Escherichia coli sp P09223 PTAA_ECOLI PTS SYSTEM, N-ACETYLGLUCOSAMINE-SPECIFIC IIABC COMPONENT [EIIA	65	50	888
338	5	4120	3170	gi1277029	biotin synthase (Bacillus subtilis)	65	49	951
343	3	1490	2800	gi143264	membrane-associated protein (Bacillus subtilis)	65	48	1311
344	4	2761	2531	gi1050540	tRNA-glutamine Synthetase (Lupinus luteus)	65	34	231
358	3	3421	3621	gi1146220	NAD+ dependent glycerol-3-phosphate dehydrogenase (Bacillus subtilis)	65	47	201
364	1	238	699	gi1340128	ORF1 (Staphylococcus aureus)	65	51	462
379	1	1	576	gi143331	alkaline phosphatase regulatory protein (Bacillus subtilis) p1r1A27650 A27650 regulatory protein phor - Bacillus subtilis sp P21545 PHOR_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN PHOR (EC 2.7.3.1)	65	40	576
379	3	3666	4346	gi143268	ldihydrolysoamide transuccinylase (odh8; EC 2.3.1.61) (Bacillus subtilis)	65	50	681
428	1	187	483	gi1420465	ORF YOR195W (Saccharomyces cerevisiae)	65	45	297
438	2	272	838	gi143498	dugS protein (Bacillus subtilis)	65	38	567
446	11	9280	10215	gi1204756	ribokinase (Haemophilus influenzae)	65	47	936
449	2	1241	1531	gi1599848	Ha/H antiporter homolog (Lactococcus lactis)	65	41	291
478	2	1452	865	gi1045942	glycyl-tRNA synthetase (Hycoplasma genitalium)	65	39	588
479	1	1032	517	gi1498192	putative (Pseudomonas aeruginosa)	65	40	516
480	6	4312	5637	gi1415662	UDP-N-acetylglucosamine 1-carboxyvinyl transferase (Acinetobacter alcoaceticus)	65	48	1326
484	1	2	430	gi146551	transmembrane protein (kdpD) (Escherichia coli)	65	44	429
499	1	54	932	gi1603456	reductase (Lactobacillus majori)	65	53	879
505	1	914	459	gi11518853	OafA (Salmonella typhimurium)	65	39	456
571	2	1509	883	gi149399	open reading frame upstream glnE (Escherichia coli) ir-[S37754][S37754 hypothetical] region XE (glnE 5' region) - Escherichia coli	65	44	627
611	2	506	270	gi10968	RAP-2 (Plasmodium falciparum)	65	40	237

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
705	1	564	283	gi 710020	nitrite reductase (nirB) [Bacillus subtilis]	65	52	282
712	1	1	177	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	65	37	177
712	2	196	354	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	65	37	159
743	1	2	631	gi 310631	ATP binding protein [Streptococcus gordonii]	65	45	610
749	2	393	779	gi 467374	single strand DNA-binding protein [Bacillus subtilis]	65	29	387
762	1	1698	850	gi 160399	multidrug resistance protein [Plasmodium falciparum]	65	48	849
788	1	85	315	gi 1129096	unknown protein [Bacillus sp.]	65	35	211
850	1	1	408	gi 1006604	hypothetical protein [Synecocystis sp.]	65	37	408
908	1	1	444	gi 1199546	2362 [Saccharomyces cerevisiae]	65	46	444
925	1	1	174	gi 1256653	DNA-binding protein [Bacillus subtilis]	65	54	174
1031	1	26	232	gi 238657	AppCycochrome d oxidase, subunit 1 homolog [Escherichia coli, K12, eptide, 514 aa]	65	47	207
1037	1	414	262	gi 1491813	gamma-glutamyltranspeptidase [Bacillus subtilis]	65	46	153
1053	1	348	175	gi 642655	unknown [Rhizobium meliloti]	65	34	174
1149	1	1399	752	gi 1162980	ribulose-5-phosphate 3-epimerase [Spinacia oleracea]	65	48	648
1214	1	801	495	gi 1205959	lactam utilization protein [Haemophilus influenzae]	65	45	387
1276	1	476	276	pis S35493 S354	site-specific DNA-methyltransferase Stal (EC 2.1.1.-) - Streptococcus sanguis	65	35	201
1276	2	900	577	gi 473794	ORF [Escherichia coli]	65	34	324
2057	1	272	138	gi 633699	TrsH [Yersinia enterocolitica]	65	21	135
2521	1	336	169	gi 1045789	hypothetical protein (GB_U14003.76) [Mycoplasma genitalium]	65	41	168
2974	1	590	297	gi 152052	enantiomerase-selective amidase [Rhodococcus sp.]	65	45	294
3031	1	306	154	pis J01024 J010	hypothetical 10K protein (lmpA140 5' region) - fruit fly (Drosophila melanogaster)	65	45	153
3069	1	3	278	gi 144906	product homologous to E.coli thioredoxin reductase; J.Biol.Chem. 1988) 261:9015-9019, and to F52a protein of alkyl hydroperoxide oxidase from S.typhimurium; J.Biol.Chem. (1990) 265:10535-10540; pen reading frame A [Clostridium pasteurianum]	65	46	276
3146	1	282	142	gi 49315	ORF1 gene product [Bacillus subtilis]	65	47	141
3170	1	679	341	gi 1507711	indolepyruvate decarboxylase [Erwinia herbicola]	65	44	339

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3546	1	1	303	gi1450688	hsdM gene of EcoPr1 gene product [Escherichia coli] p1r[S38437/S38437] hsdM protein - Escherichia coli p1r[S09629/S09629] hypothetical protein A - Escherichia coli (SUB 40-520)	65	42	303
3782	1	2	328	gi1166412	NADH-glucanase synthase (Medicago sativa)	65	42	327
3990	1	374	189	gi11009366	Respiratory nitrate reductase [Bacillus subtilis]	65	53	186
4032	1	613	308	gi11323127	ORP YCR087c [Saccharomyces cerevisiae]	65	50	306
4278	2	726	364	gi1197667	vitalogenin [Anolis pulchellus]	65	42	363
19	4	4259	5518	gi1145727	dead [Escherichia coli]	64	45	1260
19	6	7639	6926	gi11016232	ycf27 gene product [Cyanophora paradoxa]	64	36	714
20	8	7053	6454	gi1765073	autolysin [Staphylococcus aureus]	64	47	600
31	13	11206	11537	gi1416009	lpe-85d gene product [Bacillus subtilis]	64	45	1170
33	4	2388	4364	gi11204696	fructose-permease IIBC component [Haemophilus influenzae]	64	47	1977
36	3	1871	3013	gi1290503	glucamate permease [Escherichia coli]	64	40	1143
37	6	4065	4409	gi139815	orf 2 gene product [Bacillus subtilis]	64	46	345
45	9	7852	8760	gi11230585	nucleotide sugar epimerase [Vibrio cholerae O139]	64	53	909
53	3	1540	1899	gi11303961	YqjJ [Bacillus subtilis]	64	50	360
56	6	4793	3855	gi1457514	glcC [Bacillus subtilis]	64	45	939
56	24	30002	30247	gi1470331	similar to zinc fingers [Caenorhabditis elegans]	64	42	246
62	4	2759	2421	gi1642655	unknown [Rhizobium meliloti]	64	28	339
85	6	7178	6027	gi1457702	5-aminimidazole ribonucleotide-carboxylase [Pichia methanolic] p1r[S39112/S39112] phosphoribosylaminoimidazole carboxylase (EC 1.1.2.1) - yeast [Pichia methanolica]	64	46	1152
96	9	9251	10030	gi1511513	ABC transporter, probable ATP-binding subunit [Methanococcus jannaschii]	64	42	780
100	1	1	600	gi1765073	autolysin [Staphylococcus aureus]	64	44	600
106	5	3868	4854	gi1466778	lysine specific permease [Escherichia coli]	64	46	987
123	2	838	554	gi1467484	unknown [Bacillus subtilis]	64	47	285
127	8	7514	7810	gi1210061	serotype-specific antigen [African horse sickness virus] p1r[S27891/S27891] capsid protein VP2 - African horse sickness virus	64	28	297
131	7	7134	6721	gi1511160	M. jannaschii predicted coding region M1163 [Methanococcus jannaschii]	64	46	414

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
142	5	5455	4817	gi 11173517	riboflavin synthase alpha subunit [Actinobacillus pleuropneumoniae]	64	44	639
143	1	709	356	gi A32950 A329	probable reductase protein - Leishmania major	64	52	354
149	10	3555	3295	gi 398151	major surface antigen MSG2 [Pneumocystis carinii]	64	44	261
154	4	3134	2307	gi 984587	DinP [Escherichia coli]	64	50	828
161	5	3855	4880	gi 901304	ORF72 [Bacillus subtilis]	64	37	1036
165	1	33	791	gi 467483	unknown [Bacillus subtilis]	64	38	759
175	6	6355	4214	gi 1072398	phd gene product [Rhizobium meliloti]	64	42	1512
188	3	2042	2500	gi A001961	MHC class II analog [Staphylococcus aureus]	64	45	459
195	14	13667	13446	gi 396380	No definition line found [Escherichia coli]	64	47	222
206	15	16429	16938	gi 304134	argC [Bacillus stearothermophilus]	64	49	510
215	1	560	282	gi 142359	ORF 6 [Xanthobacter vinelandii]	64	39	279
243	7	7818	6928	gi 414014	ipe-90d gene product [Bacillus subtilis]	64	49	891
258	2	1330	845	gi 664754	P17 [Listeria monocytogenes]	64	38	486
259	1	462	232	gi 1499663	M. Jannaschii predicted coding region M0837 [Methanococcus jannaschii]	64	52	231
263	6	6565	5567	gi 1142828	aspartate semialdehyde dehydrogenase [Bacillus subtilis] sp Q04797 MHAS_BACSU ASPARTATE-SEMIALDEHYDE DEHYDROGENASE (EC 2.1.1.1) (ASA DEHYDROGENASE)	64	48	999
271	1	3	1163	gi 467091	hflX; B2235_C2_202 [Mycobacterium leprae]	64	44	1161
280	1	173	1450	gi 1303839	YqfR [Bacillus subtilis]	64	43	1278
293	1	2532	1267	gi 1147345	prinosomal protein n' [Escherichia coli]	64	45	1266
295	2	742	1488	gi 459266	Potential membrane spanning protein [Staphylococcus hominis] p1r S42932 S42932 potential membrane spanning protein - taphylococcus hominis	64	39	747
301	5	1625	1446	gi 580835	lysine decarboxylase [Bacillus subtilis]	64	35	180
315	4	5064	3949	gi 1141396	guinol oxidase [Bacillus subtilis]	64	45	1116
321	1	1264	635	gi 710496	transcriptional activator protein [Bacillus brevis]	64	41	630
333	5	4520	4239	gi 1314295	ORF2; putative 19 kDa protein [Listeria monocytogenes]	64	43	282
342	1	1	549	gi 142940	fcsA [Bacillus subtilis]	64	38	549
353	3	2878	2324	gi 537049	ORF_0470 [Escherichia coli]	64	44	555

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
379	2	827	3658	[gi S25295 A328	oxoglutarate dehydrogenase (liposoma) (EC 1.2.4.2) - Bacillus subtilis	64	47	2832
404	6	4429	4839	[gi A36923 A369	diacylglycerol kinase homolog - Streptococcus mutans	64	35	411
407	1	2020	1133	[gi 969026	OrfX (Bacillus subtilis)	64	41	888
425	1	1109	591	[gi 1146177	phosphotransferase system glucose-specific enzyme II (Bacillus subtilis)	64	44	519
443	6	4082	4798	[gi 147309	purine nucleoside phosphorylase (Escherichia coli)	64	51	717
450	2	1035	1604	[gi 106376	ORF_0162 (Escherichia coli)	64	38	570
470	5	1680	6107	[gi 1169948	host interacting protein (Bacteriophage B1)	64	45	4428
486	4	1911	1471	[gi 1205582	spermidine/putrescine transport system permease protein (Haemophilus influenzae)	64	35	441
497	1	2217	1159	[sp P36929 FNU_E	FNU PROTEIN	64	38	1059
501	1	3	410	[gi 142450	shrC protein (Bacillus subtilis)	64	38	408
514	1	3	290	[gi 1204496	H. influenzae predicted coding region H10238 (Haemophilus influenzae)	64	34	288
551	4	3162	3323	[gi 1204511	bacterioferritin comigratory protein (Haemophilus influenzae)	64	41	162
603	4	759	956	[gi 753823	HADH dehydrogenase F (Streptogyna amaricana)	64	35	198
653	2	940	746	[gi 1213234	dicarboxylic amino acids DiP3p permease (Saccharomyces cerevisiae)	64	41	195
660	3	3801	2257	[sp P46133 YHAI_	HYPOTHETICAL PROTEIN IN OCT 5 REGION (FRAGMENT)	64	39	1345
695	1	11	502	[gi 1001383	hypothetical protein (Synechocystis sp.)	64	41	492
702	1	3	752	[gi 142865	DNA primase (Bacillus subtilis)	64	46	750
826	1	1	339	[gi 971336	arginyl tRNA synthetase (Bacillus subtilis)	64	50	339
838	1	1831	917	[gi 1354775	pfos/R (Treponema pallidum)	64	41	915
864	3	675	944	[gi 139833	cyclomaltodextrin glucanotransferase (Bacillus stearothermophilus) i 39835	64	47	270
887	1	3	677	[gi 153002	enterotoxin type E precursor (Staphylococcus aureus) p A28179 A28179	64	46	675
					enterotoxin E precursor - Staphylococcus aureus sp P12993 ETWE_STAAU			
					ENTEROTOXIN TYPE E PRECURSOR (SEE)			
928	2	1172	963	[gi 1311976	fibrinogen-binding protein (Staphylococcus aureus) p P34270 S34270	64	41	210
					fibrinogen-binding protein - Staphylococcus aureus			
1049	2	800	606	[gi 1049115	Rap60 (Bacillus subtilis)	64	42	195
1067	2	999	748	[gi 1151072	hda precursor (Haemophilus ducreyi)	64	50	252

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1120	1	50	202	gi142439	[ATP-dependent nuclease [Bacillus subtilis]]	64	30	153
1125	1	751	377	gi1581648	[epiB gene product [Staphylococcus epidermidis]]	64	44	375
1688	1	402	214	pir1A01365 TVMS	[transforming protein K-res - mouse]	64	47	189
2472	1	2	358	gi1487282	[Hsp-ATPase subunit J [Enterococcus hirae]]	64	36	357
2989	1	520	356	gi1304134	[arQC [Bacillus stearothermophilus]]	64	50	165
3013	1	630	352	gi1551699	[cytochrome oxidase subunit I [Bacillus firmus]]	64	51	279
3034	1	546	274	gi11204349	[hypothetical protein (GB:D90212_3) [Haemophilus influenzae]]	64	50	273
3197	1	613	308	gi11009366	[Respiratory nitrate reductase [Bacillus subtilis]]	64	46	306
3303	1	90	362	gi1107839	[alginatase lyase [Pseudomonas aeruginosa]]	64	41	273
3852	2	82	288	gi1216746	[D-lactate dehydrogenase [Lactobacillus plantarum]]	64	42	207
3868	1	1	312	gi1149435	[putative [Lactococcus lactis]]	64	48	312
3918	1	660	331	gi15532	[acetyl-CoA acyltransferase [Yarrowia lipolytica]]	64	46	330
4000	1	112	378	gi194688	[unknown [Saccharomyces cerevisiae]]	64	44	267
4009	1	81	368	gi139372	[grab gene product [Bacillus brevis]]	64	41	288
4166	1	2	349	gi1149435	[putative [Lactococcus lactis]]	64	46	348
4366	1	2	307	gi1216267	[ORF2 [Bacillus megaterium]]	64	44	306
4457	1	2	400	gi1197667	[vitellogenin [Anolis pulchellus]]	64	43	399
11	3	1539	2438	gi1438228	[ORF C [Staphylococcus aureus]]	63	32	900
24	7	5611	5423	gi11369943	[al gene product [Bacteriophage B1]]	63	34	189
29	1	1	390	gi1457441	[expressed at the end of exponential growth under conditions in which the enzymes of the TCA cycle are repressed [Bacillus subtilis] gi1467441]	63	43	390
					[expressed at the end of exponential growth under conditions in which the enzymes of the TCA cycle are repressed Bacil]			
31	6	6329	5712	gi1496943	[ORF [Saccharomyces cerevisiae]]	63	47	618
44	123	114669	15019	pir1A04446 QOEC	[hypothetical protein F-92 - Escherichia coli]	63	36	351
48	6	4403	6250	gi143498	[pyruvate synthase [Halobacterium halobium]]	63	42	1848
50	5	3869	4738	gi1413967	[ipa-43d gene product [Bacillus subtilis]]	63	43	870
53	6	6764	5742	gi1474176	[regulator protein [Staphylococcus xylosus]]	63	49	1023

TABLE 2

3. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match gene name	Accession	Match	% sim	% ident	length (nt)
56	14	15880	17607	DNA polymerase III subunit [Bacillus subtilis]	gi 467409		63	44	1728
57	11	7945	7376	ORF_0158 [Escherichia coli]	gi 570316		63	39	570
62	3	2479	2114	unknown [Rhizobium meliloti]	gi 42656		63	41	366
70	8	6562	7353	PhoC [Rhizobium meliloti]	gi 1399821		63	46	792
75	2	223	927	HlsG [Lactococcus lactis]	gi 149376		63	45	705
78	5	4912	4403	ipa-26d gene product [Bacillus subtilis]	gi 411950		63	42	510
91	5	9076	7220	metH2; B2126-CL157 [Mycobacterium leprae]	gi 466997		63	41	1857
91	8	10566	9448	cystathionine gamma-synthase [Haemophilus influenzae]	gi 1204344		63	45	1119
120	1	21	1508	sulfite reductase [NADPH] flavoprotein beta subunit [Escherichia coli]	gi 882657		63	46	1488
120	4	2722	4125	hypothetical protein [Bacillus subtilis]	gi 665994		63	34	1404
127	7	6064	7566	murG gene product [Bacillus subtilis]	gi 40162		63	44	1503
149	6	2321	2106	dnaK [Eryipelothrix rhusiopathiae]	gi 148503		63	40	216
149	26	10445	10170	ORF 2, has similarity to DNA polymerase [Saccharomyces kluyveri] r[S1596]S15961 hypothetical protein 2 - yeast [Saccharomyces uverii] plasmid pSKL	gi 4870		63	42	276
164	2	507	1298	CDP-diglyceride synthetase [Escherichia coli]	gi 145476		63	44	792
166	6	9909	8164	fructose enzyme II [Rhodospirillum rubrum]	gi 151912		61	41	1746
169	4	1704	1886	elongation factor Ts (tsf) [Spiroplasma citri]	gi 152886		63	48	183
188	5	3145	2951	Gly COI 114 grp t8 protein [Podospira anserina]	gi 1334547		63	42	195
195	13	11767	12804	ORF_0335 [Escherichia coli]	gi 606100		63	40	1038
201	2	607	2283	arginyl-tRNA synthetase [Corynebacterium glutamicum] pIR[A49936] arginine-tRNA ligase [EC 6.1.1.19] - oryzae [Bacillus subtilis]	gi 433534		63	46	1677
206	14	15893	16489	N-acetyl-glucamate-gamma-semialdehyde dehydrogenase [Bacillus subtilis]	gi 580828		63	49	597
220	5	7769	5766	secA protein [Bacillus subtilis]	gi 216334		63	42	2004
221	1	74	907	AppA [Bacillus subtilis]	gi 677945		63	42	834
227	3	944	1708	cochylate acid synthase [Methanococcus jannaschii]	gi 1510558		63	46	765
261	2	804	1070	ORF VR054c [Saccharomyces cerevisiae]	gi 486511		63	45	267
269	2	3606	1960	DNA-dependent ATPase, DNA helicase [Escherichia coli] pIR[350137]BVECRQ recQ protein - Escherichia coli	gi 148221		63	42	1647

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
278	8	7417	8176	gi1699273	lysathionine gamma-synthase (Mycobacterium leprae) sp P46807 NETB_MYCLE_CYSTATHIONINE GAMMA-SYNTASE (EC 4.2.99.9) O-SUCCINYLMONOSERINE (THIOL) - LYASE	63	41	1242
287	2	738	1733	gi1405133	putative (Bacillus subtilis)	63	38	996
295	1	2	748	gi11239983	hypothetical protein (Bacillus subtilis)	63	41	747
328	3	2148	3134	gi145302	carrier protein (AA 1 - 417) (Pseudomonas aeruginosa) ir S11497 S11497 branched-chain amino acid tRNA:port protein braa - eudomonas aeruginosa	63	36	987
362	2	1226	1216	sp P35136 SERA_	10-3-PHOSPHOGLYCERATE DEHYDROGENASE (EC 1.1.1.95) (PCDH)	63	38	411
404	1	326	1051	gi11303816	Yqaz (Bacillus subtilis)	63	35	726
405	3	2101	1715	gi11303914	Yqhy (Bacillus subtilis)	63	42	387
406	1	451	227	gi1142152	sulfate permease (gtg start codon) (Synecococcus PCC6301) pir A30101 CRYCS7 sulfate transport protein - Synecococcus sp. PCC 7942	63	43	235
415	2	1048	2718	gi11205402	transport ATP-binding protein (Haemophilus influenzae)	63	41	1671
426	4	3575	2679	gi1193268	29-kilodalton protein (Streptococcus pneumoniae) sp P42362 P29K_STRPN 29 KD MEMBRANE PROTEIN IN PSAA 5'REGION ORF1	63	39	897
505	3	1347	2195	gi11418999	orf4 (Lactobacillus sake)	63	40	849
507	1	2	574	gi1546917	comK (Bacillus subtilis, E26, Peptide, 192 aa)	63	35	573
562	2	146	1084	gi143985	nifs-like gene (Lactobacillus delbrueckii)	63	45	939
675	1	427	215	gi11510994	serine aminotransferase (Methanococcus jannaschii)	63	29	213
686	1	3	230	gi1517356	nitrate reductase (NADH) (Lotus japonicus)	63	52	228
701	1	3	392	gi1881940	NorQ protein (Paracoccus denitrificans)	63	41	390
720	1	2	400	gi147168	open reading frame (Streptomyces lividans)	63	35	399
779	1	571	287	gi11261932	unknown (Mycobacterium tuberculosis)	63	41	285
907	1	22	321	gi1149445	ORF1 (Lactococcus lactis)	63	27	300
972	1	794	399	gi11511235	M. jannaschii predicted coding region MJ1232 (Methanococcus jannaschii)	63	27	396
1085	1	1154	618	gi11204277	hypothetical protein (CB:U00019_14) (Haemophilus influenzae)	63	38	537
1094	1	3	542	gi1790943	urea amidolyase (Bacillus subtilis)	63	39	540
1108	1	3	482	pir S49892 S498	regulation protein - Bacillus subtilis	63	44	480
1113	1	1231	617	gi1493017	endocarditis specific antigen (Enterococcus faecalis)	63	45	615

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1300	1	3	695	sp P33940 YOH1	HYPOTHETICAL 54.3 KD PROTEIN IN ECO-ALKB INTERGENIC REGION	63	46	693
1325	1	1	204	gi 928989	ip100 protein [Borrelia burgdorferi]	63	30	204
1814	1	3	245	gi 1303914	YqH [Bacillus subtilis]	63	34	243
2021	1	498	250	pir C33496 C334	hisc homolog - Bacillus subtilis	63	46	249
2325	1	2	193	gi 436132	product is similar to TnpA of transposon Tn554 from Staphylococcus aureus [Clostridium butyricum]	63	40	192
2335	1	1	195	gi 1184298	flagellar MS-ring protein [Borrelia burgdorferi]	63	47	195
2406	1	451	227	gi 1041785	rhoptry protein [Plasmodium yoelii]	63	33	225
2961	2	136	360	gi 312443	carbamoyl-phosphate synthase (glutamine-hydrolysing) [Bacillus subtilis]	63	52	225
2965	1	1	402	gi 1407784	orf-1, novel antigen [Staphylococcus aureus]	63	50	402
2987	1	583	293	gi 1224069	amidase [Moraxella catarrhalis]	63	35	291
2994	1	266	135	gi 836646	phosphoribosylformimino-pralic ketoisomerase [Rhodobacter phaeoides]	63	51	132
3043	1	440	252	gi 1480237	phenylacetaldehyde dehydrogenase [Escherichia coli]	63	40	169
3078	1	609	400	gi 1487982	intrinsic membrane protein [Mycoplasma hominis]	63	36	210
3139	1	2	217	gi 439126	glutamate synthase (NADPH) [Azospirillum brasilense] pir A49916 A49916 glutamate synthase (NADPH) (EC 1.4.1.13) - Azospirillum brasilense	63	47	216
3625	1	793	198	gi 623073	ORF160; putative thioester [aphane 14-II]	63	48	194
3658	1	1	399	gi 1303697	Yrka [Bacillus subtilis]	63	37	399
3659	1	3	395	gi 1256135	YbbF [Bacillus subtilis]	63	48	393
3783	1	720	361	gi 1256902	pyruvate decarboxylase isozyme 2 (Swiss Prot. accession number P16467) [Saccharomyces cerevisiae]	63	34	360
3900	1	338	171	sp P10537 ANYB	BETA-AMYLASE (EC 3.2.1.2) (1.4-ALPHA-D-GLUCAN HALTOHYDROLASE)	63	54	168
4309	1	3	176	pir A37967 A379	neural cell adhesion molecule Ng-CAM precursor - chicken	63	57	174
4367	1	1	195	gi 1121932	Per6p gene product [Pichia pastoris]	63	30	195
4432	1	1	312	gi 1151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir A44756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	63	51	312
4468	1	6	308	gi 296464	ATPase [Lactococcus lactis]	63	36	303
33	3	1411	2400	gi 153675	catagase 6-P kinase [Streptococcus mutans]	62	44	990
36	9	5985	6218	gi 1490521	INH51 [Homo sapiens]	62	51	234

TABLE 2

5. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
17	1	2	721	gi 1107531	ceuE gene product (Campylobacter coli)	62	33	720
38	15	110912	111589	gi 1222058	H. influenzae predicted coding region HIN1279 (Haemophilus influenzae)	62	38	678
38	25	119526	120329	gi 695280	ORF2 (Alcaligenes eutrophus)	62	41	804
57	2	2523	1780	gi 171234	orf1 (Haemophilus influenzae)	62	55	744
57	9	6646	6350	gi 508174	ETIB domain of PTS-dependent Gat transport and phosphorylation Escherichia coli	62	35	297
58	1	2	559	gi 755152	highly hydrophobic integral membrane protein (Bacillus subtilis) sp P43953 TAGC_BACSU TEICHOIC ACID TRANSLLOCATION PERMEASE PROTEIN AGC.	62	34	558
67	10	8250	9014	gi 470683	Shows similarity with ATP-binding proteins from other ABC-transport perons, Swiss Prot Accession Numbers P24137, P08007, P04585, P24136 Escherichia coli	62	31	765
69	8	8315	7494	gi 46816	actVA 4 gene product (Streptomyces coelicolor)	62	44	822
80	3	1793	1320	gi 39993	UDP-N-acetylauramoylalanine-D-glutamate ligase (Bacillus subtilis)	62	43	474
87	7	7034	9205	gi 217191	5'-nucleotidase precursor (Vibrio parahaemolyticus)	62	38	2172
100	3	4051	3089	gi 1511047	phosphoglycerate dehydrogenase (Methanococcus jannaschii)	62	42	963
102	1	2	520	gi 153655	mismatch repair protein (Streptococcus pneumoniae) pir(C28667) C28667 DNA mismatch repair protein hexA - Streptococcus pneumoniae	62	34	519
112	2	466	1068	gi 153741	ATP-binding protein (Streptococcus mutans)	62	37	603
114	7	6855	7562	gi 1204866	L-fucose operon activator (Haemophilus influenzae)	62	38	708
116	4	6823	5633	gi 677947	AppC (Bacillus subtilis)	62	37	1191
124	8	6855	6008	gi 853777	product similar to E.coli PRA2 protein (Bacillus subtilis) pir S55438 S55438 ymAe protein - Bacillus subtilis sp P45873 HEMK_BACSU POSSIBLE PROTOPORPHYRINOGEN OXIDASE (EC 3.3.-)	62	40	852
148	1	24	554	gi 467456	unknown (Bacillus subtilis)	62	30	531
149	20	7591	6725	gi 1205807	replicative DNA helicase (Haemophilus influenzae)	62	41	867
161	3	1503	1153	gi 40067	X gene product (Bacillus sphaericus)	62	42	351
164	15	14673	15632	gi 42219	P35 gene product (AA 1 - 314) (Escherichia coli)	62	38	960
165	2	1166	1447	gi 403936	phenylalanyl-tRNA synthetase alpha subunit (Gly294 variant) unidentified (cloning vector)	62	38	282
166	2	2084	5089	gi 308861	GTC start codon (Lactococcus lactis)	62	44	3006
171	1	1225	614	gi 1048053	hypothetical protein (SP:IP2049) (Hycoplasma genitalium)	62	41	612

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
183	1	2521	1310	gi1143045	hemY [Bacillus subtilis]	62	45	1212
200	1	3	956	gi1142439	ATP-dependent nuclease [Bacillus subtilis]	62	32	954
237	2	935	1966	gi1141695	hsc protein [Escherichia coli]	62	44	1032
261	3	4008	2605	gi1143121	ORF A; putative [Bacillus firmus]	62	42	1404
299	8	4477	4719	gi1146741	expressed at the end of exponential growth under conditions in which the enzymes of the TCA cycle are repressed [Bacillus subtilis] gi1146741	62	47	243
					expressed at the end of exponential growth under conditions in which the enzymes of the TCA cycle are repressed Bacil			
304	6	5018	3819	gi1153015	FemA protein [Staphylococcus aureus]	62	43	1200
324	1	2	262	gi1142717	cytochrome aa3 controlling protein [Bacillus subtilis] pir11439601A33960	62	30	261
					cta protein - Bacillus subtilis ep112946[CTAA_BACSU CYTOCHROME AA3 CONTROLLING PROTEIN]			
325	2	269	1207	gi11581088	methionyl-tRNA formyltransferase [Escherichia coli]	62	39	939
332	6	4894	4631	gi11499960	uridine 5'-monophosphate synthase [Methanococcus jannaschii]	62	36	264
355	1	2	370	gi1145925	fecB [Escherichia coli]	62	32	369
365	8	6628	6804	gi11413943	lipa-19d gene product [Bacillus subtilis]	62	54	177
369	2	2744	1626	pir11435771A435	regulatory protein pfor - Clostridium perfringens	62	42	1119
370	1	34	264	gi1140665	phate-glucosidase [Clostridium thermocellum]	62	37	231
415	3	2709	3176	gi11205401	transport ATP-binding protein [Haemophilus influenzae]	62	35	468
429	1	1578	790	gi11046024	Na+ ATPase subunit J [Mycoplasma genitalium]	62	40	789
444	2	704	1369	gi11581510	modulation gene; integral membrane protein; homology to Rhizobium eguminosarum nodi [Rhizobium loti]	62	37	666
477	2	751	1869	pir114384401A484	ring-infected erythrocyte surface antigen 2, RESA-2 - Plasmodium falciparum	62	44	1119
485	1	241	1707	gi117934	betaine aldehyde dehydrogenase [Beta vulgaris]	62	43	1467
487	3	1141	1311	gi1149445	ORF1 [Lactococcus lactis]	62	31	171
494	2	1134	1313	gi1166815	ribulose biphosphate carboxylase/oxygenase activase [Arabidopsis thaliana]	62	37	180
518	1	193	882	gi1153491	O-methyltransferase [Streptomyces glaucescens]	62	39	690
534	2	369	2522	gi11480429	putative transcriptional regulator [Bacillus stearothermophilus]	62	35	2154
551	6	4371	4820	gi1111113	ferric uptake regulation protein [Campylobacter jejuni]	62	37	450
574	1	1	570	gi1153000	enterotoxin B [Staphylococcus aureus]	62	43	570

TABLE 2

S aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
590	2	344	1171	gi140367	ORF1 (Clostridium acetobutylicum)	62	37	828
655	1	396	830	gi147195	phnA protein (Escherichia coli)	62	44	435
656	1	2	478	gi1205451	cell division inhibitor (Haemophilus influenzae)	62	36	477
676	1	692	348	gi1511613	methyl coenzyme M reductase system, component A2 (Methanococcus jannaschii)	62	36	345
687	1	493	248	gi149272	Asparaginase (Bacillus licheniformis)	62	48	246
700	2	267	944	gi1205822	hypothetical protein (G8:X75627_4) (Haemophilus influenzae)	62	40	678
840	2	1715	1041	gi1045865	M. genitalium predicted coding region MG181 (Mycoplasma genitalium)	62	36	675
864	4	898	1491	gi1144332	deoxyuridine nucleotidohydrolase (Homo sapiens)	62	38	594
916	1	35	400	gi1413931	lipA-7d gene product (Bacillus subtilis)	62	45	366
1071	1	1	771	gi1510649	aspartokinase I (Methanococcus jannaschii)	62	40	771
1084	1	19	609	gi1688011	Agx-1 antigen (human, infertile patient, testis, peptide, 505 aa)	62	39	591
1103	1	3	203	gi1581261	ORF homologous to E. coli motB (Herpetosiphon aurantiacus) p1r[S14030]S14030	62	51	201
1217	1	463	233	gi1460025	ORF2, putative (Streptococcus pneumoniae)	62	41	231
1533	1	644	414	gi1413968	lipA-4d gene product (Bacillus subtilis)	62	48	231
1537	1	3	257	gi1510641	galenyl-tryptophan synthetase (Methanococcus jannaschii)	62	23	255
2287	1	3	161	gi1485956	hmpC gene product (Proteus mirabilis)	62	45	159
2386	1	3	245	gi1285708	nontoxic component (Clostridium botulinum)	62	31	243
2484	1	331	167	gi142092	DNA-repair protein (recA) (Anabaena variabilis)	62	35	165
2490	1	798	400	gi1581648	lep18 gene product (Staphylococcus epidermidis)	62	42	399
3036	1	596	300	gi1710032	uroporphyrinogen III (Bacillus subtilis)	62	51	297
3116	1	1	213	gi1466883	nlfs; B1496_C2_193 (Mycobacterium leprae)	62	44	213
3297	1	823	413	gi1475715	acetyl coenzyme A acetyltransferase (thiolase) (Clostridium acetobutylicum)	62	42	411
3609	1	31	276	gi11408501	homologous to N-acyl-L-amino acid amidohydrolase of Bacillus stearothermophilus (Bacillus subtilis)	62	48	246
3665	2	584	402	gi151259	HM-CoA reductase (EC 1.1.1.88) (Pseudomonas mevalonii) p1r[A44756]A44756	62	40	183
3733	1	3	374	gi1353197	hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	62	42	372

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
3898	1	1	237	gi1153675	tagatose 6-P kinase [Streptococcus mutans]	62	45	237
4027	1	283	143	gi1330705	homologue to gene 30 (aa 1-59); putative [Bovine herpesvirus 4]	62	43	141
4109	1	727	365	gi141748	hcdM protein (AA 1-520) [Escherichia coli]	62	45	363
4303	1	1	303	gi1303813	Yqew [Bacillus subtilis]	62	43	303
4380	1	530	267	gi1235684	mevalonate pyrophosphate decarboxylase [Saccharomyces cerevisiae]	62	55	264
4494	1	2	256	gi1510692	enterotoxin H [Staphylococcus aureus]	62	34	255
4598	1	411	223	gi1763513	ORF4; putative [Streptomyces violaceoruber]	62	45	189
4624	1	1	222	gi141748	hcdM protein (AA 1-520) [Escherichia coli]	62	45	222
5	5	4288	3912	gi1928831	ORF95; putative [Lactococcus lactis phage BK5-T]	61	36	357
11	1	320	162	pir[C33356][C33]	prothymosin alpha homolog (clone 32) - human (fragment)	61	33	159
16	11	10991	11938	gi1205391	hypothetical protein (SP-P3395) [Haemophilus influenzae]	61	44	948
32	1	281	801	gi1066504	exo-beta 1,3 glucanase [Cochliobolus carbonum]	61	50	519
38	3	616	1107	gi1510864	glutamine transport ATP-binding protein Q [Methanococcus jannaschii]	61	41	492
45	4	3082	4038	gi1105686	ProX [Bacillus subtilis]	61	45	957
48	8	7118	7504	gi1498839	ORF2 [Clostridium perfringens]	61	33	387
51	9	4605	5570	gi1388269	traC [Plasmid pAD1]	61	42	966
60	6	1689	2243	gi1205893	hypothetical protein (GB:U00011_3) [Haemophilus influenzae]	61	32	555
62	9	5559	5122	gi1854656	Na/H antiporter system ORF2 [Bacillus alcalophilus]	61	38	438
67	5	4330	5646	gi1466612	nika [Escherichia coli]	61	36	1317
74	2	2400	1504	gi1204846	carbamate kinase [Haemophilus influenzae]	61	40	897
85	1	2198	1101	gi1498756	amidophosphoribosyltransferase PurF [Rhizobium etli]	61	41	1098
86	4	1995	1582	gi1499931	M. jannaschii predicted coding region M31083 [Methanococcus jannaschii]	61	44	414
97	1	74	649	gi1518679	orf [Bacillus subtilis]	61	44	576
99	2	2454	1990	gi1413958	lipa-34d gene product [Bacillus subtilis]	61	19	465
124	7	6223	5123	gi1556881	Similar to Saccharomyces cerevisiae SUA5 protein (Bacillus subtilis) pir[S49358][S49358 ipc-29d protein - Bacillus subtilis sp[P39153]VIMC_BACSU HYPOTHETICAL 37.0 KD PROTEIN IN SPO1R-OLYC INTERGENIC REGION.	61	46	1101
125	4	1668	2531	gi1491643	ORFA gene product [Chloroflexus aurantiacus]	61	43	864

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	fastq gene name	% sim	% ident	length (nt)
132	1	1250	627	pir P00259 P002	hypothetical protein 1 - Enterococcus faecalis plasmid pAM-beta-1 (fragment)	61	43	624
149	9	3617	3075	gi 1144332	deoxyuridine nucleotidohydrolase (Homo sapiens)	61	40	543
149	22	8690	7869	gi 160047	p101/acidic basic repeat antigen (Plasmodium falciparum) p1r A29232 A29232 101K malaria antigen precursor - Plasmodium falciparum (strain Camp)	61	35	822
168	3	1915	2361	gi 1499694	hlt protein, member of the Hlt-family (Methanococcus jamaohii)	61	41	447
171	9	9675	7948	gi 467416	similar to SpoVB (Bacillus subtilis)	61	38	1728
174	3	1042	2340	gi 216374	glutaryl 7-ACA acylase precursor (Bacillus laterosporus)	61	49	1299
190	4	5034	4111	gi 409286	bmrU (Bacillus subtilis)	61	37	924
216	1	2	190	gi 415861	eukaryotic initiation factor 2 beta (eIF-2 beta) (Cryptosporidium parvum)	61	29	189
227	7	4161	5048	gi 216341	ORF for methionine amino peptidase (Bacillus subtilis)	61	41	888
238	4	1959	3047	gi 909343	CbrC protein (Erwinia chrysanthemi)	61	38	1089
247	1	2	694	gi 537231	ORF-4579 (Escherichia coli)	61	38	693
247	2	678	1034	gi 142226	chvD protein (Agrobacterium tumefaciens)	61	40	357
257	2	3523	2627	gi 699379	glvr-1 protein (Mycobacterium leprae)	61	40	897
268	2	3419	3051	gi 40184	ORF1 (Clostridium acetobutylicum)	61	41	369
275	4	4621	4827	gi 1204848	hypothetical protein (gp:M87049_57) (Haemophilus influenzae)	61	36	207
277	1	1	1845	gi 784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1r A56390 A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - streptococcus pneumoniae	61	45	1845
278	9	8003	7032	gi 467462	cysteine synthetase A (Bacillus subtilis)	61	43	972
278	10	9878	8535	gi 1205919	Na+ and Cl- dependent gamma-aminobutyric acid transporter (Haemophilus influenzae)	61	38	1344
283	1	1	366	gi 755607	polyA polymerase (Bacillus subtilis)	61	36	366
288	2	1918	1496	gi 388108	cell wall enzyme (Enterococcus faecalis)	61	43	423
291	1	86	334	gi 454265	FRP3 (Petunia hybrida)	61	38	249
318	1	1104	694	gi 290531	similar to beta-glucoside transport protein (Escherichia coli) sp P31451 PTIB_ECOLI_PTS_SYSTEM_ARBUTIN-LIKE_IIB_COMPONENT PHOSPHOTRANSFERASE ENZYME II, B COMPONENT (EC 2.7.1.69)	61	47	411
330	2	1912	1190	gi 1001805	hypothetical protein (Synchocystis sp.)	61	41	723

TABLE 2

5. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
385	2	1513	1025	gi 533098	DnaD protein [Bacillus subtilis]	61	42	489
426	1	794	399	gi 1303853	YggF [Bacillus subtilis]	61	44	396
438	3	810	1421	gi 1293660	AbsA2 [Streptomyces coelicolor]	61	36	612
454	1	1580	792	gi 733522	phosphatidylinositol-4,5-diphosphate 3-kinase [Dictyostelium discoideum]	61	30	789
464	2	784	560	gi 1123120	CSJ37.5 gene product [Caenorhabditis elegans]	61	38	225
470	8	6077	7357	gi 623073	ORF360; putative [Bacteriophage LU-H]	61	47	1281
509	1	554	279	gi 467484	unknown [Bacillus subtilis]	61	45	276
555	3	1916	1296	gi 141800	anthranilate synthase glutamine amidotransferase [Acinetobacter alcoaceticus]	61	42	621
569	1	1711	857	gi 467090	R2235_C2_195 [Mycobacterium leprae]	61	47	855
585	2	961	803	sp P36686 SURE_	SURVIVAL PROTEIN SURE HOMOLOG (FRAGMENT)	61	33	159
592	3	1694	1422	gi 1221602	immunity repressor protein [Haemophilus influenzae]	61	32	273
603	1	43	357	gi 507738	hmp [Vibrio parahaemolyticus]	61	33	315
669	1	2467	1235	gi 1146243	22.4% identity with Escherichia coli DNA-damage inducible protein ...; putative [Bacillus subtilis]	61	37	1233
675	3	805	1101	gi 403373	glycerophosphoryl diester phosphodiesterase [Bacillus subtilis] pir S37251 S37251 glycerophosphoryl diester phosphodiesterase - acillus subtilis	61	36	297
703	1	1656	829	gi 537181	ORF_470 [Escherichia coli]	61	32	828
728	1	1628	816	gi 806281	DNA polymerase I [Bacillus stearothermophilus]	61	39	813
821	1	61	318	gi 709992	hypothetical protein [Bacillus subtilis]	61	38	258
856	2	2313	1567	gi 609310	portal protein gp3 [Bacteriophage HK97]	61	40	747
923	1	1081	542	gi 143213	putative [Bacillus subtilis]	61	38	540
1124	1	59	370	gi 1107541	C3J09.8 [Caenorhabditis elegans]	61	26	312
1492	1	548	276	gi 406397	unknown [Mycoplasma genitalium]	61	32	273
1602	1	46	318	gi 733522	phosphatidylinositol-4,5-diphosphate 3-kinase [Dictyostelium discoideum]	61	34	273
2500	1	577	290	gi 1045964	hypothetical protein (GB-U14001.297) [Mycoplasma genitalium]	61	31	288
2968	1	2	808	gi 397526	clumping factor [Staphylococcus aureus]	61	55	807
3076	1	3	248	gi 119373	ORF 1 [Lactococcus lactis]	61	41	246

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
3609	2	207	401	gi 11408501	homologous to N-acetyl-L-amino acid amidohydrolase of <i>Bacillus stearothermophilus</i> (Bacillus subtilis)	61	39	195
3662	1	1472	740	gi 1303813	YgeW (Bacillus subtilis)	61	42	738
3672	1	2	442	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	441
3724	1	2	220	gi 1009366	Respiratory nitrate reductase (Bacillus subtilis)	61	41	219
3728	1	3	398	gi 677943	AppB (Bacillus subtilis)	61	46	396
3884	1	3	401	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	47	399
3971	1	3	383	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	45	381
4038	1	661	359	gi 1339970	large subunit of NAM-dependent glutamate synthase (Plectonema boryanum)	61	24	303
4041	1	546	274	gi 413953	lme-29d gene product (Bacillus subtilis)	61	48	273
4047	1	1	402	gi 528991	unknown (Bacillus subtilis)	61	42	402
4102	1	1	345	gi 976025	lraA (Escherichia coli)	61	46	345
4155	1	1	336	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	336
4268	1	463	233	gi 450688	hsdM gene of EcoP1 gene product (Escherichia coli) p1c[S38437]S38437 hsdM protein - Escherichia coli p1c[S09629]S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	61	38	231
4374	1	542	273	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	50	270
4389	1	2	172	gi 147516	ribokinase (Escherichia coli)	61	35	171
4621	1	2	268	gi 1784897	beta-N-acetylhexosaminidase (Streptococcus pneumoniae) p1c[A56390]A56390 mannosyl-glycoprotein ndo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - treptococcus pneumoniae	61	47	267
4663	1	27	227	gi 976025	lraA (Escherichia coli)	61	50	201
4	6	6663	5536	gi 11408501	homologous to N-acetyl-L-amino acid amidohydrolase of <i>Bacillus stearothermophilus</i> (Bacillus subtilis)	60	43	1128

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
11	6	3426	3725	gi 410748	ring-infected erythrocyte surface antigen (Plasmodium falciparum) p1c[A25526]A25526 ring-infected erythrocyte surface antigen recursor - Plasmodium falciparum (strain FC27/Papua New Guinea) sp P13830 RESA_PLAFF RING-INFECTED ERYTHROCYTE SURFACE ANTIGEN RE	60	24	300
11	14	11035	10313	gi 1217651	carboxyl reductase (NADPH) (Rattus norvegicus)	60	28	723
16	12	11917	12930	gi 1001453	hypothetical protein (Synchocystis sp.)	60	37	1014
33	1	26	469	gi 388109	regulatory protein (Enterococcus faecalis)	60	41	449
37	13	10914	9834	gi 1336656	Orf1 (Bacillus subtilis)	60	40	981
39	4	4364	4522	gi 4872	ORF 4 (Saccharomyces kluyveri)	60	47	159
41	1	2047	1025	gi 142822	D-alanine racemase cds (Bacillus subtilis)	60	39	1023
43	4	2474	3607	gi 468046	para-nitrobenzyl esterase (Bacillus subtilis)	60	40	1134
44	10	6756	7769	gi 414234	thf (Escherichia coli)	60	52	1014
45	10	8874	9074	gi 343949	var140.0 (Saccharomyces cerevisiae)	60	44	201
56	18	27842	26430	gi 468764	mecR gene product (Rhizobium meliloti)	60	35	1413
60	2	173	388	gi 1303864	YggO (Bacillus subtilis)	60	33	216
63	2	357	1619	gi 467124	ureD: B229_C3.234 (Mycobacterium leprae)	60	43	1263
69	1	787	395	gi 1518853	OafA (Salmonella typhimurium)	60	36	193
88	1	1	1188	gi 1480429	putative transcriptional regulator (Bacillus stearothermophilus)	60	30	1188
92	6	4735	3881	gi 349227	transmembrane protein (Escherichia coli)	60	37	855
92	7	5996	4923	gi 466613	nikB (Escherichia coli)	60	38	1074
93	1	949	476	gi 1510925	coenzyme F420-reducing hydrogenase, beta subunit (Methanococcus jannaschii)	60	27	474
96	6	7366	7478	gi 472715	accessory protein (Carnobacterium piscicola)	60	30	213
98	6	3212	4069	gi 467425	unknown (Bacillus subtilis)	60	42	858
102	10	7158	7430	gi 143092	acetolactate synthase small subunit (Bacillus subtilis) sp P37252 ILUN_BACSU ACETOLACTATE SYNTHASE SMALL SUBUNIT (EC 1.1.1.18) (ANAS) (ACETOHYDROXY-ACID SYNTHASE SMALL SUBUNIT) (ALS)	60	37	273
109	11	9127	10515	gi 1255259	D-succinylbenzoic acid (OSB) CoA ligase (Staphylococcus aureus)	60	28	1189
109	12	10499	11656	gi 141954	beta-ketothiolase (Mycobacterium tuberculosis)	60	41	1158
119	2	4630	3134	gi 1524280	unknown (Mycobacterium tuberculosis)	60	45	1497

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	naïch gene name	% sim	% ident	length (nt)
121	9	6957	7646	gi1107529	ceuc gene product [Campylobacter coli]	60	35	690
140	7	7704	6013	gi1146547	kdpA [Escherichia coli]	60	45	1692
145	1	2	703	gi11460077	unknown [Mycobacterium tuberculosis]	60	23	702
150	3	2809	2216	gi1146230	putative [Bacillus subtilis]	60	40	594
157	2	1389	961	gi11303975	VqjX [Bacillus subtilis]	60	30	429
158	5	5125	4769	gi11449288	unknown [Mycobacterium tuberculosis]	60	36	357
159	1	511	257	gi1580912	murD gene product [Bacillus subtilis]	60	43	255
160	1	159	1187	gi1204532	hypothetical protein (CB:U19201_29) [Haemophilus influenzae]	60	34	1029
161	14	8249	7866	gi11496003	ORF1: PepY; putative oligoendopeptidase based on homology with Lactococcus lactis PepF (Genbank Accession Number 232522) [Caldicellulosiruptor saccharolyticus]	60	34	384
172	3	1331	2110	gi1485280	28.2 kDa protein [Streptococcus pneumoniae]	60	33	780
173	2	4082	2460	gi1524397	glycine betaine transporter Opd [Bacillus subtilis]	60	41	1623
173	4	5963	4953	gi11100737	NAUP dependent leukotriene b6 12-hydroxylase [Sus scrofa]	60	44	1011
198	1	3	995	gi1413943	Ipa-13d gene product [Bacillus subtilis]	60	42	993
201	4	3641	4573	sp P37028 VADT_	HYPOTHETICAL 29.4 KD PROTEIN IN MEMU-PFS INTERGENIC REGION PRECURSOR.	60	37	933
203	3	3269	2415	gi1927798	IP9719.34p; CAI: 0.14 [Saccharomyces cerevisiae]	60	43	855
206	9	12234	12515	sp P37347 VECD_	HYPOTHETICAL 21.4 KD PROTEIN IN ASP5 5'-REGION.	60	47	282
212	4	1213	1810	gi1332711	hemagglutinin-neuraminidase fusion protein [Human parainfluenza virus 3]	60	34	198
214	1	65	1153	gi1204366	hypothetical protein (CB:U14003_130) [Haemophilus influenzae]	60	36	1089
237	1	2	937	gi1149377	HLSD [Lactococcus lactis]	60	40	936
241	6	5696	4998	gi11046160	hypothetical protein (CB:000021_3) [Mycoplasma genitalium]	60	37	699
260	6	5919	6485	gi1431950	similar to a B.subtilis gene (GB: BACHEM9_5) [Clostridium acetabularum]	60	35	567
264	1	2432	1218	gi1397526	clumping factor [Staphylococcus aureus]	60	53	1215
267	1	3	1409	gi1148316	NaN-antiporter protein [Enterococcus hirae]	60	27	1407
275	3	3804	4595	sp P36889 F368	leuB 3'-region hypothetical protein - Lactococcus lactis subsp. lactis (strain IL1403)	60	35	792
291	3	860	1198	gi11208489	coded for by C. elegans cDNA yk130e12.5; contains C2H2-type zinc fingers [Caenorhabditis elegans]	60	33	339

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
307	6	3421	3176	gi11070014	protein-dependent (Bacillus subtilis)	60	36	246
316	8	4957	5023	gi1413952	lipo-28d gene product (Bacillus subtilis)	60	41	867
328	4	2996	3484	gi11204484	membrane-associated component, branched amino acid transport system (Haemophilus influenzae)	60	39	489
332	5	4887	4363	gi11205449	colicin V production protein (pur regulon) (Haemophilus influenzae)	60	37	525
357	1	1062	532	gi1087842	single-stranded DNA-specific exonuclease (Escherichia coli)	60	41	531
375	2	96	362	gi11057	adenylyl cyclase gene product (Saccharomyces kluyveri) r[301145]OVBK adenylyl cyclase (EC 4.6.1.1) - yeast ccharomyces kluyveri	60	47	267
397	1	66	416	gi1709999	glucuronate dehydratase (Bacillus subtilis)	60	37	351
409	1	2	163	gi1499700	glycogen phosphorylase (Saccharomyces cerevisiae)	60	35	162
453	4	914	1237	gi11196899	unknown protein (Staphylococcus aureus)	60	36	324
453	7	1818	3620	sp P12222 YCF1_HYPOTHETICAL 226 KD PROTEIN (ORF 1901)	hypothetical 226 kd protein (ORF 1901)	60	31	219
470	2	622	945	gi1307821S107	integrin homolog - yeast (Saccharomyces cerevisiae)	60	31	324
500	1	118	606	gi1467407	unknown (Bacillus subtilis)	60	36	489
503	3	752	982	gi1167835	myosin heavy chain (Dictyostellium discoideum)	60	34	231
505	4	2238	3563	gi11510732	NADH oxidase (Methanococcus jannaschii)	60	26	1326
523	1	3	1043	gi1143331	alkaline phosphatase regulatory protein (Bacillus subtilis) pir[427650]427650 regulatory protein phoB - Bacillus subtilis sp P23545 PHO_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN HOB (EC 2.7.3.-)	60	41	1041
543	1	1	465	gi11511103	cobalt transport ATP-binding protein O (Methanococcus jannaschii)	60	40	465
545	1	1	726	gi11498192	putative (Pseudomonas aeruginosa)	60	40	726
556	1	2	1054	gi11477402	kex gene product (Bordetella pertussis)	60	42	1053
578	1	974	489	gi1205129	H. influenzae predicted coding region H10882 (Haemophilus influenzae)	60	42	486
594	1	1	624	gi11212755	adenylyl cyclase (Aeromonas hydrophila)	60	45	624
604	1	3	530	gi1145925	fecB (Escherichia coli)	60	42	528
620	1	926	465	gi1205483	bicyclic resistance protein (Haemophilus influenzae)	60	33	462
630	2	871	1122	gi11486242	unknown (Bacillus subtilis)	60	41	252
645	2	574	425	gi12051136	serine hydromethyltransferase (serine methylase) (Haemophilus influenzae)	60	28	150

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% Ident	length (nt)
684	1	1082	843	gi11205538	hypothetical protein (GB-U14003_302) (Haemophilus influenzae)	60	39	240
786	1	967	485	gi11402944	orfRm1 gene product (Bacillus subtilis)	60	46	483
844	1	588	346	gi1790943	urea amidolyase (Bacillus subtilis)	60	40	243
851	1	1	726	gi1159663	GMP reductase (Ascaris lumbricoides)	60	41	726
871	1	1746	874	gi11001493	hypothetical protein (Synchocystis sp.)	60	39	873
896	1	1558	839	gi1604926	NADH dehydrogenase, subunit 5 (Schizosaccharomyces pombe) ap1P50368 NUSC_SCHCO NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 6.5.3.1)	60	39	720
908	2	448	753	gi1662880	novel hemolytic factor (Bacillus cereus)	60	31	306
979	1	2	595	gi11429255	putative; orf1 (Bacillus subtilis)	60	30	594
1078	1	669	502	gi1581055	inner membrane copper tolerance protein (Escherichia coli) gi1871029 diaulphide isomerase like protein (Escherichia coli) pir1847295 547295 inner membrane copper tolerance protein - Escherichia coli	60	40	168
1112	1	1150	620	gi1407885	ORF3 (Streptomyces griseus)	60	34	531
1135	1	484	275	gi1171407	Vps8p (Saccharomyces cerevisiae)	60	36	210
1146	1	17	562	gi11239981	hypothetical protein (Bacillus subtilis)	60	36	546
1291	1	716	360	pir155730 S575	carboxyl esterase - Acinetobacter calcoaceticus	60	30	357
1332	1	336	169	gi11222056	aminotransferase (Haemophilus influenzae)	60	44	168
1429	1	3	146	gi11205619	ferritin like protein (Haemophilus influenzae)	60	39	144
1722	1	570	286	gi1240052	dihydroflavonol-4-reductase, DFR (Hordeum vulgare-barley, cv. Gula, aptide, 354 aa)	60	36	285
2350	1	385	200	gi1497626	ORF 1 (Plasmid pAQ1)	60	20	186
2936	1	519	310	gi1508981	prephenate dehydratase (Bacillus subtilis)	60	48	210
3027	1	568	302	gi1146199	putative (Bacillus subtilis)	60	37	267
3084	1	20	208	gi11407784	orf-1; novel antigen (Staphylococcus aureus)	60	51	189
3155	1	2	226	gi11046097	cytadherence-accessory protein (Mycoplasma genitalium)	60	34	225
3603	1	368	186	gi1510108	mitochondrial long-chain enoyl-CoA hydratase/3-hydroxycyl-CoA ehydrogenase alpha-subunit (Rattus norvegicus)	60	42	183
3665	1	486	244	gi1151259	NH2-CoA reductase (EC 1.1.1.88) (Pseudomonas moventonii) pir1444756 A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	60	42	243
3747	1	3	146	gi1474192	lucC gene product (Escherichia coli)	60	36	144

TABLE 2

S aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match junction	Match gene name	% sim	% ident	length (nt)
3912	1	3	335	gi 488695	novel antigen: orf-2 (Staphylococcus aureus)	60	44	333
4072	1	3	272	gi 405879	lyeH (Escherichia coli)	60	33	270
4134	1	510	352	gi 780656	chemoreceptor protein (Rhizobium leguminosarum bv. viciae) gi 780656 chemoreceptor protein (Rhizobium leguminosarum bv. iciae)	60	28	159
4207	2	677	402	gi 602031	similar to trimethylamine DH (Mycoplasmaceae capricolus) pir S49950 S49950 probable trimethylamine dehydrogenase (EC 5.99.7) - Mycoplasma capricolus (SOC3) (fragment)	60	41	276
4243	1	127	324	gi 899317	peptide synthetase module (Microcystis aeruginosa) pir S49111 S49111 probable amino acid activating domain - Microcystis aeruginosa (fragment) (SUB 144-528)	60	42	198
4110	1	624	313	gi 308980	pheB (Bacillus subtilis)	60	28	312
4345	1	343	173	gi 510108	mitochondrial long-chain enoyl-CoA hydratase/3-hydroxacyl-CoA dehydrogenase alpha-subunit (Rattus norvegicus)	60	42	171
4382	1	498	280	gi 47382	acyl-CoA-dehydrogenase (Streptomyces purpurascens)	60	48	219
4474	1	53	223	gi 510108	mitochondrial long-chain enoyl-CoA hydratase/3-hydroxacyl-CoA dehydrogenase alpha-subunit (Rattus norvegicus)	60	42	171
23	4	4518	3523	gi 426446	VipA protein (Salmonella typhi)	59	39	996
33	2	707	1483	gi S48604 S486	hypothetical protein - Mycoplasma capricolus (SOC3) (fragment)	59	33	777
33	5	4651	5853	gi 6721	P5982.3 (Caenorhabditis elegans)	59	33	1203
17	2	3228	3299	gi 142833	ORF2 (Bacillus subtilis)	59	37	910
38	21	16784	16593	gi 912576	BIP (Phasodactylum tricornutum)	59	40	192
52	3	2648	2349	gi 534972	ORF_090a (Escherichia coli)	59	44	300
54	12	14181	13402	gi 4403940	transcription regulator (Bacillus subtilis)	59	37	780
57	3	4397	3339	gi 508176	Gat-1-P-DH, NAD dependent (Escherichia coli)	59	40	1059
66	1	986	495	gi 1103901	Yqhr (Bacillus subtilis)	59	34	492
67	7	6552	7460	gi 912461	NixC (Escherichia coli)	59	37	909
70	7	5383	6366	gi 1339822	PhoD precursor (Rhizobium meliloti)	59	46	984
78	1	1	1449	gi 971345	unknown, similar to E. coli cardiolipin synthase (Bacillus subtilis) sp P45860 VME_BACSU HYPOTHETICAL 58.2 PROTEIN IN NARI-ACDA NTERGENIC REGION	59	39	1449
82	10	14329	15536	gi 490328	LORE F (unidentified)	59	44	1206

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
89	2	1602	958	gi 642801	unknown [Saccharomyces cerevisiae]	59	32	645
96	4	4940	5473	gi 1333802	protein of unknown function [Rhodobacter capsulatus]	59	33	534
98	1	2	820	gi 467421	similar to B. subtilis DnaH [Bacillus subtilis]	59	34	819
119	1	166	1557	gi 143122	ORF B; putative [Bacillus firmus]	59	36	1392
120	10	6214	6756	gi 15354	ORF 55.9 [Bacteriophage T4]	59	39	543
120	16	12476	13510	gi 1086575	BetA [Rhizobium meliloti]	59	44	1035
123	1	386	195	gi 984737	catalase [Campylobacter jejuni]	59	38	192
130	1	370	645	gi 1256634	25-8% identity over 120 aa with the Synenococcus sp. MpeV protein; putative [Bacillus subtilis]	59	31	276
131	4	5278	5712	gi 1510655	hypothetical protein (SP:P42297) [Methanococcus jannaschii]	59	30	435
164	1	3	509	gi 1001342	hypothetical protein [Synecocystis sp.]	59	41	507
164	4	1529	2821	gi 1205165	hypothetical protein (SP:P3784) [Haemophilus influenzae]	59	35	1293
164	10	17643	21376	gi 1001381	hypothetical protein [Synecocystis sp.]	59	34	1734
173	3	4727	3717	gi 1184121	leuxin-induced protein [Vigna radiata]	59	50	1011
179	2	2218	1688	gi 1430316	unidentified gene product [Bacillus subtilis]	59	33	531
195	12	11269	11503	gi 762778	WIFS gene product [Anabaena azollae]	59	41	1167
201	5	4702	5670	gi 1510240	hemolysase [Methanococcus jannaschii]	59	32	969
201	7	5719	6315	gi 1511456	M. jannaschii predicted coding region MJ1437 [Methanococcus jannaschii]	59	34	597
209	1	102	461	gi 1204666	hypothetical protein (CB:X73124_53) [Haemophilus influenzae]	59	42	360
214	3	1050	2234	gi 551531	12-nitropropane dioxygenase [Williopsis saturnus]	59	36	1185
214	5	3393	4135	gi 1301709	YrkJ [Bacillus subtilis]	59	32	843
217	2	3381	2167	gi 1204489	ldp (CG Site No. 18430) [Escherichia coli]	59	44	1215
237	5	3078	3785	gi 149182	JulIA [Lactococcus lactis]	59	38	708
251	2	376	960	gi 1203791	YqjJ [Bacillus subtilis]	59	34	585
286	1	1621	812	gi 146551	transmembrane protein (tkpD) [Escherichia coli]	59	31	810
316	5	4978	3860	gi 405879	YehH [Escherichia coli]	59	32	1119
370	3	600	761	gi 1303794	Yqem [Bacillus subtilis]	59	35	162

TABLE 2

S. aureus - putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
392	1	1009	506	gi 547513	orf3 (Haemophilus influenzae)	59	34	504
391	3	1620	1273	gi 152901	ORF 3 (Spirochaeta aurantia)	59	37	348
406	3	2805	1705	gi 709992	hypothetical protein (Bacillus subtilis)	59	34	1101
426	5	3802	3245	gi 1204610	iron(III) dictrate transport ATP-binding protein FECE (Haemophilus influenzae)	59	36	558
429	2	1513	1148	gi 1064809	homologous to sp.HTRA_ECOLI (Bacillus subtilis)	59	42	366
460	2	708	1301	gi 466882	ppsl: B1496_C2_189 (Mycobacterium leprae)	59	37	594
461	4	2212	3135	gi 1498295	homoserine kinase homolog (Streptococcus pneumoniae)	59	37	924
473	1	2929	1607	gi 147989	trigger factor (Escherichia coli)	59	40	1123
480	8	5862	6110	gi 1205311	(3R)-hydroxymyristol acyl carrier protein dehydrase (Haemophilus influenzae)	59	40	249
521	1	14	1354	pir A25620 A256	staphylocoagulase - Staphylococcus aureus (fragment)	59	32	1341
534	4	2994	4073	gi 153746	mammilot-phosphate dehydrogenase (Streptococcus mutans) pir C44798 C44798	59	36	1080
535	1	1	954	gi 1469939	group B oligonucleotide Pept (Streptococcus agalactiae)	59	33	954
551	3	2836	3186	gi 1204511	bacterioferritin consigratory protein (Haemophilus influenzae)	59	45	351
573	2	449	940	gi 386681	ORF VAL022 (Saccharomyces cerevisiae)	59	36	492
650	1	5	748	gi 396400	similar to eukaryotic Na+/H+ exchanger (Escherichia coli) sp P32703 VJCE_ECOLI HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS NTERGENIC REGION (0549)	59	30	744
664	1	566	285	gi 1262748	luciferase-like component (Staphylococcus aureus)	59	33	282
670	1	3	455	gi 1122758	unknown (Bacillus subtilis)	59	42	453
674	3	543	929	gi 293033	integrase (Bacteriophage phi-1C3)	59	46	387
758	1	349	176	gi 1500472	M. jannaschii predicted coding region MJ1577 (Methanococcus jannaschii)	59	37	174
771	2	2270	1461	gi 522150	bromoperoxidase BPO-A1 (Streptomyces aureofaciens) sp P33912 BPA1_STRAU NON-HAEM BROMOPEROXIDASE BPO-A1 (EC 1.11.1.-) BROMIDE PEROXIDASE (BPO1) (SUB 2-275)	59	44	810
825	1	2191	1057	gi 397526	clumping factor (Staphylococcus aureus)	59	47	1095
1052	2	1094	723	gi 289262	comE ORF3 (Bacillus subtilis)	59	36	372
1152	1	373	188	gi 1276669	ORF238 gene product (Porphyra purpurea)	59	37	186

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	Length (nt)
1198	1	492	247	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	59	26	246
1441	1	468	235	gi 1045942	glycyl-tRNA synthetase [Mycoplasma genitalium]	59	37	234
2103	1	1	186	gi 459250	triacylglycerol lipase [Galactomyces geotrichum]	59	33	186
2205	1	793	398	gi 1303794	Yqem [Bacillus subtilis]	59	38	396
2578	1	484	284	gi 258003	insulin-like growth factor binding protein complex acid-labile subunit [rat, liver, peptide, 60] aa	59	48	201
2967	2	145	348	gi 1212730	Yqhk [Bacillus subtilis]	59	44	204
3012	1	3	248	gi 773571	neurofilament protein NF70 [Helix aspersa]	59	31	246
3544	1	3	401	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	399
3548	1	3	401	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	399
3580	1	698	351	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	348
3720	1	722	363	gi 1408494	homologous to penicillin acylase [Bacillus subtilis]	59	36	360
4171	1	3	296	gi 1055218	crotonase [Clostridium acetobutylicum]	59	42	294
4305	1	618	310	gi 1524193	unknown [Mycobacterium tuberculosis]	59	39	309
18	1	1242	622	gi 146913	N-acetylglucosamine transport protein [Escherichia coli] pir B29895 MOEC2W phosphotransferase system enzyme II (EC 7.1.69), N-acetylglucosamine-specific - Escherichia coli gi P09313 PTAA-ECOLI PTS SYSTEM, N-ACETYLGLUCOSAMINE-SPECIFIC IIABC ORPHANT (E11A)	58	43	621
20	7	7020	5845	gi 50502	collagen alpha chain precursor (AA -27 to 1127) [Mus musculus]	58	50	1176
21	5	3234	3626	gi 1054860	phosphoribosyl anthranilate isomerase [Thermotoga maritima]	58	32	391
23	2	2841	1669	gi 1276880	EpsG [Streptococcus thermophilus]	58	29	1173
23	10	9301	8090	pir A31133 A311	diaminopimelate decarboxylase (EC 4.1.1.20) - Pseudomonas aeruginosa	58	37	1212
36	29	22555	22884	gi 973249	vestitone reductase [Medicago sativa]	58	37	330
44	1	2	406	gi 289272	ferrichrome-binding protein [Bacillus subtilis]	58	33	405
45	1	1	552	gi 29464	embryonic myosin heavy chain (1085 AA) (Homo sapiens) ir S12460 S12460 myosin beta heavy chain - human	58	33	552
55	2	759	538	gi 158852	glucose regulated protein [Echinococcus multilocularis]	58	32	222
62	13	8493	8068	gi 975353	kinase-associated protein B [Bacillus subtilis]	58	35	426
63	3	1553	1717	gi 166926	[Arabidopsis thaliana unidentified mRNA sequence, complete cds.], one product [Arabidopsis thaliana]	58	35	165

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
67	13	12017	11229	gi 1228083	NADH dehydrogenase subunit 2 (Chorthippus parallelus)	58	41	789
96	8	8208	9167	gi 709992	hypothetical protein (Bacillus subtilis)	58	42	960
107	2	2065	1364	gi 806327	Escherichia coli hrpA gene for A protein similar to yeast PRP16 and RP22 (Escherichia coli)	58	37	702
112	7	4519	5613	gi 155588	glucose-fructose oxidoreductase (Zymomonas mobilis) pfr[M42289]A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis	58	38	1095
114	6	7118	6503	gi 1377843	unknown (Bacillus subtilis)	58	38	816
143	2	2261	1395	gi A45605 A456	mature-parasite-infected erythrocyte surface antigen MESA - Plasmodium falciparum	58	31	867
151	2	717	950	gi 1370261	unknown (Mycobacterium tuberculosis)	58	31	234
154	6	6015	4627	gi 1209277	pCTH01 gene product (Chlamydia trachomatis)	58	41	1389
154	16	14281	13541	gi 146613	DNA ligase (EC 6.5.1.2) (Escherichia coli)	58	39	741
155	3	2269	1892	gi 1303917	YqjB (Bacillus subtilis)	58	34	378
174	1	1056	539	gi 904198	hypothetical protein (Bacillus subtilis)	58	26	528
189	4	1533	1769	gi 467383	DNA binding protein (probable) (Bacillus subtilis)	58	25	237
201	3	2669	3307	gi 1511453	endonuclease III (Methanococcus jannaschii)	58	34	639
208	1	2	238	gi 1276729	phycobilisome linker polypeptide (Porphyra purpurea)	58	29	237
220	11	14575	13058	gi 397526	clumping factor (Staphylococcus aureus)	58	51	1518
231	3	1629	1474	gi 1002520	Huts (Bacillus subtilis)	58	45	156
233	6	4201	3497	gi 1463023	No definition line found (Caenorhabditis elegans)	58	39	705
243	10	9303	10082	gi 537207	ORF_4277 (Escherichia coli)	58	32	780
257	1	331	1183	gi 1340128	ORF1 (Staphylococcus aureus)	58	44	813
302	2	460	801	gi 140174	ORF X (Bacillus subtilis)	58	34	342
307	11	6984	6127	gi 1303842	YnfU (Bacillus subtilis)	58	30	458
321	3	1914	2747	gi 1239996	hypothetical protein (Bacillus subtilis)	58	41	834
342	4	2724	3497	gi 454838	ORF 6; putative (Pseudomonas aeruginosa)	58	41	774
348	1	1	663	gi 467478	unknown (Bacillus subtilis)	58	36	663
401	2	384	605	gi 143407	para-aminobenzoic acid synthase, component 1 (pab) (Bacillus subtilis)	58	53	222

TABLE 2

B. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
437	1	325	1554	gi 1301866	YngS (Bacillus subtilis)	58	35	1230
445	1	105	1442	gi 581583	protein A (Staphylococcus aureus)	58	32	1338
453	3	789	965	gi 1009455	unknown (Schistosoma haematobium)	58	34	177
453	5	2748	2047	gi 537214	Yjg gene product (Escherichia coli)	58	40	702
479	2	731	1444	gi 1256621	26.7% of identity in 165 aa to a Thermophilic bacterium hypothetical protein 6; putative (Bacillus subtilis)	58	36	714
490	1	909	547	gi 580920	rodd (gtaA) polypeptide (AA 1-673) (Bacillus subtilis) pir[S06048]S06048 probable rodd protein - Bacillus subtilis sp P13484 TAGE_BACSU PROBABLE POLY(GLYCEROL-PHOSPHATE) LPGA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS PROTEIN E)	58	36	363
517	1	1	1164	gi 1947264 Y018	HYPOTHETICAL HELICASE MG018	58	30	1164
517	6	4182	4544	gi 453422	orf268 gene product (Mycoplasma hominis)	58	29	363
546	3	2802	4019	gi 186052	restriction modification system S subunit (Spiroplasma citri) gi 886052 restriction modification system S subunit (Spiroplasma citri)	58	37	1218
562	1	3	179	gi 43831	nifs protein (AA 1-400) (Riesbiella pneumoniae)	58	34	177
600	2	1347	1156	gi 1181839	unknown (Pseudomonas aeruginosa)	58	48	192
604	2	1231	1001	gi 1001353	hypothetical protein (Synechocystis sp.)	58	41	231
619	1	1	504	gi 903748	integral membrane protein (Homo sapiens)	58	43	504
625	1	2	364	gi 1208474	hypothetical protein (Synechocystis sp.)	58	41	363
635	1	1492	755	gi 1510995	transaldolase (Methanococcus jannaschii)	58	41	738
645	1	1	846	gi 677882	ileal sodium-dependent bile acid transporter (Rattus norvegicus) gi 677882 ileal sodium-dependent bile acid transporter (Rattus norvegicus)	58	33	846
645	3	906	1556	gi 1239999	hypothetical protein (Bacillus subtilis)	58	41	651
665	1	771	532	gi 1204262	hypothetical protein (GB:U0128.61) (Haemophilus influenzae)	58	39	240
674	1	635	327	gi 498817	ORF; homologous to small subunit of phage terminases (Bacillus subtilis)	58	39	309
675	2	1312	806	gi 42181	osmC gene product (Escherichia coli)	58	28	507
745	1	618	310	gi 1205432	coenzyme PQQ synthase protein III (pqqIII) (Haemophilus influenzae)	58	32	309
799	2	242	1174	gi 1204669	collagenase (Haemophilus influenzae)	58	36	933
800	2	1096	614	gi 171963	tRNA isopentenyl transferase (Saccharomyces cerevisiae) sp P07884 MO05_YEAST TRNA ISOPENTENYLTRANSFERASE (EC 2.5.1.8) ISOPENTENYL-DIPHOSPHATE: TRNA ISOPENTENYLTRANSFERASE (IPP TRANSFERASE) (IPPT)	58	37	483

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
854	1	1108	605	gi 466778	lysine specific permease [Escherichia coli]	58	44	504
885	1	481	242	gi 861199	protoporphyrin IX Mg-chelatase subunit precursor [Mordeum vulgare]	58	33	240
891	1	3	527	gi 1293660	AbsA2 [Streptomyces coelicolor]	58	31	525
942	1	931	467	gi 405567	traH [Plasmid pSK41]	58	30	465
1002	1	952	521	gi 577649	prelUHA [Staphylococcus aureus]	58	34	432
1438	1	1	261	gi 581558	isoleucyl tRNA synthetase [Staphylococcus aureus] sp P41368 SYIP_STAAU ISOLEUCYL-TRNA SYNTHETASE, RUPIROCIN RESISTANT EC 6.1.1.5 [ISOLEUCINE--TRNA LIGASE] (ILERS) (RUPIROCIN RESISTANCE ROTIN)	58	30	261
1442	1	2	463	gi 971394	similar to Acc.No. D26185 [Escherichia coli]	58	34	462
1873	1	480	241	gi 1133951	small subunit of NADH-dependent glutamate synthase [Plectonema boryanum]	58	38	240
1876	1	3	158	gi 529216	No definition line found [Caenorhabditis elegans] sp P46503 YUX7_CABEL HYPOTHETICAL 7.3 KD PROTEIN F21F12.7 IN HROMOSOME III	58	33	156
1889	1	108	401	gi 1405458	YneR [Bacillus subtilis]	58	29	294
2109	1	3	401	gi 1001801	hypothetical protein [Synechocystis sp.]	58	31	399
2473	1	288	145	gi 510140	ligandopeptidase F [Lactococcus lactis]	58	38	144
2523	1	452	228	gi 644873	catabolic dehydroquinase dehydratase [Acinetobacter calcoaceticus]	58	37	225
3041	1	2	211	gi 1205367	oligopeptide transport ATP-binding protein [Haemophilus influenzae]	58	39	210
3074	1	3	263	gi 1185288	isochoformate synthase [Bacillus subtilis]	58	38	261
3706	1	3	383	gi 456614	mevalonate kinase [Arabidopsis thaliana]	58	48	381
3854	1	1	402	gi 808869	human gcp172 [Homo sapiens]	58	32	402
4082	1	51	224	gi 508551	ribulose-1,5 biphosphate carboxylase large subunit -methyltransferase [Pisum sativum]	58	37	174
4278	1	3	206	gi 180189	cerebellar-degeneration-related antigen (CDR34) [Homo sapiens] gi 182737 cerebellar degeneration-associated protein [Homo sapiens] p1r A29770 A29770 cerebellar degeneration-related protein - human	58	37	204
19	7	7818	7363	gi 1001516	hypothetical protein [Synechocystis sp.]	57	31	456
23	11	9663	8872	gi 606066	ORF_f256 [Escherichia coli]	57	29	792
31	1	4801	2402	gi 153146	ORF3 [Streptomyces coelicolor]	57	32	2400
38	14	11611	10796	gi 144859	ORF B [Clostridium perfringens]	57	31	816
46	14	112063	113046	gi 1001319	hypothetical protein [Synechocystis sp.]	57	25	984

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
51	3	1411	1187	gi 33856 b338	hypothetical 80K protein - <i>Bacillus sphaericus</i>	57	38	225
54	1	1	453	gi 684950	[staphylococcal accessory regulator A (Staphylococcus aureus)]	57	31	453
75	1	3	239	gi 1000470	C27B7.7 [Caenorhabditis elegans]	57	42	237
92	5	3855	3061	gi 143607	sporulation protein [Bacillus subtilis]	57	35	795
96	3	4006	4773	gi 144297	acetyl esterase (XyNC) [Caldocellum saccharolyticum] pir B37202 B37202 acetyl esterase (EC 3.1.1.6) (XyNC) - Caldocellum saccharolyticum	57	34	768
107	3	1480	2076	gi 460955	TagE [Vibrio cholerae]	57	42	597
109	8	5340	5933	gi 1438846	[unknown [Bacillus subtilis]]	57	41	594
112	9	6679	7701	gi 1486250	[unknown [Bacillus subtilis]]	57	33	1023
114	4	6384	4108	gi 871456	putative alpha subunit of formate dehydrogenase [Methanobacterium thermoautotrophicum]	57	37	2277
126	2	430	1053	gi 288301	ORF2 gene product [Bacillus megaterium]	57	37	624
131	5	6517	6277	gi 1511160	M. jannaschii predicted coding region M31163 [Methanococcus jannaschii]	57	38	261
133	3	2668	2201	gi 1303912	YqkM [Bacillus subtilis]	57	40	468
133	4	3383	2784	gi 1221886	[urea7] amidolyase [Haemophilus influenzae]	57	37	600
147	4	2164	1694	gi 467469	[unknown [Bacillus subtilis]]	57	33	471
160	2	1293	1060	gi 558604	chitin synthase 2 [Neurospora crassa]	57	28	234
163	8	5687	4764	gi 145580	[rarp gene product [Escherichia coli]]	57	38	924
168	6	4336	5325	gi 39782	33KDa lipoprotein [Bacillus subtilis]	57	32	990
170	5	3297	3455	gi 603404	[ver16p [Saccharomyces cerevisiae]]	57	37	159
221	6	8026	6809	gi 1136221	[carboxypeptidase [Sulfolobus solfataricus]]	57	32	1218
228	3	1348	1791	gi 288969	[fibronectin binding protein [Streptococcus dysgalactiae] pir S33850 S33850] [fibronectin-binding protein - Streptococcus dysgalactiae]	57	32	444
263	4	4411	3686	gi 1183002	[dihydrodipicolinate reductase [Pseudomonas syringae pv. tabaci]]	57	42	726
276	1	494	255	gi 396380	[no definition line found [Escherichia coli]]	57	40	240
283	2	335	1324	gi 773349	[Bla protein [Bacillus subtilis]]	57	32	990
297	1	469	236	gi 1334820	[reading frame V [cauliflower mosaic virus]]	57	46	234
342	3	1993	2805	gi 1204431	[hypothetical protein [SP-P33644] [Haemophilus influenzae]]	57	35	813

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
375	6	3340	3741	gi1385177	cell division protein (Bacillus subtilis)	57	26	402
433	6	3286	4011	gi1524117	alpha-acetolactate decarboxylase (Lactococcus lactis)	57	40	726
470	3	903	1145	gi1804819	protein serine/threonine kinase (Toxoplasma gondii)	57	30	243
487	5	1391	1723	gi1507323	ORF1 (Bacillus stearothermophilus)	57	28	333
498	1	274	852	gi1114549	NADH-ubiquinone oxidoreductase subunit 4L (Podospore anserina)	57	34	579
503	1	343	173	gi1502283	organic cation transporter OCT2 (Rattus norvegicus)	57	30	171
505	2	1619	1284	gi1466884	gi1496_C2_194 (Mycobacterium lepreae)	57	40	336
519	2	1182	2549	gi11303707	YcrKH (Bacillus subtilis)	57	34	1368
522	2	1234	1945	gi11064809	homologous to ep-HRNA_ECOLI (Bacillus subtilis)	57	36	1290
538	2	909	1415	gi1153179	phosphorothioic n-acetyltransferase (Streptomyces coelicolor) p1rJH0246/JH0246 phosphothioic n-acetyltransferase (EC 2.3.1.-) Streptomyces coelicolor	57	40	507
547	1	968	486	gi1467340	unknown (Bacillus subtilis)	57	50	483
599	1	1062	532	gi12069217YBA	PREPHENATE DEHYDROGENASE (EC 1.1.1.121) (PDH)	57	41	531
620	2	757	572	gi11107894	unknown (Schizosaccharomyces pombe)	57	38	186
622	2	1600	1130	gi1173028	thioredoxin II (Saccharomyces cerevisiae)	57	39	471
625	2	362	1114	gi11262366	hypothetical protein (Mycobacterium lepreae)	57	34	753
680	1	1	204	gi143544	RNA polymerase sigma-30 factor (Bacillus subtilis) p1rJ28625/J28625 transcription initiation factor sigma H - acillus subtilis	57	30	204
690	1	3	629	gi1466520	pocR (Salmonella typhimurium)	57	29	627
696	1	2	433	gi1413972	ipe-48r gene product (Bacillus subtilis)	57	33	432
704	1	36	638	gi1149931	M. jamaeschil predicted coding region MJI083 (Methanococcus jamaeschilii)	57	36	603
732	1	2316	1621	gi1418999	orf4 (Lactobacillus sake)	57	37	696
746	1	451	227	gi1392973	Rab3 (Aplysia californica)	57	42	225
757	1	20	466	gi143979	L. curvatus small cryptic plasmid gene for rep protein (Lactobacillus curvatus)	57	45	447
862	1	2	295	gi13303827	YqfI (Bacillus subtilis)	57	21	294
1049	1	907	455	gi1510108	ORF-1 (Agrobacterium tumefaciens)	57	35	453
1117	1	1387	695	gi1896285	ORF2 terminus uncertain (Leishmania tarentolae)	57	28	693

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1136	1	2	322	gi11303853	YqgF (Bacillus subtilis)	57	38	321
1144	2	1033	611	gi1310083	voltage-activated calcium channel alpha-1 subunit (Rattus oryngicus)	57	46	423
1172	1	1472	738	gi11511146	M. jannaschii predicted coding region M1144 (Methanococcus jannaschii)	57	28	735
1500	2	746	558	gi1142780	putative membrane protein; putative (Bacillus subtilis)	57	35	189
1676	1	659	399	gi1313777	uracil permease (Escherichia coli)	57	31	261
2481	1	2	400	gi11237015	ORF4 (Bacillus subtilis)	57	23	399
3099	1	3	230	gi11204540	isochorismate synthase (Haemophilus influenzae)	57	19	228
3122	1	360	181	gi1882472	ORF_0464 (Escherichia coli)	57	40	180
3560	1	2	361	gi1153490	tetracycline C resistance and export protein (Streptomyces laevis)	57	37	360
3850	1	856	434	gi1155588	glucose-fructose oxidoreductase (Zymomonas mobilis) p1r/A42289/A42289	57	40	423
3931	1	704	354	gi1413953	lpa-29d gene product (Bacillus subtilis)	57	36	351
1993	1	1	384	gi1151259	HMG-CoA reductase (EC 1.1.1.88) (Pseudomonas mevaloni) p1r/A44756/A44756	57	39	384
4065	1	793	398	p1r/JV0037/RDEC	hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	57	31	396
4100	1	596	300	gi11086633	nitrate reductase (EC 1.7.99.4) alpha chain - Escherichia coli	57	37	297
4163	1	571	287	gi121512	T06C10.5 gene product (Caenorhabditis elegans)	57	47	285
4267	2	631	335	gi11000365	patatin (Solanum tuberosum)	57	50	297
4358	1	3	302	gi1298032	SpotIAG (Bacillus subtilis)	57	32	300
4389	2	108	290	gi1405894	EF (Streptococcus suis)	57	37	183
4399	1	2	232	gi11483603	1-phosphofructokinase (Escherichia coli)	57	35	231
4481	1	572	288	gi1405879	Pristinamycin I synthase 1 (Streptomyces pristinaespiralis)	57	44	285
4486	1	512	258	gi1515938	yeH (Escherichia coli)	57	42	255
4510	1	481	242	gi1205301	glutamate synthase (ferredoxin) (Synecocystis sp.1 p1r/S46957/S46957)	57	38	240
4617	1	468	256	gi11511222	glutamate synthase (ferredoxin) (EC 1.4.7.1) - Synecocystis sp.	57	35	213
4	11	12201	11524	gi149204	leukotoxin secretion ATP-binding protein (Haemophilus influenzae)	56	31	678
					restriction modification enzyme, subunit M1 (Methanococcus jannaschii)			
					histidine utilization repressor G (Klebsiella aerogenes) p1r/A36730/A36730			
					hucG protein - Klebsiella pneumoniae (fragment) ep1p19452/HUTG_KLEAB			
					FORMININGLUTAMASE (EC 3.5.3.8) FORMININGLUTAMATE HYDROLASE (HISTIDINE UTILIZATION PROTEIN G) FRAGMENT			

TABLE 2

E. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
22	8	4248	5177	gi 1322222	RACH1 (Homo sapiens)	56	33	930
38	28	21139	22264	gi 1480705	lipote-protein ligase (Mycoplasma capricolum)	56	34	1086
44	3	1861	2421	gi 490320	Y gene product (unidentified)	56	31	561
44	15	10103	10606	gi 1205099	hypothetical protein (G8:LI920_1) (Haemophilus influenzae)	56	39	504
50	6	4820	5161	gi 209931	fiber protein (Human adenovirus type 5)	56	48	342
53	4	2076	2972	gi 623476	transcriptional activator (Providencia stuartii) sp P43463 AARP_PROST TRANSCRIPTIONAL ACTIVATOR AARP.	56	30	897
67	6	5656	6594	gi 466613	nkg (Escherichia coli)	56	32	939
89	3	2364	1810	gi 482922	protein with homology to pail repressor of B. subtilis (Lactobacillus elbrueckii)	56	39	555
96	1	203	913	gi 145594	cAMP receptor protein (crp) (Escherichia coli)	56	35	711
109	121	118250	117846	gi 1204367	hypothetical protein (G8:U1400_278) (Haemophilus influenzae)	56	27	405
112	8	5611	6678	gi 155588	glucose-fructose oxidoreductase (Zymomonas mobilis) pfr A42289 A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) reductor - Zymomonas mobilis	56	40	1068
131	3	6404	5100	gi 619724	MgtE (Bacillus firmus)	56	30	1305
138	2	65	232	gi 413948	lpa-24d gene product (Bacillus subtilis)	56	31	168
138	4	823	1521	gi 580868	lpa-22r gene product (Bacillus subtilis)	56	31	699
146	2	740	447	gi 1046009	M. genitalium predicted coding region MG309 (Mycoplasma genitalium)	56	37	294
147	2	1639	1067	gi 945380	terminase small subunit (Bacteriophage L4-II)	56	35	573
163	1	2	223	gi 143947	glutamine synthetase (Bacteroides fragilis)	56	30	222
166	5	6745	6449	gi 405792	ORF154 (Pseudomonas putida)	56	26	297
187	1	31	393	gi 311237	HLI-1-transferring ATP synthase (Zea mays)	56	30	363
190	1	2	373	gi 1109686	Prox (Bacillus subtilis)	56	35	372
191	8	11538	9943	gi 561070	acyl coenzyme A synthetase (Escherichia coli)	56	35	1596
195	1	1291	647	gi 1510242	collagenase (Methanococcus jannaschii)	56	34	645
230	3	2123	2072	gi 40363	heat shock protein (Clostridium acetobutylicum)	56	39	252
238	5	3383	3775	gi 1477533	sarA (Staphylococcus aureus)	56	31	393
270	2	813	1712	gi 765073	autolysin (Staphylococcus aureus)	56	41	900

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
290	1	3221	1632	gi1547513	orf3 [Haemophilus influenzae]	56	34	1590
297	5	1140	1373	gi1511556	[M. jannaschii predicted coding region M1561 [Methanococcus jannaschii]]	56	40	234
321	2	2947	1799	gi1001801	hypothetical protein [Synecocystis sp.]	56	31	1149
359	2	1279	641	gi146336	inl1 gene product [Rhizobium meliloti]	56	26	639
371	2	360	1823	gi1145304	U-ribulokinase [Bacterichia coli]	56	39	1464
391	4	1762	2409	gi11001634	hypothetical protein [Synecocystis sp.]	56	34	648
402	1	380	192	gi11438904	5-Hr4L receptor [Homo sapiens]	56	48	189
416	4	2480	2109	gi11408486	HS74A gene product [Bacillus subtilis]	56	31	372
424	3	1756	2334	gi1142471	acetolactate decarboxylase [Bacillus subtilis]	56	32	579
457	1	1907	1017	gi11205194	[formamidopyrimidine-DNA glycosylase [Haemophilus influenzae]]	56	36	891
458	2	2423	1812	gi115466	terminase [Bacteriophage SP1]	56	37	612
504	2	2152	1283	gi1142681	lpp38 [Pasteurella haemolytica]	56	38	870
511	1	1	1284	gi1217049	brnG protein [Salmonella typhimurium]	56	37	1284
601	3	1099	1701	gi1467109	rifM; 30S ribosomal protein S18 alanine acetyltransferase; 229_C1_170 [Mycobacterium leprae]	56	43	603
660	5	3547	3774	gi11229106	2K930.1 [Caenorhabditis elegans]	56	30	228
707	1	35	400	gi1151929	[NADPH-sulfite reductase flavoprotein component [Salmonella typhimurium]]	56	38	366
709	2	1385	1095	gi11510801	hydrogenase accessory protein [Methanococcus jannaschii]	56	38	291
718	1	1	495	gi1413948	[lpa-24d gene product [Bacillus subtilis]]	56	35	495
743	1	87	677	gi1928836	repressor protein [Lactococcus lactis phage BK5-7]	56	35	591
790	1	776	399	gi11511313	ABC transporter, probable ATP-binding subunit [Methanococcus jannaschii]	56	33	378
795	1	3	407	gi11205382	cell division protein [Haemophilus influenzae]	56	34	405
813	1	19	930	gi11222161	permease [Haemophilus influenzae]	56	28	912
855	1	3	515	gi11356621	36.7% of identity in 165 aa to a thermophilic bacterium hypothetical protein 6; putative [Bacillus subtilis]	56	33	513
968	1	2	466	gi1547513	orf3 [Haemophilus influenzae]	56	37	465
973	2	1049	732	gi1866022	HexR [Pseudomonas aeruginosa]	56	31	318
1203	1	5	223	gi1184251	[HMG-1 [Homo sapiens]]	56	34	219

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
1976	1	452	237	gi 19806	lysine-rich aspartic acid-rich protein [Plasmodium chebaudi] r[S2218]S2218 lysine/aspartic acid-rich protein - Plasmodium baudi	56	33	216
2161	1	2	400	gi 1237015	ORF4 [Bacillus subtilis]	56	27	399
2958	1	362	183	gi 466685	No definition line found [Escherichia coli]	56	26	180
2979	1	421	212	gi 1204354	spore germination and vegetative growth protein [Haemophilus influenzae]	56	40	210
2994	2	526	326	gi 1816646	phosphoribosylformimino-praic ketolase [Rhodobacter phaeoideae]	56	29	201
3026	1	177	328	gi 1141106	penicillin V amidase [Bacillus sphaericus]	56	30	150
3189	1	289	146	gi 1166604	Similar to aldehyde dehydrogenase [Caenorhabditis elegans]	56	37	144
3770	1	63	401	gi 1129145	acetyl-CoA C-acyltransferase [Mangifera indica]	56	43	339
4054	2	720	361	gi 1205355	Na ⁺ /H ⁺ antiporter [Haemophilus influenzae]	56	31	360
4165	1	1	324	gi 126095	long-chain acyl-CoA dehydrogenase [Mus musculus]	56	36	324
4200	1	505	254	gi 155588	glucose-fructose oxidoreductase [Zymomonas mobilis] pf-[A42289]A42289 glucose-fructose oxidoreductase (EC 1.1.-.-) recursor - Zymomonas mobilis	56	40	252
4273	1	675	355	gi 108861	GTG start codon [Lactococcus lactis]	56	33	321
1	3	4095	3436	gi 15341	Purative orf YCUX8c, len:192 [Saccharomyces cerevisiae] r[S53591]S53591 hypothetical protein - yeast [Saccharomyces evisiae]	55	25	660
11	12	9377	8505	gi 216773	haloacetate dehalogenase H-1 [Haloalkala sp.]	55	32	873
12	4	5133	4534	gi 467337	unknown [Bacillus subtilis]	55	26	600
19	5	5404	5844	gi 1001719	hypothetical protein [Synecocystis sp.]	55	25	441
23	13	14087	12139	gi 474190	lucA gene product [Escherichia coli]	55	30	1749
32	7	5364	6888	gi 1340036	unknown [Mycobacterium tuberculosis]	55	37	1521
34	3	2569	1808	gi 1303968	YqjQ [Bacillus subtilis]	55	39	762
34	5	3960	3412	gi 1303962	YqjK [Bacillus subtilis]	55	33	549
36	1	1291	647	gi 1606045	ORF_0118 [Escherichia coli]	55	27	645
36	6	6220	5243	gi 1001341	hypothetical protein [Synecocystis sp.]	55	31	978
47	3	3054	3821	gi 1001819	hypothetical protein [Synecocystis sp.]	55	21	768
49	1	2065	1127	gi 103373	[glycerophosphoryl] diester phosphodiesterase [Bacillus subtilis] pf-[S3725]S3725 glycerophosphoryl diester phosphodiesterase - actillus subtilis	55	36	919

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
67	11	8966	9565	gi 151053	norA199 protein [Staphylococcus aureus]	55	23	600
75	3	881	1273	gi 121698	L-histidinol: NAD ⁺ oxidoreductase (EC 1.1.1.23) (aa 1-434) [Escherichia coli]	55	33	393
82	9	11587	14194	gi 1136221	carboxypeptidase [Sulfolobus solfataricus]	55	35	1194
87	4	3517	4917	gi 1064812	function unknown [Bacillus subtilis]	55	26	1401
88	2	1172	1636	gi 882463	protein-Nip1-phosphohistidine-sugar phosphotransferase [Escherichia coli]	55	35	465
92	1	127	516	gi 1377832	unknown [Bacillus subtilis]	55	36	390
100	2	836	2035	gi 1170274	seaxanthin epoxidase [Nicotiana glauca]	55	36	1200
100	5	5137	4658	gi 396660	unknown open reading frame [Buchnera aphidicola]	55	29	480
108	3	4266	2986	gi 1499866	M. jamae-like predicted coding region MJ1024 [Methanococcus jamae]	55	31	1281
114	3	2616	1834	gi 1511367	formate dehydrogenase, alpha subunit [Methanococcus jamae]	55	29	783
144	3	1805	1476	gi 1100787	unknown [Saccharomyces cerevisiae]	55	35	330
165	5	6212	5508	gi 1045884	M. genitalium predicted coding region MG199 [Mycoplasma genitalium]	55	27	705
189	5	2205	2576	gi 142569	ATP synthase a subunit [Bacillus firmus]	55	35	372
191	6	9136	6857	gi 559411	B0272.3 [Caenorhabditis elegans]	55	39	2280
194	2	364	636	gi 1145768	K7 kinesin-like protein [Dictyostelium discoideum]	55	34	273
209	4	1335	1676	gi 473357	thi4 gene product [Schizosaccharomyces pombe]	55	35	342
211	2	1693	1145	gi 410130	ORFX6 [Bacillus subtilis]	55	37	549
213	2	644	1372	gi 433692	TraA [Yersinia enterocolitica]	55	28	729
214	7	4144	5481	gi 1001793	hypothetical protein [Synecocystis sp.]	55	30	1338
221	7	11473	9197	gi 466520	pocR [Salmonella typhimurium]	55	32	2277
233	8	5908	4817	gi 1237063	unknown [Mycobacterium tuberculosis]	55	38	1092
236	4	1375	2340	gi 1146199	putative [Bacillus subtilis]	55	32	966
243	2	380	1885	gi 459907	mercuric reductase [Plasmodium falciparum]	55	29	1506
258	1	786	394	gi 455006	orf6 [Rhodococcus fascians]	55	36	393
281	1	126	938	gi 1408493	homologous to SwissProt:YIDA_ECOLI hypothetical protein [Bacillus subtilis]	55	35	813
316	3	1323	2102	gi 1486447	LuxA homologue [Rhizobium sp.]	55	30	780
326	5	2968	2744	gi 1296824	proline isomerase [Lactobacillus helveticus]	55	36	225

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
351	2	2322	1429	gi1204820	hydrogen peroxide-inducible activator (Haemophilus influenzae)	55	28	894
353	4	2197	2412	gi1272475	chitin synthase (Emerella nidulans)	55	50	216
380	1	14	379	gi142554	ATP synthase I subunit (Bacillus megaterium)	55	37	366
383	1	462	232	gi149272	ferrichrome-binding protein (Bacillus subtilis)	55	36	231
386	1	3	938	gi1510251	DNA helicase, putative (Methanococcus jannaschii)	55	30	916
410	2	1208	1891	gi1205144	multidrug resistance protein (Haemophilus influenzae)	55	27	684
483	2	411	833	gi143934	lipo-10r gene product (Bacillus subtilis)	55	26	423
529	3	1777	1433	gi1606150	ORF_F309 (Escherichia coli)	55	33	345
555	1	1088	585	gi143407	para-aminobenzoic acid synthase, component I (pab) (Bacillus subtilis)	55	28	504
565	1	402	202	gi1223961	CDP-cytosine epimerase (Yersinia pseudotuberculosis)	55	41	201
582	1	751	452	gi1256643	20.2k identity with NADH dehydrogenase of the Leishmania major mitochondrion; putative (Bacillus subtilis)	55	36	300
645	5	2260	2057	gi1210824	fusion protein F (bovine respiratory syncytial virus) p12J01481 VGN28A fusion glycoprotein precursor - bovine respiratory syncytial virus (strain A51908)	55	25	204
672	2	957	2216	gi11511333	M. jannaschii predicted coding region MJ1322 (Methanococcus jannaschii)	55	36	1260
730	1	955	479	gi1537007	ORF_F379 (Escherichia coli)	55	30	477
737	1	1859	945	gi1536963	CG Site No. 18166 (Escherichia coli)	55	30	915
742	2	228	572	gi1304160	product unknown (Bacillus subtilis)	55	38	345
817	2	1211	903	gi11136289	histidine kinase A (Dictyostelium discoideum)	55	29	309
819	1	582	355	gi1558073	polymorphic antigen (Plasmodium falciparum)	55	22	228
832	2	1152	724	gi140167	ORF_C (Clostridium acetobutylicum)	55	32	429
840	1	769	386	gi11205875	pseudouridylylase synthase I (Haemophilus influenzae)	55	39	384
1021	1	23	529	gi148563	beta-lactamase (Yersinia enterocolitica)	55	38	507
1026	1	60	335	gi147804	Opp C (AAI-301) (Salmonella typhimurium)	55	26	276
1525	1	1	282	gi1147533	lsrA (Staphylococcus aureus)	55	29	282
1814	2	224	985	gi11046078	H. genitalium predicted coding region MG369 (Mycoplasma genitalium)	55	38	762
3254	1	427	254	gi1413968	lipo-44d gene product (Bacillus subtilis)	55	30	174

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3695	1	686	345	gi 216773	[haloacetate dehalogenase H-1 (Moraxella sp.)	55	32	342
3721	1	1	312	gi 42029	[ORF1 gene product (Escherichia coli)]	55	31	312
3799	1	3	272	gi 42029	[ORF1 gene product (Escherichia coli)]	55	38	270
3889	1	22	423	gi 1129145	[acetyl-CoA C-acyltransferase (Mangifera indica)]	55	45	402
3916	1	2	385	gi 529754	[apeC (Streptococcus pyogenes)]	55	38	384
3945	1	4	198	gi 476252	[phase 1 flagellin (Salmonella enterica)]	55	36	195
4074	1	488	246	gi 42029	[ORF1 gene product (Escherichia coli)]	55	38	243
4184	1	2	343	gi 1524267	[unknown (Mycobacterium tuberculosis)]	55	28	342
4284	1	14	208	gi 1100774	[ferredoxin-dependent glutamate synthase (Synechocystis sp.)	55	36	195
4457	2	644	378	gi 180189	[cerebellar-degeneration-related antigen (CDR34) (Homo sapiens)]	55	38	267
				pir A29770 A29770 cerebellar degeneration-related protein - human				
4514	1	2	244	gi 216773	[haloacetate dehalogenase H-1 (Moraxella sp.)	55	32	243
4599	1	432	217	gi 1129145	[acetyl-CoA C-acyltransferase (Mangifera indica)]	55	42	216
4606	1	416	210	gi 386120	[myosin alpha heavy chain (S2 subfragment) (rabbits, measser, eptide Partial, 234 aa)]	55	27	207
5	6	5348	4932	gi 536069	[ORF YBL047c (Saccharomyces cerevisiae)]	54	27	417
12	7	7166	6165	gi 1205504	[homoserine acetyltransferase (Haemophilus influenzae)]	54	30	1002
23	16	17086	15326	gi 474192	[lucC gene product (Escherichia coli)]	54	31	1761
35	1	2	979	gi 48054	[small subunit of soluble hydrogenase (AA 1-384) (Synechococcus sp.)	54	36	978
				ir S06919 HQ7CS soluble hydrogenase (EC 1.12.-.-) small chain -				
				nechococcus sp. (PCC 6716)				
37	11	9437	8667	gi 537207	[ORF_277 (Escherichia coli)]	54	38	771
37	12	8165	8332	gi 1160967	[palmitoyl-protein thioesterase (Homo sapiens)]	54	37	168
46	15	13025	13804	gi 438473	[protein is hydrophobic, with homology to E. coli ProW; putative Bacillus subtilis]	54	28	780
56	2	203	736	gi 1256139	[YbbJ (Bacillus subtilis)]	54	34	536
57	13	11117	10179	gi 1151248	[inosine-uridine preferring nucleoside hydrolase (Crithidia fasciculata)]	54	32	939
66	2	516	1133	gi 1335781	[Cap (Drosophila melanogaster)]	54	29	618
70	10	8116	8646	gi 1399823	[PhoE (Rhizobium meliloti)]	54	31	531

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
70	15	12556	11801	[sp P02993 TCR_S TETRAVULINE RESISTANCE PROTEIN.		54	29	756
87	5	4915	5706	[gi 1064811 function unknown (Bacillus subtilis)		54	33	792
92	4	3005	2289	[gi 1205366 oligopeptide transport ATP-binding protein (Haemophilus influenzae)		54	33	717
103	2	2596	1556	[gi 710495 protein kinase (Bacillus brevis)		54	33	1041
105	2	3585	2095	[gi 143727 putative (Bacillus subtilis)		54	30	1491
112	4	2137	2732	[gi 153724 HsfC (Streptococcus pneumoniae)		54	41	396
127	2	1720	2493	[gi 144297 acetyl esterase (XynC) (Caldocellum saccharolyticum) pir B37202 B37202 acetyl esterase (EC 3.1.1.6) (XynC) - Caldocellum saccharolyticum		54	34	774
138	5	1600	3306	[gi 42473 pyruvate oxidase (Escherichia coli)		54	36	1707
152	2	525	1172	[gi 1377834 unknown (Bacillus subtilis)		54	23	648
161	9	4831	5469	[gi 903305 ORF73 (Bacillus subtilis)		54	28	639
161	13	6694	7251	[gi 1511039 phosphate transport system regulatory protein (Methanococcus jannaschii)		54	32	558
164	6	3263	4543	[gi 1204976 proyl-tRNA synthetase (Haemophilus influenzae)		54	34	1281
164	20	21602	22243	[gi 143582 spolIIIEA protein (Bacillus subtilis)		54	32	642
171	6	5683	4250	[gi 436965 malA gene products (Bacillus stearothermophilus) pir S03914 S03914 hypothetical protein 1 - Bacillus stearothermophilus		54	37	1434
206	18	19208	19720	[gi 1240016 RO9E10.3 (Caenorhabditis elegans)		54	38	513
218	2	1090	1905	[gi 467378 unknown (Bacillus subtilis)		54	26	816
220	1	1322	663	[gi 1333761 myosin II heavy chain (Naegleria fowleri)		54	22	660
220	13	12655	13059	[pir S00485 S004 gene 11-1 protein precursor - Plasmodium falciparum (fragmenta)		54	35	405
221	3	2030	3709	[gi 1303813 yagW (Bacillus subtilis)		54	34	1680
272	7	5055	4219	[gi 62964 arylamine N-acetyltransferase (AA-1-290) (Gallus gallus) ir S06652 KYCHY3 aryamine N-acetyltransferase (EC 2.3.1.5) (clone NAT-3) - chicken		54	33	837
316	7	4141	4701	[gi 682769 accE gene product (Escherichia coli)		54	31	561
316	10	6994	8742	[gi 413951 lpa-27d gene product (Bacillus subtilis)		54	28	1749
338	3	3377	2214	[gi 490328 LORF F (unidentified)		54	28	1163
341	4	3201	3614	[gi 171959 myosin-like protein (Saccharomyces cerevisiae)		54	25	414

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
346	1	1820	912	gi 396400	similar to eukaryotic Na ⁺ /H ⁺ exchanger [Escherichia coli] sp P32703 YCEC_ECOLI_HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS NTERGENIC REGION [0549]	54	34	909
348	2	623	1351	gi 537109	ORF_334a [Escherichia coli]	54	34	729
378	2	1007	1942	sp P02983 TCR_S	TETRACYCLINE RESISTANCE PROTEIN	54	31	936
408	6	4351	5301	gi 474190	lucA gene product [Escherichia coli]	54	29	951
444	9	7934	8854	gi 216267	ORF2 [Bacillus megaterium]	54	32	921
463	2	2717	2229	gi 104160	product unknown [Bacillus subtilis]	54	50	489
502	2	1696	1133	gi 1205015	hypothetical protein (sp P10120) [Haemophilus influenzae]	54	38	564
505	6	6262	5357	gi 1500558	12-hydroxyhepta-2,4-diene-1,7-dioate isomerase [Methanococcus jannaschii]	54	41	906
550	1	2736	1522	gi 40100	rodC (tag) polypeptide (AA 1-746) [Bacillus subtilis] tr S06049 S06049 rodC protein - Bacillus subtilis p P13485 TAGP_BACSU TECHOIC ACID BIOSYNTHESIS PROTEIN F	54	35	1215
551	5	3305	4279	gi 950197	unknown [Corynebacterium glutamicum]	54	34	975
55n	2	1356	958	gi 485090	No definition line found [Caenorhabditis elegans]	54	32	399
580	1	91	936	gi 331906	fused envelope glycoprotein precursor [Friend spleen focus-forming virus]	54	45	846
603	3	554	757	gi 1323423	ORF YCR234w [Saccharomyces cerevisiae]	54	36	204
617	1	25	249	gi 219959	ornithine transcarbamylase [Homo sapiens]	54	40	225
622	3	1097	1480	gi 1103873	YogZ [Bacillus subtilis]	54	25	384
623	1	3	404	gi 1063250	low homology to p20 protein of Bacillus licheniformis and bleomycin acetyltransferase of Streptomyces verticillus [Bacillus subtilis]	54	45	402
689	1	1547	1011	gi 552446	NADH dehydrogenase subunit 4 [Apis mellifera ligustical p r S52968 S52968 NADH dehydrogenase chain 4 - honeybee mitochondrion (SGC4)]	54	30	537
725	2	686	1441	gi 987096	sensory protein kinase [Streptomyces hygroscopicus]	54	26	756
956	1	1	249	p r S10782 S107	integrin homolog - yeast [Saccharomyces cerevisiae]	54	24	249
978	2	1137	859	gi 1101994	ORF YNL091w [Saccharomyces cerevisiae]	54	33	279
1314	1	3	281	gi 1001108	hypothetical protein [Synecocystis sp.]	54	33	279
2450	1	1	228	gi 1045057	Ich-TOG [Homo sapiens]	54	32	228
2934	1	1	387	gi 580870	ipe-37d qxa gene product [Bacillus subtilis]	54	36	387
2970	1	499	251	sp P3734P YECE	HYPOTHETICAL PROTEIN IN ASP5 5'REGION (FRAGMENT)	54	42	249

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	magp gene name	% sim	% ident	length (nt)
3002	1	1	309	gi144027	Tma protein [Lactococcus lactis]	54	33	309
3561	1	9	464	gi1151259	HMG-CoA reductase (EC 1.1.1.88) [Pseudomonas mevalonii] pir[A44756] A44756 hydroxymethylglutaryl-CoA reductase (EC 1.1.1.88) Pseudomonas sp.	54	35	456
3572	1	72	401	gi1450688	hcdm gene of EcoRII gene product [Escherichia coli] pir[S18437] S18437 hcdm protein - Escherichia coli pir[S09629] S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	54	36	330
3829	1	798	400	gi11322245	mevalonate pyrophosphate decarboxylase [Rattus norvegicus]	54	29	399
3909	1	1	273	gi129865	CENP-E [Homo sapiens]	54	30	273
3921	1	3	209	pir[S24325] S243	glucan 1,4-beta-glucosidase (EC 3.2.1.74) - Pseudomonas fluorescens subsp. cellulosa	54	34	207
4438	1	566	285	gi1156657	unknown protein [Mycoplasma pneumoniae]	54	30	282
4459	1	3	272	gi11046081	hypothetical protein (GB:D26185.10) [Mycoplasma genitalium]	54	38	270
4564	1	3	221	gi1216267	ORF2 [Bacillus megaterium]	54	38	219
23	12	12538	10685	gi1474192	lucC gene product [Escherichia coli]	53	35	1854
23	14	14841	13579	gi142029	ORF1 gene product [Escherichia coli]	53	32	1263
24	3	4440	3940	gi11369947	ic2 gene product [Bacteriophage B1]	53	36	501
26	4	3818	4618	gi11486247	unknown [Bacillus subtilis]	53	37	801
38	6	2856	3998	gi1405880	yell [Escherichia coli]	53	40	1143
38	10	9380	7806	gi11399954	thyroid sodium/iodide symporter NIS [Rattus norvegicus]	53	29	1575
56	10	112324	12100	pir[A54592] A545	110k actin filam. associated protein - chicken	53	32	225
57	6	5047	4583	pir[A00341] DE2P	alcohol dehydrogenase (EC 1.1.1.3) - fission yeast (Schizosaccharomyces pombe)	53	39	465
57	12	110515	8932	gi11480429	putative transcriptional regulator [Bacillus stearothermophilus]	53	30	1584
67	12	9496	10218	gi11511555	quinolone resistance norA protein protein [Methanococcus jamaacensis]	53	31	723
69	3	3125	2382	gi11087017	arabinogalactan-protein, AGP [Nicotiana glauca, cell-suspension culture filtrate, Peptide, 461 aa]	53	30	744
79	1	3	1031	gi11523802	glucanase [Anabaena variabilis]	53	32	1029
80	1	673	338	gi1152428	ATPase 3 [Plasmodium falciparum]	53	36	336
88	4	1910	2524	gi117034	ORF_0488 [Escherichia coli]	53	25	615
88	5	2467	3282	gi11537034	ORF_0488 [Escherichia coli]	53	29	816

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
92	8	5870	5505	gi1399598	amphotropic murine retrovirus receptor [Rattus norvegicus]	53	33	366
94	5	4417	3239	gi1173038	tropomyosin (TPM1) [Saccharomyces cerevisiae]	53	25	1179
99	5	4207	5433	sp P28246 RCR_E	8-CYCLOXYCIN RESISTANCE PROTEIN (SULFONAMIDE RESISTANCE PROTEIN)	53	30	1227
120	3	1639	2262	gi1576655	ORF1 [Vibrio anguillarum]	53	35	624
120	11	7257	8897	gi1524397	glycine betaine transporter OpuD [Bacillus subtilis]	53	33	1641
127	6	6893	5685	gi1256630	putative [Bacillus subtilis]	53	32	1209
147	2	255	557	gi1581648	epib gene product [Staphylococcus epidermidis]	53	34	303
158	4	4705	4256	gi151004	mucoicid regulatory protein Algr [Pseudomonas aeruginosa] pir A32807 A32802 regulatory protein algr - Pseudomonas aeruginosa sp P26275 ALGR_PSEAE POSITIVE ALGINATE BIOSYNTHESIS REGULATORY PROTEIN	53	32	450
171	7	5717	5421	gi1510669	hypothetical protein (GP:D64044.18) [Methanococcus jannaschii]	53	34	297
191	9	13087	11483	gi1298085	acetate decarboxylase [Clostridium acetobutylicum] pir A9346 A9346 butyrate-acetate CoA-transferase (EC 2.8.3.9) small chain - Clostridium acetobutylicum sp P33752 CTFA_CLOAB BUTYRATE-ACETOACETATE COA-TRANSFERASE SUBUNIT (EC 2.8.3.9) (CONT A)	53	31	1605
203	5	3763	4326	gi1343456	ipoE protein (tgg start codon) [Bacillus subtilis]	53	29	564
206	17	18204	18971	gi1304136	acetylglutamate kinase [Bacillus stearothermophilus] sp Q07905 ARGB_BACST ACETYLGLUTAMATE KINASE (EC 2.7.2.8) (NAG INASE) (AGK) (N-ACETYL-L-GLUTAMATE 5-PHOSPHOTRANSFERASE)	53	36	768
212	10	4021	4221	gi1987A	protein kinase [Plasmodium falciparum]	53	28	201
231	2	1580	1350	gi1537506	paramyosin [Dirofilaria immitis]	53	34	231
272	6	2719	3249	pir A33141 A331	hypothetical protein (gcfD 3' region) - Streptococcus mutans	53	34	531
308	3	927	2576	gi1606292	ORF_0696 [Secherichia coli]	53	33	1650
320	7	5645	5884	gi1160596	RNA polymerase III largest subunit [Plasmodium falciparum] sp P27625 RPCL_PLAFA DNA-DIRECTED RNA POLYMERASE III LARGEST SUBUNIT (EC 2.7.7.6)	53	33	240
327	1	218	901	gi1854601	unknown [Schizosaccharomyces pombe]	53	31	684
341	2	212	2500	gi1633732	ORF1 [Campylobacter jejuni]	53	31	2289
351	1	763	383	sp P31675 VABH	HYPOTHETICAL 42.7 KD PROTEIN IN TBPA-LEUD INTERGENIC REGION (ORF104)	53	32	381
433	7	5087	4731	gi11001961	MHC class II analog [Staphylococcus aureus]	53	30	357
454	2	1240	980	pir A60328 A603	40K cell wall protein precursor (ser 5' region) - Streptococcus mutans (strain OM2175, serotype f)	53	27	261

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	AsiQh gene name	% sim	% ident	length (nt)
470	4	1123	1761	gi 516826	rat GC9360 (Rattus rattus)	53	30	639
483	1	432	217	gi 1480429	putative transcriptional regulator (Bacillus stearothermophilus)	53	53	216
544	1	516	1259	gi 46587	ORF 1 (AA 1 - 121) (1 is 2nd base in codon) (Staphylococcus aureus) tr S15745 S15765 hypothetical protein 1 (hib 5' region) - aphylococcus aureus (fragment)	53	38	744
558	10	1957	3754	gi 15140	res gene (Bacteriophage P1)	53	32	204
603	2	339	620	gi 507738	Hmp (Vibrio parahaemolyticus)	53	26	282
693	1	1669	941	gi 153123	toxic shock syndrome toxin-1 precursor (Staphylococcus aureus) pti A24606 KCSA91 toxic shock syndrome toxin-1 precursor - taphylococcus aureus	53	38	729
766	1	2	673	gi 687600	orfA2; orfA2 forms an operon with orfA1 (Listeria monocytogenes)	53	43	672
781	1	667	335	gi 1204551	pillin biogenesis protein (Haemophilus influenzae)	53	26	333
801	1	3	545	gi 1279400	SepA protein (Escherichia coli)	53	25	543
803	1	2	910	gi 695278	lipase-like enzyme (Alcaligenes eutrophus)	53	30	909
872	1	1177	590	gi 298032	EF (Streptococcus suis)	53	30	588
910	1	2	184	gi 1044936	unknown (Schizosaccharomyces pombe)	53	29	183
943	1	794	399	gi 390508	similar to unidentified ORF near 47 minutes (Escherichia coli) sp P31436 YICK_EC001 HYPOTHETICAL 43.5 KD PROTEIN IN SELC-NLPA HYPERGENIC REGION.	53	30	396
988	1	1004	504	gi 142441	ORF 3; putative (Bacillus subtilis)	53	28	501
1064	1	3	434	gi 305080	myosin heavy chain (Entamoeba histolytica)	53	26	432
1366	1	3	452	gi 308852	transmembrane protein (Lactococcus lactis)	53	33	450
1758	1	792	397	gi 1001724	hypothetical protein (Synechocystis sp.)	53	30	396
1897	1	1	447	gi 1303949	YqIX (Bacillus subtilis)	53	27	447
2301	1	798	400	gi 1146243	22.4% identity with Escherichia coli DNA-damage inducible protein ...; putative (Bacillus subtilis)	53	37	399
3537	1	1	327	gi 450688	hscM gene of Ecoprr1 gene product (Escherichia coli) pti S38437 S38437 hscM protein - Escherichia coli pti S09629 S09629 hypothetical protein A - Escherichia coli (SUB 40-520)	53	35	327
3747	2	117	397	gi 1477486	transposase (Burkholderia cepacia)	53	53	261
11	5	3049	3441	gi 868224	[no definition line found (Caenorhabditis elegans)]	52	33	393

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
15	5	2205	2369	gi 215966	G41 protein (qtg start codon) (Bacteriophage T4)	52	34	165
19	3	2429	3808	gi 1205379	UDP-murac-pentapeptide synthetase (Haemophilus influenzae)	52	31	1380
24	1	6920	3462	gi 579124	Predicted 86.4kd protein; 52kd observed (Mycobacteriophage J5) p1r[S30971][S30971] gene 26 protein - Mycobacteriophage J5 sp[005233][VG28_BPM15 MINOR TAIL PROTEIN GP26. (SUB 2-837)	52	32	3459
37	5	3015	3935	gi 1500543	P115 protein (Methanococcus jannaschii)	52	25	921
38	13	8795	9703	gi 46851	glucose kinase (Streptomyces coelicolor)	52	29	909
44	16	10617	11066	gi 42012	moaE gene product (Escherichia coli)	52	36	450
46	1	3	521	gi 1040957	NADH dehydrogenase subunit 6 (Anopheles trinkae)	52	25	519
51	10	5531	6280	gi 388269	trac (Plasmid pAD1)	52	32	750
56	5	3968	2826	gi 161949	endothelial differentiation protein (edg-1) (Homo sapiens) p1r[A35300][A35300] G protein-coupled receptor edg-1 - human sp[921453][EDG1_HUMAN PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.	52	23	1143
57	5	4850	4173	gi 304153	sorbitol dehydrogenase (Bacillus subtilis)	52	27	678
62	5	3364	2870	gi 1072399	phaE gene product (Rhizobium meliloti)	52	25	495
62	6	4445	3651	gi 46485	NADH dehydrogenase (Synecococcus PCC7942)	52	27	795
67	14	11355	12962	gi 1511365	glutamate synthase (NADPH), subunit alpha (Methanococcus jannaschii)	52	30	1608
67	21	16935	18158	gi 1204393	hypothetical protein (SP:P31122) (Haemophilus influenzae)	52	25	1224
70	4	2185	1997	gi 7227	cytoplasmic dynein heavy chain (Dictyostelium discoideum) r[A44357][A44357] dynein heavy chain, cytosolic - slime mold cytoskeleton (dictyodema)	52	36	189
96	10	10005	10664	gi 1408485	B650 gene product (Bacillus subtilis)	52	26	660
103	5	3986	3351	gi 1009168	respiratory nitrate reductase (Bacillus subtilis)	52	42	636
109	3	4102	3350	gi 699274	lmbE gene product (Mycobacterium leprae)	52	39	753
109	19	15732	17300	gi 1526981	amino acid permease YeeF like protein (Salmonella typhimurium)	52	30	1569
121	3	1412	981	gi 722931	unknown (Saccharomyces cerevisiae)	52	32	432
125	3	865	1680	gi 1296975	put gene product (Porphyromonas gingivalis)	52	38	816
130	2	659	1807	gi 1256634	25.8% identity over 120 aa with the Synecococcus sp. MpeV protein; putative (Bacillus subtilis)	52	36	1149
149	1	1164	583	gi 1225943	P88X terminase (Bacillus subtilis)	52	31	582
149	14	4687	4415	gi 1510368	M. jannaschii predicted coding region M0372 (Methanococcus jannaschii)	52	35	273

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
167	1	216	1001	gi 146025	cell division protein [Escherichia coli]	52	43	786
188	1	120	1256	gi 474915	orf 337; translated orf similarity to SW: BCR ECOLI bicyclicacycline resistance protein of Escherichia coli [Coxiella burnetii] pit[S44207/S44207] hypothetical protein 337 - Coxiella burnetii (SUB-138)	52	26	1137
195	9	9161	8760	gi 3028	mitochondrial outer membrane 72K protein [Neurospora crassa] r[A3682/A3682 72K mitochondrial outer membrane protein - rosepora crassa]	52	25	402
200	3	2065	2607	gi 142439	ATP-dependent nuclease [Bacillus subtilis]	52	35	543
203	4	2776	3684	gi 1303698	BtCD [Bacillus subtilis]	52	25	909
227	8	5250	5651	gi 305080	myosin heavy chain [Entamoeba histolytica]	52	24	402
242	1	21	1424	gi 1060877	[Eav] [Escherichia coli]	52	32	1404
249	5	4526	4753	pir C37222/C372	cytochrome P450 1A1, hepatic - dog (fragment)	52	23	228
255	1	2107	1055	gi 143290	penicillin-binding protein [Bacillus subtilis]	52	28	1053
276	7	3963	3664	gi 1001610	hypothetical protein [Synchocystis sp.]	52	30	300
276	8	4456	4055	gi 416235	orf L3 [Mycoplasma capricolum]	52	26	402
289	2	1856	1449	gi 150900	GTP phosphohydrolase [Proteus vulgaris]	52	34	408
325	1	1	279	gi 1204874	polypeptide deformylase (formylmethionine deformylase) [Haemophilus influenzae]	52	33	279
340	1	2017	1010	gi 1215695	peptide transport system protein SapF homolog; SapF homolog [Mycoplasma pneumoniae]	52	33	1008
375	3	340	1878	gi 467446	[similar to SpoVB [Bacillus subtilis]	52	28	1539
424	4	4104	3262	gi 1478239	[unknown [Mycobacterium tuberculosis]	52	34	843
430	1	3	575	pir A42606/A426	orfA 5' to orf405 - Saccharopolyspora erythraea (fragment)	52	28	573
444	4	4728	3712	gi 1408494	homologous to penicillin acylase [Bacillus subtilis]	52	31	1017
465	1	1802	903	gi 143331	alkaline phosphatase regulatory protein [Bacillus subtilis] pir[A27650/A27650 regulatory protein phoR - Bacillus subtilis sp P33545 PHO_R_BACSU ALKALINE PHOSPHATASE SYNTHESIS SENSOR PROTEIN PHO R (BC 2.7.3.-)]	52	36	900
469	5	4705	4169	gi 755152	highly hydrophobic integral membrane protein [Bacillus subtilis] sp P42953 TAGO_BACSU TEICHOIC ACID TRANSLOCATION PERMEASE PROTEIN AGO.	52	32	537
495	1	1262	633	gi 1204607	transcription activator [Haemophilus influenzae]	52	25	630
505	7	6004	5762	gi 142440	ATP-dependent nuclease [Bacillus subtilis]	52	28	243

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
517	2	1162	gi166162	Bacteriophage phi-11 int gene activator [Staphylococcus acterlophage phi 11]	52	35	453
543	2	444	gi1215693	putative orf; GT9 orf434 [Mycoplasma pneumoniae]	52	25	852
586	1	1	gi1581648	lepB gene product [Staphylococcus epidermidis]	52	36	336
773	1	848	gi1279769	fdhC [Methanobacterium thermoformicicum]	52	30	423
1120	2	100	gi142439	ATP-dependent nuclease [Bacillus subtilis]	52	35	231
1614	1	691	gi289262	comE ORF [Bacillus subtilis]	52	28	345
2495	1	1	gi216151	DNA polymerase (gene L; ttg start codon) [Bacteriophage SP02] gi1579197 SP02 DNA polymerase (aa 1-648) [Bacteriophage SP02] p1r[A21498] [UBPS2 DNA-directed DNA polymerase (EC 2.7.7.7) - phage P02]	52	34	324
2931	1	566	gi1256136	YbhG [Bacillus subtilis]	52	30	282
2943	1	577	gi141713	hlaA ORF (AA 1-245) [Escherichia coli]	52	35	258
2993	1	588	gi1298032	EF [Streptococcus suis]	52	34	294
3667	1	612	gi1849025	hypothetical 64.7-kDa protein [Bacillus subtilis]	52	36	306
3944	1	478	gi1218040	RAA [Bacillus licheniformis]	52	36	219
3954	2	613	gi1854064	U87 [Human herpesvirus 6]	52	50	267
3986	1	90	gi11205919	Na ⁺ and Cl ⁻ dependent gamma-aminobutyric acid transporter [Haemophilus influenzae]	52	33	312
4002	1	3	gi140003	oxoglutarate dehydrogenase (NADP+) [Bacillus subtilis] p123129 [ODOL-BACSU 2-OXOGLOUTARATE DEHYDROGENASE E1 COMPONENT (EC 2.4.2) (ALPHA-KETOGLUTARATE DEHYDROGENASE)]	52	42	387
4020	1	1	gi1159388	ornithine decarboxylase [Leishmania donovani]	52	47	249
4098	1	438	gi1409795	No definition line found [Escherichia coli]	52	32	219
4248	1	3	gi1965077	AdpP [Saccharomyces cerevisiae]	52	40	210
7	1	3	gi1895747	putative cel operon regulator [Bacillus subtilis]	51	28	573
21	4	2479	gi11510962	indole-3-glycerol phosphate synthase [Methanococcus jannaschii]	51	32	798
22	9	5301	gi11303933	VqjN [Bacillus subtilis]	51	25	666
43	3	1516	gi11519460	Srpl [Schizosaccharomyces pombe]	51	31	234
44	17	11042	gi142011	moaD gene product [Escherichia coli]	51	35	264
51	11	6453	gi1495471	vacuolating toxin [Helicobacter pylori]	51	37	279

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
52	4	2537	2995	gi1256652	25% identity to the E.coli regulatory protein MprA; putative [Bacillus subtilis]	51	32	459
57	10	7331	6843	gi1508173	81% domain of PTS-dependent GAT transport and phosphorylation [Escherichia coli]	51	32	489
59	1	29	1111	gi1299163	alanine dehydrogenase [Bacillus subtilis]	51	33	1083
67	20	15791	16576	gi1510977	M. jannaschii predicted coding region M0938 [Methanococcus jannaschii]	51	24	786
69	2	1559	1218	gi1467359	unknown [Bacillus subtilis]	51	34	342
71	1	3	1196	gi1298032	EF [Streptococcus suis]	51	32	1194
78	2	349	176	gi11161242	proliferating cell nuclear antigen [Styela clava]	51	28	174
99	4	3357	4040	gi1642795	TFIID subunit TAF155 [Homo sapiens]	51	25	684
109	1	2852	1428	gi1580920	rodD (gtaA) polypeptide (AA 1-673) [Bacillus subtilis] p[806048]S06048 probable rodD protein - Bacillus subtilis sp[133484]TAGE-BACSU PROBABLE POLY(GLYCEROL-PHOSPHATE) LPHAS-OLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS PROTEIN E)	51	27	1425
109	9	6007	6693	gi11204815	hypothetical protein (SP:P32662) [Haemophilus influenzae]	51	23	687
112	3	1066	2352	gi1505330[S053]	maltoase-binding protein precursor - Enterobacter aerogenes	51	42	1287
112	13	14432	12855	gi1405857	yeu [Escherichia coli]	51	29	1578
114	9	9725	8967	gi1435098	orf1 [Mycoplasma capricolum]	51	30	759
115	1	1	912	gi1141110	ORF YH085w [Saccharomycos euviolau]	51	25	912
127	10	9647	10477	gi11204314	M. influenzae predicted coding region H10056 [Haemophilus influenzae]	51	37	831
152	9	6814	7356	gi1431929	Muni regulatory protein [Mycoplasma sp.]	51	38	543
154	2	575	1153	gi11237044	unknown [Mycobacterium tuberculosis]	51	36	579
154	7	6587	5634	gi1409286	barU [Bacillus subtilis]	51	27	954
171	8	6943	6236	gi11205484	hypothetical protein (SP:P3918) [Haemophilus influenzae]	51	32	708
184	1	1	291	gi1466886	B1496_C1_206 [Mycobacterium leprae]	51	33	291
212	5	1501	2139	gi145605[A456]	mature-parasite-infected erythrocyte surface antigen MESA - Plasmodium falciparum	51	23	639
228	2	707	1378	gi18204	nuclear protein [Drosophila melanogaster]	51	27	672
236	8	8137	7483	gi149272	Asparaginase [Bacillus licheniformis]	51	31	657
243	4	4637	3546	gi11511102	melvalonate kinase [Methanococcus jannaschii]	51	29	1092

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match region	Match gene name	% sim	% ident	length (nt)
257	4	3540	3373	gi 204579	H. influenzae predicted coding region H10326 [Haemophilus influenzae]	51	22	168
258	3	2397	1609	gi 160299	glutamic acid-rich protein [Plasmodium falciparum] pfrA54514[A54514]	51	34	789
265	5	2419	3591	gi 580841	PI [Bacillus subtilis]	51	32	1173
298	2	518	748	gi 1336162	SCPb [Streptococcus agalactiae]	51	34	231
316	9	5817	7049	gi 413953	Ipa-29d gene product [Bacillus subtilis]	51	19	1233
332	2	3775	2057	gi 1209012	mutS [Thermus aquaticus thermophilus]	51	26	1719
364	4	3816	4991	gi 528991	unknown [Bacillus subtilis]	51	32	1176
440	2	448	684	gi 2819	transferase [GM10] (AA 1 - 687) [Kluyveromyces lactis] r[S01407]KUVK	51	32	237
					UDPglucose 4-epimerase (EC 5.1.3.2) - yeast uyveromyces marxianus var. lactis)			
495	2	1353	1177	gi 297861	protease G [Erwinia chrysanthemi]	51	41	177
495	3	2287	1718	gi 1513317	serine rich protein [Entamoeba histolytica]	51	25	570
506	1	840	421	gi 455320	ctfI protein [Bacteriophage P4]	51	33	420
600	1	1474	983	gi 587532	orf, len: 201, CAI: 0.16 [Saccharomyces cerevisiae] pit[S48818]S48818	51	30	492
					hypothetical protein - yeast [Saccharomyces erevisiae]			
607	3	479	934	gi 3511524	hypothetical protein (SP-P37002) [Methanococcus jannaschii]	51	40	456
686	2	127	600	gi 493017	endocarditis specific antigen [Enterococcus faecalis]	51	30	474
726	1	33	230	gi 1353851	unknown [Prochlorococcus marinus]	51	45	198
861	1	176	652	gi 410145	dehydroquinase dehydratase [Bacillus subtilis]	51	34	477
869	1	782	393	gi 40100	rodC (tag) polypeptide (AA 1-746) [Bacillus subtilis] ir[S06049]S06049	51	23	390
					rodC protein - Bacillus subtilis p[13485]TAGF_BACSU TECHNIC ACID BIOSYNTHESIS PROTEIN F.			
1003	1	642	322	gi 1279707	hypothetical phosphoglycerate mutase [Saccharomyces cerevisiae]	51	39	321
1046	2	866	624	gi 510257	glycosyltransferase [Escherichia coli]	51	29	243
1467	1	702	352	gi 1511175	M. jannaschii predicted coding region MJ1177 [Methanococcus jannaschii]	51	32	351
2558	1	457	230	sp P10582 DPOH_	DNA POLYMERASE (EC 2.7.7.7) (S-1 DNA ORF 3)	51	26	228
3003	1	779	399	gi 809543	CbrC protein [Erwinia chrysanthemi]	51	27	381
3604	1	1	399	gi JC6210 JC42	J-hydroxycyl-CoA dehydrogenase (EC 1.1.1.35) - mouse	51	37	399
3732	1	2	316	gi 145906	acyl-CoA synthetase [Escherichia coli]	51	33	315

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
3791	1	2	274	gi 1061351	semaphorin III family homolog (Homo sapiens)	51	37	273
3995	1	46	336	gi 216346	surfactin synthetase (Bacillus subtilis)	51	38	291
4193	1	612	307	gi 42749	ribosomal protein L12 (AA 1-179) (Escherichia coli) fr[S04776]XREIPL peptide N-acetyltransferase (tmL DEC 2.3.1.-) - chericchia coli	51	25	306
4539	1	367	185	gi 1408494	homologous to penicillin acylase (Bacillus subtilis)	51	40	183
4562	1	442	239	gi 1458280	coded for by C. elegans cDNA cno1a7; Similar to hydroxymethylglutaryl-CoA synthase (Caenorhabditis elegans)	51	35	204
1	4	3576	4859	gi 559160	GRAFL score: null; cap site and late promoter motifs present pstream; putative (Autographa californica nuclear polyhedrosis virus)	50	44	1284
11	7	4044	5165	gi 1146207	putative (Bacillus subtilis)	50	35	1122
11	13	10509	9496	gi 1208451	hypothetical protein (Synecocystis sp.)	50	39	1014
19	3	2034	1018	gi 411966	ipa-42d gene product (Bacillus subtilis)	50	29	1017
20	11	8586	8407	gi 1332159	ORF YGR103W (Saccharomyces cerevisiae)	50	28	180
24	5	5408	4824	gi 496280	structural protein (Bacteriophage Tuc2009)	50	29	585
34	4	1926	2759	gi 1303966	YqjO (Bacillus subtilis)	50	36	834
38	30	22865	23440	gi 1072179	Similar to dihydroflavonol-4-reductase (maize, petunia, tomato) (Caenorhabditis elegans)	50	32	576
47	2	1705	2976	gi 153015	FemA protein (Staphylococcus aureus)	50	29	1272
56	13	15290	15841	gi 606096	ORF_f167; end overlaps end of o100 by 14 bases; start overlaps f174, ther starts possible (Escherichia coli)	50	30	552
57	1	2135	1077	gi 440922	xyitol dehydrogenase (unidentified hemiascomycete)	50	29	1059
58	2	628	1761	gi 143725	putative (Bacillus subtilis)	50	29	1134
88	6	4393	3884	gi 1072179	Similar to dihydroflavonol-4-reductase (maize, petunia, tomato) (Caenorhabditis elegans)	50	32	510
89	5	3700	3356	gi 1276658	ORF174 gene product (Porphyra purpurea)	50	25	345
141	1	3	239	gi 476024	[carbamoyl phosphate synthetase II (Plasmodium falciparum)]	50	33	237
151	1	186	626	gi 1403441	unknown (Mycobacterium tuberculosis)	50	35	441
166	7	11065	9623	gi 895747	putative cel operon regulator (Bacillus subtilis)	50	32	1443
201	6	5284	5096	gi 160229	circumsporozoite protein (Plasmodium reichenowi)	50	42	189
206	22	30784	29555	gi 1052754	LarP integral membrane protein (Lactococcus lactis)	50	24	1230

TABLE 2

S aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
211	4	1523	1927	gi 410131	ORF7 [Bacillus subtilis]	50	29	405
214	4	2411	3295	sp P37348 VECE	HYPOTHETICAL PROTEIN IN ASP5 5' REGION [FRAGMENT]	50	37	885
228	7	5068	4406	gi 313580	envelope protein [human immunodeficiency virus type 1] p1r[S35835]S35835 envelope protein - human immunodeficiency virus type 1 (fragment) (SUB 1-77)	50	35	663
272	2	3048	1723	gi 1408485	B65G gene product [Bacillus subtilis]	50	22	1326
273	2	1616	984	gi 184186	phosphoglycerate mutase [Saccharomyces cerevisiae]	50	28	613
328	2	2507	1605	gi 148896	flippase [Haemophilus influenzae]	50	26	903
332	4	5469	3802	gi 1526547	DNA polymerase family X [Thermus aquaticus]	50	27	1668
342	5	1473	1931	gi 456562	G-box binding factor [Dictyostelium discoideum]	50	35	459
352	1	1478	741	gi 288301	ORF2 gene product [Bacillus megaterium]	50	29	738
408	7	5299	5523	gi 111665	ORF2116 [Marchantia polymorpha]	50	27	225
420	3	650	1825	gi 1757842	UDP-sugar hydrolase [Escherichia coli]	50	30	1176
464	1	1	591	gi 487282	Na ⁺ -ATPase subunit J [Enterococcus hirae]	50	29	591
472	2	1418	864	gi 551875	BglR [Lactococcus lactis]	50	23	555
520	1	23	541	gi 567036	CapE [Staphylococcus aureus]	50	27	519
529	1	6	410	gi 1256632	25% identity to the E.coli regulatory protein MprA; putative [Bacillus subtilis]	50	34	405
534	5	7726	6059	gi 1295671	selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependent RNA polymerase I and III [Saccharomyces cerevisiae]	50	18	1668
647	1	2990	1497	gi 405568	TraI protein shares sequence similarity with a family of topoisomerases [Plasmid pSK41]	50	31	1494
664	3	1133	711	gi 410007	leukocidin F component [Staphylococcus aureus, MRSA No. 4, Peptide, 23 aa]	50	32	423
678	1	1	627	gi 138032	EF [Streptococcus suis]	50	29	627
755	3	947	1171	gi 150572	cytochrome c1 precursor [EC 1.10.2.2] [Paracoccus denitrificans] gi 45465 cytochrome c1 (AA 1-450) [Paracoccus denitrificans] p1r[C29413]C29413 ubiquinol--cytochrome-c reductase [EC 1.10.2.2] cytochrome c1 precursor - Paracoccus denitrificans sp p1627 CV1	50	37	225
827	1	1363	683	gi 142020	heterocyst differentiation protein [Anabaena sp.]	50	21	681
892	1	3	752	gi 408485	B65G gene product [Bacillus subtilis]	50	27	750
910	2	438	887	gi 104727	tyrosine-specific transport protein [Haemophilus influenzae]	50	25	450

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	match gene name	% sim	% ident	length (nt)
933	1	524	760	gi1205451	cell division inhibitor [Haemophilus influenzae]	50	32	237
973	1	424	236	gi1886947	orf3 gene product [Saccharomyces cerevisiae]	50	40	189
1009	1	653	429	gi1153727	M protein (group G streptococcus)	50	28	225
1027	1	511	257	gi1419314	[ipa-10r gene product (Bacillus subtilis)]	50	25	255
1153	2	556	326	gi1773676	nccA [Alcaligenes xylosoxydans]	50	36	231
1222	1	798	400	gi1408485	B65G gene product (Bacillus subtilis)	50	21	399
1350	1	692	399	gi1289272	ferrichrome-binding protein [Bacillus subtilis]	50	32	294
2945	1	366	184	gi1171704	hexaprenyl pyrophosphate synthetase (COQ1) [Saccharomyces erevisiae]	50	34	183
2968	2	1604	804	gi1397526	clumping factor [Staphylococcus aureus]	50	33	801
2998	2	657	394	gi1495696	F54E7.3 gene product [Caenorhabditis elegans]	50	40	264
3046	2	506	306	gi15138191S138	acyl carrier protein - Anabaena variabilis (fragment)	50	32	201
3063	1	547	275	gi1174190	lucA gene product [Escherichia coli]	50	29	273
3174	1	3	146	gi151900	alcohol dehydrogenase [Rhodobacter sphaeroides]	50	31	144
3792	1	625	314	gi11001423	hypothetical protein [Synecocystis sp.]	50	35	312
3800	1	2	262	gi1144733	NAD-dependent beta-hydroxybutyryl coenzyme A dehydrogenase Clostridium acetobutylicum	50	28	261
3946	1	373	188	gi1576765	cytochrome b [Myrmecia pilosula]	50	38	186
3984	1	578	291	sp P37348 YECC_	HYPOTHETICAL PROTEIN IN ASP5 5'-REGION (FRAGMENT)	50	37	288
37	10	8250	7885	gi1204367	hypothetical protein (GB:U14003.278) [Haemophilus influenzae]	49	30	366
46	16	13802	14848	gi1466860	lact: B1308_F1_34 [Mycobacterium leprae]	49	24	1047
59	5	2267	3601	gi1606304	ORF_0462 [Escherichia coli]	49	27	1335
112	118	17884	18615	gi1559502	N04 protein (AA 1 - 409) [Caenorhabditis elegans]	49	25	732
138	9	6973	7902	gi1303953	esterase [Acinetobacter calcoaceticus]	49	29	930
217	6	4401	5138	gi1496254	[fibronectin/fibrinogen-binding protein (Streptococcus pyogenes)]	49	31	738
220	12	11803	12657	gi1397526	clumping factor [Staphylococcus aureus]	49	31	855
228	4	1842	2492	gi1523692S236	hypothetical protein 9 - Plasmodium falciparum	49	24	651
268	1	5016	2614	gi143047	ORF8 [Bacillus subtilis]	49	26	2403

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
271	2	1164	1373	gi 1001257	hypothetical protein [Synecocystis sp.]	49	38	210
300	3	4340	3180	gi 1510796	hypothetical protein (GPX91006_2) [Methanococcus jannaschii]	49	26	1161
381	1	2281	1142	gi 396301	matches ES0041: Bacterial regulatory proteins, arae family signature [Escherichia coli]	49	29	1140
466	1	3	947	gi 303863	Yggp [Bacillus subtilis]	49	26	945
666	1	379	191	gi 633112	ORF1 [Streptococcus sobrinus]	49	29	189
670	2	403	1014	gi 1122758	unknown [Bacillus subtilis]	49	32	612
709	1	1433	795	gi 143830	XPAC [Bacillus subtilis]	49	29	639
831	1	943	473	gi 401786	phosphomannomutase [Mycoplasma pirum]	49	29	471
1052	1	422	213	gi 1303799	YgeN [Bacillus subtilis]	49	21	210
1800	1	342	172	gi 216300	peptidoglycan synthase enzyme [Bacillus subtilis] sp p37585 HURG_BACSU HURG PROTEIN UPD-N-ACETYLGLUCOSAMINE-N-ACETYLURAMYL-PENTAPEPTIDE PYROPHOSPHORYL-UNDECAPRENOL N-ACETYLGLUCOSAMINE RANSFERASE)	49	28	171
2430	1	2	376	sp P27434 YFGA_	HYPOTHETICAL 36.2 KD PROTEIN IN NDK-OCPE INTERGENIC REGION.	49	26	375
3096	1	542	273	gi 516360	surfactin synthetase [Bacillus subtilis]	49	25	270
32	4	3771	3100	gi 1217963	hepatocyte nuclear factor 4 gamma [HNF4gamma] [Homo sapiens]	48	36	672
38	1	1	609	gi 1205790	H. influenzae predicted coding region H1555 [Haemophilus influenzae]	48	28	609
45	6	5021	6427	gi 1534267	unknown [Mycobacterium tuberculosis]	48	20	1407
59	14	16346	31096	gi 1197336	ImpJ protein [Mycoplasma hominis]	48	28	14751
61	1	3	608	gi 1511555	quinolone resistance norA protein protein [Methanococcus jannaschii]	48	30	606
61	3	3311	3646	gi 1303893	YghU [Bacillus subtilis]	48	29	336
114	1	98	415	gi 671708	[sua] homolog: similar to Drosophila melanogaster suppressor of able [sua]) protein, Swiss-Prot Accession Number P22293 [Drosophila virilis]	48	25	318
121	1	1131	610	gi 1314584	unknown [Sphingomonas S88]	48	29	522
136	1	2014	1280	gi 1205968	H. influenzae predicted coding region H11718 [Haemophilus influenzae]	48	21	735
171	10	8220	9557	gi 1208454	hypothetical protein [Synecocystis sp.]	48	34	1338
175	1	3625	1814	gi 396400	similar to eukaryotic Na+/H. exchanger [Escherichia coli] sp p32703 XUCE_ECOLI HYPOTHETICAL 60.5 KD PROTEIN IN SOXR-ACS INTERGENIC REGION (0549).	48	29	1812
194	1	2	385	gi 1510493	H. jannaschii predicted coding region H20419 [Methanococcus jannaschii]	48	25	384

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
197	1	901	452	gi 1045716	ispermidine/putrescine transport ATP-binding protein [Mycoplasma genitalium]	48	25	450
203	1	1	396	gi 940288	protein localized in the nucleoli of pea nuclei; ORF, putative Pisum sativum	48	29	396
204	1	1363	698	gi 529202	No definition line found [Caenorhabditis elegans]	48	25	666
206	20	14815	27760	gi 511490	gramicidin S synthetase 2 [Bacillus brevis]	48	27	7056
212	1	2	166	gi 295899	nucleolin [Menopus laevis]	48	34	165
220	10	12652	11426	gi 44073	[SecY protein [Lactococcus lactis]	48	23	1227
243	6	6450	5491	gi 1184118	mevalonate kinase [Methanobacterium thermoautotrophicum]	48	30	960
264	4	5434	3308	gi 1015903	ORF YJR15C [Saccharomyces cerevisiae]	48	26	2127
441	1	1532	768	gi 142863	replication initiation protein [Bacillus subtilis] pir B26580 B26580 replication initiation protein - Bacillus subtilis	48	23	765
444	5	3898	5298	gi 145816	putative [Escherichia coli]	48	24	1401
484	2	388	1110	gi 146551	transmembrane protein (kdpD) [Escherichia coli]	48	18	723
542	3	1425	2000	gi 528969 5289	N-carbamoylserine amidohydrolase (EC 3.5.1.59) - Arthrobacter sp.	48	27	576
566	1	3	1019	gi 153490	tetracycline C resistance and export protein [Streptomyces laevis]	48	24	1017
611	1	2	730	gi 1103507	unknown [Schizosaccharomyces pombe]	48	38	729
624	1	1255	665	gi 144859	ORF B [Clostridium perfringens]	48	26	591
846	1	1014	508	gi 537506	paramyosin [Drosophila melanogaster]	48	27	507
1020	1	66	950	gi 1499876	meqasium and cohesin transport protein [Methanococcus jannaschii]	48	30	885
1227	1	1	174	gi 493730	lipoygenase [Pisum sativum]	48	35	174
1266	1	1	405	gi 982452	ORF_F211; alternate name y99h; orf5 of X14436 [Escherichia coli] gi 41425 ORF5 (AA 1-197) [Escherichia coli] (SUB 15-211)	48	24	405
2071	1	707	381	gi 1408486	HS74A gene product [Bacillus subtilis]	48	25	327
2198	1	463	233	gi 1500401	reverse gyrase [Methanococcus jannaschii]	48	40	231
2425	1	476	246	pir H48563 H485	[C1 protein - fowlpox virus (strain HP444) (fragment)]	48	40	231
2432	1	446	225	gi 1353703	Trfo [Homo sapiens]	48	33	222
2453	1	794	399	gi 142850	division initiation protein [Bacillus subtilis]	48	29	396
2998	1	469	236	gi 577569	pepV [Lactobacillus delbrueckii]	48	31	234

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
3042	1	16	280	gi 945219	lucin [Homo sapiens]	48	35	267
3686	1	1	405	gi 145836	putative [Escherichia coli]	48	25	405
4027	2	492	301	pir-S51177 S511	trans-activator protein - Equine infectious anemia virus	48	32	192
4	2	3641	2232	gi 1103989	YqkI [Bacillus subtilis]	47	24	1410
24	2	599	1084	gi 340083	PCA-1 gene product [Bradysia hygida]	47	28	486
36	10	7524	6925	gi 1109223	esterase [Acinetobacter lwoffii]	47	26	600
43	2	196	1884	gi 1403455	unknown [Mycobacterium tuberculosis]	47	27	1689
44	22	16118	15108	gi 1511555	quinolone resistance nraA protein protein [Methanococcus jannaschii]	47	31	1011
69	7	7141	6710	gi 438466	Possible operon with nrfG. Hydrophilic, no homologue in the database; putative [Bacillus subtilis]	47	29	432
81	4	5022	4279	gi 466882	lpvsl: B1496_C2_189 [Mycobacterium leprae]	47	24	744
120	12	9135	8863	gi 927340	09509.27p; CAI: 0.12 [Saccharomyces cerevisiae]	47	38	273
142	1	2022	1174	gi 486143	ORF YKL094w [Saccharomyces cerevisiae]	47	32	849
168	1	2178	1093	gi 1177254	hypothetical scsB protein [Bacillus subtilis]	47	29	1086
263	1	1884	943	gi 142822	10-alanine racemase cds [Bacillus subtilis]	47	34	942
279	1	1109	561	gi 516608	2 predicted membrane helices, homology with B. subtilis non Orf3 Ecoulard et. al unpublished Accession number M74183, approximately 1 minutes on updated Radd map; putative [Escherichia coli] sp P37353 YFBB_ECOLI_HYPOTHETICAL_26.7 KD PROTEIN IN WEND-WEND	47	31	549
345	2	2620	1676	gi 1204835	hippuricase [Haemophilus influenzae]	47	28	945
389	2	152	400	gi 456562	G-box binding factor [Dictyostellium discoideum]	47	32	249
391	1	1	831	gi 1420856	myo-inositol transporter [Schizosaccharomyces pombe]	47	19	831
404	3	2072	2773	gi 1255425	C33G8.2 gene product [Caenorhabditis elegans]	47	17	702
529	5	2145	3107	gi 1103973	YqjV [Bacillus subtilis]	47	29	963
565	2	2321	1257	gi 142824	processing protease [Bacillus subtilis]	47	28	1065
654	1	962	463	gi 243353	ORF 5' of ECRF3 (herpesvirus salmairi IVS, host-squirrel monkey, epitide, 407 aa)	47	23	480
692	1	115	633	gi 150756	40 kDa protein [Plasmid pOH1]	47	25	519
765	1	1634	819	gi 1256621	26.7% of identity in 165 aa to a Thermophilic bacterium hypothetical protein 6; putative [Bacillus subtilis]	47	28	816

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
825	2	211	1023	gi 397526	clumping factor (Staphylococcus aureus)	47	32	813
914	1	1	615	gi 558073	polymorphic antigen (Plasmodium falciparum)	47	29	615
1076	1	1	753	gi 1147557	Aspartate aminotransferase (Bacillus circulans)	47	33	753
1351	1	793	398	gi 755153	ATP-binding protein (Bacillus subtilis)	47	20	396
4192	1	3	293	gi 145836	putative (Escherichia coli)	47	24	291
5	6	4708	4361	gi 305080	myosin heavy chain (Entamoeba histolytica)	46	30	348
11	4	2777	3058	gi 603639	Yel040p (Saccharomyces cerevisiae)	46	28	282
46	11	10518	10300	gi 1246901	ATP-dependent DNA ligase (Candida albicans)	46	28	219
61	4	3941	7930	gi 298032	EF (Streptococcus suis)	46	35	3990
132	4	5028	4093	gi 1511057	hypothetical protein SP:P45869 (Methanococcus jannaschii)	46	25	936
170	4	4719	3652	pir 551910 S519	G4 protein - Sauroleishmania tarentolae	46	26	1068
191	7	9543	8284	gi 1041334	F5405.7 (Caenorhabditis elegans)	46	25	1260
253	1	1	396	gi 1204449	dihydrolipamide acetyltransferase (Haemophilus influenzae)	46	35	396
264	3	437	973	gi 180189	cerebellar-degeneration-related antigen (CDR34) (Homo sapiens) gi 182737 cerebellar degeneration-associated protein (Homo sapiens) pir A29770 A29770 cerebellar degeneration-related protein - human	46	29	537
273	1	485	285	gi 607573	envelope glycoprotein C20 region (Human immunodeficiency virus type 1)	46	35	201
350	1	3	563	gi 537052	ORF_f286 (Escherichia coli)	46	35	561
384	1	2	862	gi 1221884	(urea7) amidolyase (Haemophilus influenzae)	46	31	861
410	4	1876	2490	gi 1110518	proton antiporter efflux pump (Mycobacterium smegmatis)	46	24	615
432	1	2663	1455	gi 1197634	orf4: putative transporter: Method: conceptual translation supplied by author (Mycobacterium smegmatis)	46	27	1209
458	1	2419	1211	gi 15470	portal protein (Bacteriophage SP1)	46	30	1209
517	5	2477	4192	gi 1523812	orf5 (Bacteriophage A2)	46	23	1716
540	3	1512	1285	gi 215635	pacA (Bacteriophage P1)	46	30	228
587	2	649	1242	gi 537148	ORF_f181 (Escherichia coli)	46	29	594
1218	1	747	391	gi 1205456	single-stranded-DNA-specific exonuclease (Haemophilus influenzae)	46	30	357

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match region	match gene name	% sim	% ident	length (nt)
3685	1	1	402	gi1450688	hsm gene of Ecoprr1 gene product [Escherichia coli] pir138437138437 hsm protein - Escherichia coli pir15096291509629 hypothetical protein A - Escherichia coli (SUB 40-520)	46	33	402
4176	1	673	338	gi1351460	[FIM-C.1 gene product (Xenopus laevis)]	46	31	336
37	7	4813	5922	gi1606064	ORF 408 [Escherichia coli]	45	24	1110
38	16	11699	12004	gi1452192	protein tyrosine phosphatase (PTP-BAS, type 2) (Homo sapiens)	45	24	306
87	2	1748	2407	gi11064813	homologous to sp:PHOR_BACSU [Bacillus subtilis]	45	23	660
103	12	14182	13385	gi11001307	hypothetical protein (Synecocystis sp.)	45	22	798
112	14	14791	13811	gi11204389	[H. influenzae predicted coding region H10131 [Haemophilus influenzae]]	45	23	981
145	4	4483	3461	gi1220578	open reading frame [Mus musculus]	45	20	1023
170	6	6329	4965	gi1238657	AppCyclochrome d oxidase, subunit 1 homolog [Escherichia coli, K12, optide, 514 aa]	45	27	1365
206	2	5230	4346	gi11222056	aminotransferase [Haemophilus influenzae]	45	27	885
228	1	60	716	gi1160299	glutamic acid-rich protein (Plasmodium falciparum) pir13465141346514 glutamic acid-rich protein precursor - Plasmodium falciparum	45	23	657
288	1	2	1015	gi11255425	[C33GR.2 gene product (Caenorhabditis elegans)]	45	23	1014
313	3	4339	3128	gi1581140	[NADH dehydrogenase [Escherichia coli]]	45	30	1212
332	1	914	459	gi1870966	[F47A4.2 (Caenorhabditis elegans)]	45	20	456
344	1	3	221	gi1171225	kinesin-related protein [Saccharomyces cerevisiae]	45	26	219
441	2	1501	1073	gi1142863	replication initiation protein [Bacillus subtilis] pir18265801826580 replication initiation protein - Bacillus subtilis	45	27	429
672	1	2	982	gi11511334	[W. jannaschii predicted coding region WJ1323 (Methanococcus jannaschii)]	45	22	981
763	3	1345	851	gi1606180	ORF_f310 [Escherichia coli]	45	24	495
886	3	379	846	gi1726426	similar to protein kinases and C. elegans proteins F37C12.8 and 37C12.5 [Caenorhabditis elegans]	45	30	468
948	1	3	473	gi1156400	myosin heavy chain (isozyme unc-54) [Caenorhabditis elegans] pir13919581391958 myosin heavy chain B - Caenorhabditis elegans sp102566102566 MYOSIN HEAVY CHAIN B (HNC B)	45	25	471
1158	1	2	376	gi11441155	transmission-blocking target antigen [Plasmodium falciparum]	45	35	375
2551	1	4	285	gi11276705	ORF287 gene product [Porphyra purpurea]	45	28	282
3967	1	42	374	gi1976025	[HraA [Escherichia coli]]	45	28	333

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	Match accession	Match gene name	% sim	% ident	length (nt)
52	7	6931	5846	gi 467378	unknown [Bacillus subtilis]	44	22	1086
138	8	6475	6849	gi 173028	thioredoxin II [Saccharomyces cerevisiae]	44	28	375
221	5	7032	5617	gi 153490	tetracenomycin C resistance and export protein [Streptomyces laevis]	44	21	1416
252	2	1331	1122	gi 1204989	hypothetical protein [GB:000022_9] [Haemophilus influenzae]	44	30	210
263	2	1265	2093	gi 1136221	carboxypeptidase [Sulfolobus solfataricus]	44	26	1173
365	4	4963	3524	gi 1296822	orf1 gene product [Lactobacillus helveticus]	44	31	1440
543	3	1315	1833	gi 1063250	low homology to P20 protein of Bacillus licheniformis and bleomycin acetyltransferase of Streptomyces verticillius [Bacillus subtilis]	44	24	519
544	4	3942	4892	gi 951460	EFIM-C.1 gene product [Xenopus laevis]	44	32	951
792	1	1224	613	gi 205680	high molecular weight neurofilament [Rattus norvegicus]	44	28	612
44	18	11303	11911	gi 1511614	molybdopterin-guanine dinucleotide biosynthesis protein A [Methanococcus jannaschii]	43	27	609
59	8	3665	5128	gi 1151490	tetracenomycin C resistance and export protein [Streptomyces laevis]	43	21	1464
59	10	5536	7527	gi 153022	lipase [Staphylococcus epidermidis]	43	22	1992
99	1	1346	681	gi 11419051	unknown [Mycobacterium tuberculosis]	43	21	666
310	8	9402	12134	gi 397526	clumping factor [Staphylococcus aureus]	43	21	2733
412	3	2782	2303	pir A60540 A605	sporozoite surface protein 2 - Plasmodium yoelii (fragment)	43	29	480
519	3	2547	3122	sp 006530 DISU_1.8.2.-1 (FC) (FCSD)	SULFIDE DEHYDROGENASE (FLAVOCTOCHROME C) FLAVOPROTEIN CHAIN PRECURSOR (PC)	43	23	576
4	13	12053	13321	gi 295671	selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependent RNA polymerase I and III [Saccharomyces cerevisiae]	42	18	1269
94	2	1768	1091	gi 501027	ORF2 [Trypanosoma brucei]	42	31	678
127	4	5791	4550	gi 42029	ORF1 gene product [Escherichia coli]	42	21	1242
287	3	1515	1036	gi 142790	ORF1, putative [Bacillus firmus]	42	25	480
344	6	4097	3525	gi 40320	ORF 2 (AA 1-203) [Bacillus thuringiensis]	42	30	573
512	1	2167	1115	gi 405957	yeeF [Escherichia coli]	42	23	1053
631	1	2434	1223	gi 580920	rodD (gtaA) polypeptide (AA 1-673) [Bacillus subtilis] pir S06048 S06048 probable rodD protein - Bacillus subtilis sp P13484 TAGE_BACSU PROBABLE POLY(GLYCEROL-PHOSPHATE) LPHA-GLUCOSYLTRANSFERASE (EC 2.4.1.52) (TECHOIC ACID BIOSYNTHESIS PROTEIN E)	42	24	1212

TABLE 2

S. aureus - Putative coding regions of novel proteins similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)	match accession	match gene name	% sim	% ident	length (nt)
685	3	2359	1739	gi11303784	[vqeD [Bacillus subtilis]	42	19	621
4132	1	787	395	gi11022910	[protein tyrosine phosphatase [Dictyostelium discoideum]	42	25	393
86	2	1375	884	gi1309506	[spermidine/spermine N1-acetyltransferase [Mus saxicola] pir[S43430][S43430]	41	30	492
191	12	14797	14075	gi11124957	[spermidine/spermine N1-acetyltransferase - spiny ouse [Mus saxicola]			
212	6	2150	3127	gi135873	[orf4 gene product [Methanosarcina barkeri]	41	22	723
213	3	1263	2000	gi1633692	[observed 35.2kd protein [Mycobacteriophage 15]	41	26	978
408	4	2625	3386	gi11197634	[Trsa [Yersinia enterocolitica]	41	18	738
542	1	3	1103	gi1457146	[orf4; putative transporter; Method: conceptual translation supplied by author [Mycobacterium smegmatis]	41	24	762
924	1	2	475	gi1190148	[rhostry protein [Plasmodium yoelii]	41	21	1101
1562	1	1	402	gi1552184	[nucleolin - rat]	41	30	474
2395	1	518	261	gi1542251	[asparagine-rich antigen Pf35-2 [Plasmodium falciparum] pir[S27826][S27826]	40	20	402
4077	1	3	305	gi11055055	[asparagine-rich antigen Pf35-2 - Plasmodium falciparum (fragment)]			
918	3	1003	503	gi11255425	[hypothetical protein 5 - fowlpox virus]	40	18	258
59	12	8294	10636	gi1535260	[coded for by C. elegans cDNA YK3791.5; coded for by C. elegans cDNA YK5C9.5; coded for by C. elegans cDNA YK1A9.5; alternatively spliced form of F32C9.8b [Caenorhabditis elegans]	39	21	303
63	5	3550	8079	gi1298032	[C33G8.2 gene product [Caenorhabditis elegans]	37	25	501
544	3	2507	3601	gi11015903	[STARP antigen [Plasmodium reichenowi]	36	24	2343
63	4	1949	3574	gi1552195	[EF [Streptococcus suis]	36	19	4530
					[ORF YJRI51c [Saccharomyces cerevisiae]	35	22	1095
					[circumsporozoite protein [Plasmodium falciparum] spIP05691][CSP_PLAFL]	32	27	1626
					[CIRCUMSPOROZOITE PROTEIN (CS) (FRAGMENT)]			

TABLE 2

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
4	1	1234	692
4	3	1712	2278
4	4	3703	3032
4	14	13073	12585
5	2	2539	1601
5	3	1532	1771
5	7	4741	4550
5	9	7939	6422
5	12	8711	8547
6	4	2359	1982
8	1	349	176
11	8	5144	5983
11	9	5968	6498
11	10	6472	6284
11	16	10954	11271
12	5	5352	4942
12	6	4596	4862
15	3	1895	1650
16	10	11263	10835
18	2	1093	917
20	9	9125	7764
20	10	8571	8230
20	12	9201	8801
20	13	12158	10470
23	1	674	319
23	6	6138	5485
23	8	6376	5942

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Qatig ID	ORF ID	Start (nt)	Stop (nt)
23	9	7631	6881
23	15	12618	12830
24	4	4556	4185
24	6	5642	5241
25	2	1824	2402
31	2	505	849
31	3	1177	1524
31	4	2454	3005
32	2	765	1388
32	9	7952	8575
32	10	8591	8728
32	11	9738	9379
32	12	10797	10087
34	2	1315	1049
36	7	5226	5801
36	11	7575	7261
36	12	7428	7621
37	4	3158	2964
38	2	1585	980
38	11	8425	6868
38	20	16982	16371
38	26	20253	20804
38	27	20722	21264
39	1	1	627
40	1	805	404
43	1	796	428
44	4	2674	2324

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
44	5	2484	3263
44	14	10587	10129
44	20	13724	13536
44	21	13596	13994
45	7	6575	6297
46	8	6365	6520
46	12	10449	10976
46	17	15032	15424
47	1	288	1079
48	9	7620	7778
50	1	1612	962
50	2	1621	1316
51	1	738	370
51	5	2520	2245
53	1	442	287
53	7	6705	6319
54	7	9014	8709
55	1	592	326
55	3	1052	786
56	1	1	261
56	3	1551	1228
56	4	1970	1560
56	17	19092	18712
57	4	3694	3521
57	8	5436	5822
58	9	8885	8553
59	3	1365	1509

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
59	6	3026	2802
59	7	3770	3570
59	9	4946	4563
59	11	7518	8178
59	13	10401	11603
62	2	2696	1521
62	11	5440	5757
63	1	1	336
67	1	900	1781
67	2	1774	2610
67	3	2591	3904
67	8	7110	6955
68	1	78	326
70	6	6761	5199
70	11	8935	8645
77	3	1590	1192
79	2	1509	1228
79	3	1411	1791
83	1	2	403
85	9	8300	8653
85	10	8969	8781
86	3	1426	1232
87	8	9187	9166
88	3	1620	1922
89	1	3	161
89	7	5042	4878
91	1	1098	550

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
91	3	3938	3141
92	2	449	928
92	3	1958	1467
92	9	5638	6024
94	1	661	332
94	3	2445	1813
94	4	2583	2197
96	11	10601	11050
99	6	4672	4523
99	7	5014	4784
100	8	7658	7287
102	7	4697	4368
103	3	2496	2035
104	1	2	694
104	2	699	1277
105	1	1235	693
105	3	3233	2655
106	1	3	221
106	3	1209	1355
107	1	1081	542
109	4	4025	3651
109	13	11625	11996
109	14	11981	12268
109	20	17401	17688
110	1	2	760
114	10	8764	9384
116	1	1	309

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Coding No	ORF No	Start (nt)	Stop (nt)
116	3	6273	4662
116	8	11049	9976
116	9	10113	10158
120	5	3703	3320
120	6	4270	3869
120	13	9290	9844
121	2	417	569
126	3	1050	818
127	3	2648	3196
127	5	4084	4395
131	6	6773	6438
132	2	715	1695
134	1	2	667
135	2	512	258
135	3	1124	729
138	1	3	152
138	7	6008	6463
140	1	2060	1032
140	2	2019	1513
140	5	2387	2743
142	2	1360	2388
142	7	8830	7586
143	7	7290	6502
144	1	1227	640
146	1	2	511
146	3	502	1350
146	4	3673	2540

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
146	5	2874	3071
147	1	1	339
149	11	3956	3615
149	12	4036	3785
149	13	4507	4145
149	15	4807	4610
149	16	5495	5049
149	18	5739	5491
149	21	7416	7054
149	23	9216	8521
149	24	9681	9106
149	25	10679	9897
150	2	2103	1587
154	3	1795	1508
154	8	6586	6398
154	14	12706	12147
154	15	13531	12803
156	1	315	593
157	3	1183	2232
158	2	1471	1064
159	3	452	808
161	2	876	1808
161	6	4653	4279
161	7	4803	4540
161	8	4896	4717
161	11	5817	5638
163	2	1604	840

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
163	5	2796	2344
163	7	2952	2647
163	9	4905	5132
164	3	1338	1147
166	3	5213	4854
168	4	2500	2868
168	5	3595	4158
170	3	2517	2777
171	2	2277	1450
171	11	12576	11125
172	1	3	278
172	2	1940	1149
173	1	1289	708
173	5	7001	6114
174	2	593	1105
175	3	2552	2890
175	5	3820	3335
175	7	4342	4506
182	4	5477	4986
184	5	6043	5702
188	2	1210	1755
188	4	2647	2994
189	6	2614	3039
190	3	1998	2364
191	1	1	153
191	2	950	669
191	10	11786	13039

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Coding ID	ORF ID	Start (nt)	Stop (nt)
191	11	12902	12363
192	1	91	426
195	3	2106	1932
195	5	2899	2606
198	2	1016	1591
201	1	170	625
203	2	783	1466
206	6	8930	7015
206	12	13947	13616
206	21	28208	27960
212	2	170	817
212	3	796	1167
212	7	3128	3436
212	9	3749	4075
213	1	1	705
214	2	1076	570
214	6	4064	3738
214	9	6600	6995
214	10	7864	7469
217	1	1927	965
218	1	178	657
218	3	1776	2156
220	2	1851	1369
220	3	3251	2262
220	7	8275	7208
220	8	10244	8661
220	9	11796	10216

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
221	4	3095	2613
221	9	11428	10757
226	1	3	659
226	2	2196	1459
226	3	1476	1961
227	1	2	487
227	2	460	975
227	4	1855	2121
227	5	2052	2345
227	6	4760	3768
227	9	5591	6367
228	5	2503	2877
228	6	2846	3526
233	7	3944	3762
236	2	809	579
238	2	1975	1391
239	2	1417	905
241	5	4495	4334
242	2	1677	1363
243	1	127	576
244	1	1291	647
244	2	3035	1962
245	2	1614	1258
246	1	69	215
246	4	738	1733
249	3	3906	3712
250	1	494	249

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig Id	ORF ID	Start (nt)	Stop (nt)
254	1	1	156
256	2	956	1144
257	3	3700	3227
260	4	4906	4580
261	4	2196	2606
261	6	3214	3681
264	2	155	439
264	5	5252	4533
264	6	4739	5107
267	2	1323	931
268	4	5140	4700
272	1	862	446
272	3	1200	1439
272	9	4691	4909
272	10	6469	6035
276	4	1746	1901
278	3	224	553
278	5	3299	3468
278	7	4849	5127
285	2	551	736
288	3	1756	1950
288	5	2055	2276
289	1	2107	1055
290	2	2234	1932
291	2	332	622
291	5	1545	2051
295	3	1606	1349

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
295	4	2728	2141
295	5	2220	2762
297	2	788	465
298	1	2	205
300	2	2380	1928
301	7	2194	2624
304	1	3	194
306	1	109	654
306	5	4036	4257
307	1	674	339
307	8	3645	3995
308	1	1	654
308	2	1120	599
308	4	2643	2332
313	2	2314	1919
314	1	10	702
316	2	982	1341
316	6	2758	3165
317	1	2	1114
317	3	4570	3458
321	6	5645	5217
321	7	6319	6140
321	8	7450	6794
322	2	827	543
326	2	165	1112
326	3	1117	1467
328	1	936	469

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig id	ORF id	Start (nt)	Stop (nt)
328	5	3452	3276
329	1	3	719
329	2	781	1212
329	3	1471	1833
330	1	576	289
330	3	1447	1623
332	3	2353	2204
332	7	4971	5138
333	2	3295	3128
335	1	864	433
337	2	95	526
340	2	1658	1356
341	1	3	281
341	3	2476	3152
341	5	3618	3944
341	6	3929	4558
344	5	3197	2889
345	1	1532	768
346	2	221	592
350	3	1410	1558
352	2	2178	1765
352	3	7316	4596
352	7	7967	8404
352	8	8906	9247
352	9	10171	9854
359	1	1	546
362	1	3	656

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig Id	ORF Id	Start (nt)	Stop (nt)
364	2	2158	1808
364	8	10974	10714
365	2	1612	1313
365	5	4680	4090
365	7	4980	6239
366	3	520	1719
367	3	906	1085
368	1	748	494
375	1	2	136
380	3	1331	1097
389	1	1	276
390	1	2	877
390	2	1373	1549
391	2	751	560
395	1	391	197
396	1	2132	1068
398	3	1344	1141
399	1	176	669
401	3	566	847
402	2	100	465
404	8	5561	5370
408	2	3507	2269
408	3	2875	2672
408	5	3524	4423
410	3	2111	1890
413	1	880	488
416	1	607	320

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
416	2	578	847
416	3	2195	1590
417	1	3	179
417	2	161	616
420	2	788	513
422	2	357	677
431	2	856	1407
432	2	446	1084
433	1	1	417
433	3	2311	2033
434	1	942	535
434	2	2089	1235
440	1	1	450
442	2	1269	3320
443	3	1873	1370
444	1	1	656
444	7	6761	6366
451	1	940	614
453	2	896	636
453	8	3833	4786
453	9	4718	4512
453	10	4937	4731
455	1	434	219
455	2	472	930
459	1	265	687
462	1	2	247
466	2	1494	907

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
467	1	654	149
468	1	2	250
469	1	1488	925
469	3	2386	3372
469	4	3464	3706
470	1	77	538
470	6	4098	3694
470	7	6330	5686
470	9	7351	8181
470	10	8175	9773
471	1	940	500
471	2	1562	1017
476	1	70	267
477	1	2	760
477	3	1764	2081
477	4	2066	2332
480	5	4016	4261
481	2	956	480
486	3	613	774
487	6	1795	2112
488	1	715	359
492	1	127	675
493	1	2	520
493	2	496	1242
502	3	1149	1571
504	1	690	346
505	5	4566	4150

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
511	2	1741	1232
512	2	583	747
515	1	609	812
517	4	2179	2511
520	4	2097	2360
520	6	1908	3669
522	1	1	498
528	1	637	335
529	2	1679	1104
530	7	5296	5534
536	1	308	156
538	1	1362	736
538	3	2203	2880
538	5	3531	3121
538	6	4348	3731
540	1	956	664
540	2	1495	1031
541	1	89	433
541	2	719	432
542	2	1048	1272
545	2	1012	734
551	1	2145	1129
555	2	892	704
558	3	1357	1154
558	4	1760	1458
558	5	2105	1821
558	6	2166	2020

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
558	7	2636	2322
558	8	3053	2802
558	9	3986	3453
560	1	475	921
565	3	1206	1485
571	1	308	156
571	3	994	1206
577	1	2	199
577	2	163	453
579	1	1	477
579	2	1784	1200
583	1	1988	996
585	1	946	539
587	1	22	573
588	2	1896	1372
588	3	1762	1554
590	1	47	334
592	2	1455	1141
593	1	2	775
593	2	817	1122
595	1	87	890
596	3	1593	1435
602	1	8	169
603	5	1071	1469
606	1	322	768
607	5	1444	1226
610	1	1029	541

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
612	1	3	500
616	1	991	650
617	2	736	491
622	1	36	347
625	4	2046	2549
627	1	67	210
628	1	901	452
631	3	4789	4004
634	1	1448	759
636	1	189	388
636	2	1929	1063
637	2	2323	1994
638	1	227	1061
639	1	518	281
639	2	1377	811
641	1	118	444
642	3	1615	1331
642	4	2260	1847
643	1	3	608
645	4	1534	1758
645	6	2025	2321
645	7	2940	2488
648	1	2	1045
660	1	77	601
660	2	576	872
661	1	1725	961
664	2	89	304

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig no	ORF ID	Start (nt)	Stop (nt)
667	1	3	413
668	1	1	330
671	2	812	516
673	1	3	338
674	2	865	584
679	1	1	237
679	3	1589	1906
688	1	1236	835
688	2	1352	1077
694	1	3	143
696	2	818	432
706	1	367	224
709	3	1183	1449
711	1	3	908
715	1	3	167
716	1	2	637
721	1	133	570
722	1	763	383
723	1	1656	829
723	2	1498	1112
727	1	2	472
729	1	268	441
731	1	130	828
735	1	2	214
736	1	3	782
738	1	2	298
742	1	3	230

TABLE 3

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
745	3	1148	780
746	2	282	464
749	1	685	344
751	1	901	452
755	1	97	522
755	2	520	918
758	2	663	400
764	2	1033	746
767	1	1	405
768	1	2	373
771	1	1058	534
778	1	1735	902
785	1	1790	1023
787	1	1260	631
791	1	3	224
799	1	15	260
804	1	304	711
805	1	3	680
808	1	219	842
810	1	2221	1112
810	2	1774	1442
812	1	38	979
817	1	714	358
818	2	487	1104
819	2	1529	1032
819	3	1746	1419
820	1	135	1084

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
828	1	506	255
829	1	48	800
830	1	578	291
832	1	594	298
835	1	320	795
840	3	491	709
845	1	912	457
850	2	303	449
853	1	715	359
860	1	2	256
864	1	18	410
864	2	383	715
864	6	1676	1828
870	1	1	588
873	1	906	454
875	1	584	294
877	1	1661	1020
878	1	981	544
879	1	1567	785
881	1	1	243
882	1	389	604
890	1	2	508
905	1	793	398
906	1	852	544
912	1	173	188
913	1	3	290
913	2	1092	547

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
915	1	6	161
915	2	169	402
921	1	126	186
927	1	1578	808
928	1	2	185
929	1	2	400
932	1	2	400
934	1	1	184
936	1	1052	528
937	1	2	616
945	1	220	645
945	2	649	1242
946	1	1702	950
949	1	1	270
951	1	3	362
955	1	3	143
960	1	723	400
963	1	1	162
965	1	690	346
966	1	1079	606
969	1	3	302
971	1	12	170
974	1	319	161
976	1	692	348
977	1	2	211
982	1	1326	982
984	1	589	296

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	Start (nt)	Stop (nt)
987	1	467
993	1	525
994	1	549
1004	1	316
1014	1	313
1015	1	463
1016	1	145
1019	1	660
1022	1	474
1024	1	299
1024	2	431
1030	1	338
1032	1	179
1040	1	399
1043	1	269
1044	2	399
1047	1	159
1051	1	354
1051	2	733
1063	1	400
1069	1	148
1069	2	513
1075	1	399
1077	1	405
1081	1	436
1086	1	384
1087	2	431

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
1088	1	3	374
1096	1	474	238
1098	1	1015	509
1100	1	1020	511
1100	2	1520	1158
1101	1	703	353
1102	1	385	194
1107	1	2	580
1114	1	3	422
1115	1	2	268
1119	1	22	267
1129	1	40	342
1132	1	360	181
1133	1	609	376
1144	1	446	225
1147	1	558	280
1153	1	1	153
1154	1	3	818
1159	1	1	330
1161	1	341	186
1164	1	427	254
1171	1	19	240
1171	2	108	299
1183	1	2	379
1195	1	385	179
1196	1	1	189
1200	1	33	197

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Config ID	ORF ID	Start (nt)	Stop (nt)
	1203	2	129
	1204	2	129
	1222	2	105
	1223	1	1
	1240	1	2
	1247	1	520
	1271	1	412
	1286	1	2
	1295	1	1
	1306	1	367
	1314	2	158
	1316	1	58
	1359	1	384
	1370	1	1
	1371	1	1
	1374	1	710
	1378	1	2
	1392	1	3
	1411	1	202
	1433	1	331
	1450	1	2
	1453	1	295
	1471	1	721
	1477	1	869
	1502	1	794
	1518	1	126
	1534	1	283
	1546	1	3

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
1547	1	506	255
1583	1	3	350
1587	1	3	563
1602	2	170	679
1629	1	1	402
1665	1	468	235
1760	1	625	314
1762	1	3	200
1876	2	119	286
1895	1	2	379
1931	1	798	400
1976	2	715	383
2055	2	252	401
2056	1	331	167
2150	1	523	263
2157	1	794	399
2164	1	564	283
2175	1	218	400
2212	1	492	331
2338	1	732	367
2342	1	3	167
2352	1	330	166
2352	2	622	398
2355	1	47	352
2356	1	679	341
2359	1	301	152
2421	1	296	150

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig	ORF	Start	Stop
(no)	(no)	(nt)	(nt)
3046	1	367	185
3049	1	553	278
3050	1	3	314
3052	1	504	253
3065	1	2	157
3070	1	357	190
3075	1	440	222
3080	1	1	285
3092	1	320	162
3093	1	411	250
3100	1	52	237
3103	1	47	298
3118	1	344	174
3123	1	2	145
3127	1	1	147
3138	1	336	169
3142	1	388	203
3144	1	664	386
3151	1	337	170
3155	2	202	384
3168	1	12	176
3205	1	288	145
3282	1	1	150
3303	2	239	400
3371	2	211	399
3556	1	2	148
3558	2	36	401

TABLE 3

S. aureus - putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
3046	1	367	185
3049	1	553	278
3050	1	3	314
3052	1	504	253
3065	1	2	157
3070	1	357	190
3075	1	440	222
3080	1	1	285
3092	1	320	162
3093	1	411	250
3100	1	52	237
3103	1	47	298
3118	1	344	174
3123	1	2	145
3127	1	1	147
3138	1	336	169
3142	1	388	203
3144	1	664	386
3151	1	337	170
3155	2	202	384
3168	1	12	176
3205	1	288	145
3282	1	1	150
3303	2	239	400
3371	2	211	399
3558	1	2	148
3558	2	36	401

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
3568	1	751	377
3595	1	757	380
3618	1	2	238
3618	2	130	402
3622	1	86	358
3622	2	664	398
3642	1	876	439
3649	1	781	398
3651	1	625	314
3664	1	467	637
3674	1	55	402
3677	1	619	311
3704	1	1	402
3726	1	535	269
3765	1	510	256
3779	1	554	357
3794	1	266	135
3794	2	667	377
3796	2	638	375
3801	1	474	262
3806	1	453	298
3807	1	42	389
3815	1	798	400
3827	1	3	320
3842	1	781	392
3853	1	671	399
3855	1	1	324

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
3857	1	2	235
3861	1	590	297
3865	1	695	399
3897	1	3	173
3897	2	143	400
3898	2	225	401
3921	2	103	342
3927	1	70	375
3930	1	76	234
3946	2	651	382
3951	2	105	377
3965	1	646	344
3973	1	795	400
3981	1	3	311
3998	1	3	356
4001	1	481	296
4003	1	90	335
4018	1	2	259
4018	2	186	401
4021	1	1	345
4043	1	3	348
4054	1	3	344
4066	1	1	150
4070	1	1	324
4072	2	187	390
4073	1	1	285
4077	2	127	372

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
4083	1	3	359
4090	1	27	368
4101	1	103	397
4105	1	1	306
4107	1	570	786
4119	1	629	139
4121	1	740	372
4123	1	3	230
4127	1	3	341
4128	1	2	331
4130	1	768	415
4146	1	97	381
4157	1	3	206
4186	1	505	254
4224	1	510	256
4239	1	1	348
4242	1	709	356
4252	1	589	296
4253	1	1	174
4255	1	568	323
4258	2	498	334
4267	1	284	144
4271	1	2	304
4287	1	303	163
4289	1	471	319
4302	1	153	305
4304	1	1	186

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
4304	2	96	314
4306	1	2	151
4318	1	376	289
4322	1	5	148
4331	1	439	221
4331	2	528	364
4338	1	728	399
4346	1	471	277
4367	2	117	311
4373	1	2	268
4381	1	574	326
4384	1	614	309
4397	1	9	311
4402	1	1	249
4403	1	606	328
4406	1	3	317
4411	1	2	280
4411	2	497	398
4412	1	2	364
4418	1	3	230
4424	1	601	398
4443	1	427	215
4471	1	643	323
4478	1	540	271
4482	1	50	289
4489	1	601	302
4491	1	12	206

TABLE 3

S. aureus - Putative coding regions of novel proteins not similar to known proteins

Contig ID	ORF ID	Start (nt)	Stop (nt)
4495	1	3	179
4496	1	500	252
4500	1	130	306
4511	1	493	248
4518	1	1	266
4526	1	480	241
4527	1	2	163
4532	1	3	239
4542	1	11	175
4567	1	36	200
4573	1	1	231
4578	1	642	322
4619	1	1	180
4620	1	349	176
4662	1	1	246
4669	1	2	157
4680	1	28	183
4690	1	344	174

Table 4

ORF	SEQ ID NO	BLAST HOMOLOG	Antigenic Regions			
			Region 1	Region 2	Region 3	Region 4
168_6	5192	lipoprotein	36-45	84-103	152-161	176-185
238_1	5193	chrA	21-39	48-58	84-95	232-249
51_2	5194	OppB gene product (B. sub	20-36	70-79	100-112	121-131
278_3	5195	lipoprotein 1	20-29	59-73	85-97	162-171
276_2	5196	lipoprotein	21-33	65-74	177-186	211-220
45_4	5197	ProX	28-37	59-69	85-100	120-129
316_8	5198	hypothetical protein	45-54	88-97	182-192	243-253
154_15	5199	unknown	31-40	48-58	79-88	95-104
228_3	5200	unknown	25-38	40-52	64-74	80-89
228_6	5201	unknown	29-41	89-101	128-143	173-184
50_1	5202	unknown	21-33	52-61	168-182	197-206
112_7	5203	iron-binding periplasmic	21-31	58-67	92-101	111-120
442_1	5204	unknown	30-39	91-100	122-137	182-192
66_2	5205	unknown	50-59	104-116	127-136	167-182
304_2	5206	Q-binding periplasmic	19-28	48-57	75-84	103-116
44_1	5207	hypothetical protein	27-36	86-95	129-138	192-201
161_4	5208	SphX	27-44	149-161	166-175	201-210
46_5	5209	cmpC (permease)	21-33	61-70	83-92	100-109
942_1	5210	traH [Plasmid pSK41]	83-92	109-118	127-142	
5_4	5211	ORF (S. aureus)	12-22	87-96	111-120	151-160
20_4	5212	peptidoglycan hydrolase (S.	24-34	129-138	141-150	161-171
328_2	5213	lipoprotein (H. flu)	81-90	123-133	290-299	
520_2	5214	fibronectin binding protein	44-54	63-79	81-90	95-110
771_1	5215	emm1 gene product (S. pyo	30-39	65-82	96-106	112-121
999_1	5216	predicted trithorax prot. (D	7-16	120-129	157-166	
853_1	5217	ORF2136 (Marchantia polyr	43-52	88-97	102-111	
287_1	5218	psaA homolog	13-22	28-44	72-82	114-124
288_2	5219	cell wall enzyme	14-23	89-98		
596_2	5220	penicillin binding protein 2b	40-49	59-68	76-87	106-115
217_5	5221	fibronectin/fibrinogen bindi	28-37	40-49	62-71	93-111
217_6	5222	fibronectin/fibrinogen bp	10-19	31-40	54-62	73-92
528_3	5223	myosin cross reactive prote	4-13	29-47	60-73	90-99
171_11	5224	EF	20-31	91-110		
63_4	5225	penicillin binding protein 2b	12-21	59-68	95-104	
353_2	5226		46-55	62-71		
743_1	5227	29 kDa protein in fimA regi	23-32	68-79	94-103	175-184
342_4	5228	Twitching motility	10-19	48-60	83-92	111-121
69_3	5229	arabinogalactan protein	97-106	132-141	158-167	180-189
70_6	5230	nodulin	36-45	48-57	137-160	179-188
129_2	5231	glycerol diester phosphodie	8-17	41-50	55-74	97-106
58_5	5232	PBP (S. aureus)	26-35	70-79	117-126	152-161
188_3	5233	MHC class II analog (S. aure	72-81	94-103	115-124	136-145
236_6	5234	histidine kinase domain (Dic	24-33	52-67	81-94	106-121
310_8	5235	clumping factor (S. aureus)	59-71	77-86	93-102	118-127
601_1	5236	novel antigen/ORF2 (S. aur	45-54	91-104	108-117	186-195
544_3	5237	ORF YJR151c (S. cerevisae	76-90	101-111	131-140	154-164
662_1	5238	MHC class II analog (S. aure	22-32	71-80	89-98	114-122
87_7	5239	5' nucleotidase precursor ('	29-45	62-71	105-114	125-137
120_1	5240	B65G gene product (B. sub	102-111			

Table 4

	ORF	Antigenic Regions (cont)				
		Region 5	Region 6	Region 7	Region 8	Region 9
5	168_6	244-272	303-315			
	238_1	260-269	291-301	308-317		
	51_2	140-152	188-208	211-220	256-266	273-283
10	278_3	198-209				
	276_2	255-268				
	45_4	177-199	221-230	234-243	268-279	284-293
	316_8					304-313
	154_15	148-157	177-187	202-211		
15	228_3	101-119	139-154	166-181		
	228_6					
	50_1					
	112_7	136-149	197-211	218-229	253-273	
	442_1	199-210	247-257	264-277	287-309	
20	66_2					
	304_2	178-187	250-259			
	44_1					
	161_4					
	46_5	131-141	162-176	206-215	243-252	264-273
25	942_1					285-294
	5_4	189-205	230-239	246-264	301-318	340-354
	20_4	202-212	217-234	260-275	314-336	378-387
	328_2					380-391
	520_2					
30	771_1	145-154				
	999_1					
	853_1					
	287_1	154-164				
	288_2					
35	596_2	121-130				
	217_5	244-253	259-268	288-297	302-311	
	217_6	144-158	174-183	188-197	207-216	226-242
	528_3					
	171_11					
40	63_4					
	353_2					
	743_1	197-207				
	342_4					
	69_3	195-211				
45	70_6	206-215	263-272	291-301	331-340	358-371
	129_2	117-127	141-157	168-183	202-211	390-414
	58_5	184-203	260-269	275-299	330-344	372-381
	188_3					424-433
	236_6	138-147	163-172	187-198	244-261	268-278
50	310_8	131-140	144-153	177-186	190-199	204-213
	601_1	208-218				216-227
	544_3	170-179	184-193	224-235	274-287	327-336
	662_1					352-361
	87_7					
55	120_1					

Table 4

5	ORF	Antigenic		Regions		(cont)	
		Region 11	Region 12	Region 13	Region 14	Region 15	Region 16
	168_6						
	238_1						
	51_2						
10	278_3						
	276_2						
	45_4						
	316_8						
	154_15						
15	228_3						
	228_6						
	50_1						
	112_7						
	442_1						
20	66_2						
	304_2						
	44_1						
	161_4						
	46_5	306-315					
25	942_1						
	5_4	393-407	416-426	456-465			
	20_4	396-405	410-419	461-481			
	328_2						
	520_2						
30	771_1						
	999_1						
	853_1						
	287_1						
	288_2						
35	596_2						
	217_5						
	217_6						
	528_3						
	171_11						
40	63_4						
	353_2						
	743_1						
	342_4						
	69_3						
45	70_6	453-471	506-515				
	129_2	296-315					
	58_5						
	188_3						
50	236_6	358-377	410-423	428-439	442-457	467-476	480-493
	310_8	238-251	256-275	281-290	296-310	314-333	338-347
	601_1						
	544_3						
	662_1						
55	87_7						
	120_1						

EP 0 786 519 A2

Table 4

5	ORF	Antigenic Regions (cont)				
		Region 17	Region 18	Region 19	Region 20	Region 21
	168_6					
	238_1					
	51_2					
10	278_3					
	276_2					
	45_4					
	316_8					
	154_15					
15	228_3					
	228_6					
	50_1					
	112_7					
	442_1					
20	66_2					
	304_2					
	44_1					
	161_4					
	46_5					
25	942_1					
	5_4					
	20_4					
	328_2					
	520_2					
30	771_1					
	999_1					
	853_1					
	287_1					
	288_2					
35	596_2					
	217_5					
	217_6					
	528_3					
	171_11					
40	63_4					
	353_2					
	743_1					
	342_4					
	69_3					
45	70_6					
	129_2					
	58_5					
	188_3					
	236_6					
50	310_8	357-366	370-379	429-438	443-452	478-487
	601_1					
	544_3					
	662_1					
55	87_7					
	120_1					

Table 4

ORF	Antigenic Regions (cont)					
	Region 23	Region 24	Region 25	Region 26	Region 27	Region 28
168_6						
238_1						
51_2						
278_3						
276_2						
45_4						
316_8						
154_15						
228_3						
228_6						
50_1						
112_7						
442_1						
66_2						
304_2						
44_1						
161_4						
46_5						
942_1						
5_1						
20_4						
328_2						
520_2						
771_1						
999_1						
853_1						
287_1						
288_2						
596_2						
217_5						
217_6						
528_3						
171_11						
63_4						
353_2						
743_1						
342_4						
69_3						
70_6						
129_2						
58_5						
188_3						
236_6						
310_8	622-632	670-685	708-718	823-836	858-867	877-886
601_1						
544_3						
662_1						
87_7						
120_1						

Table 4

5

10

15

20

25

30

35

40

45

50

55

ORF	Antigenic Region 29	Regions Region 30	(cont)
168_6			
238_1			
51_2			
278_3			
276_2			
45_4			
316_8			
154_15			
228_3			
228_6			
50_1			
112_7			
442_1			
66_2			
304_2			
44_1			
161_4			
46_5			
942_1			
5_4			
20_4			
328_2			
520_2			
771_1			
999_1			
853_1			
287_1			
288_2			
596_2			
217_5			
217_6			
528_3			
171_11			
63_4			
353_2			
743_1			
342_4			
69_3			
70_6			
129_2			
58_5			
188_3			
236_6			
310_8			
601_1			
544_3			
662_1			
87_7			
120_1			

Table 4

ORF	BLAST		Antigenic Regions			
	HOMOLOG		Region 1	Region 2	Region 3	Region 4
46_1	5241	aldehyde dehydrogenase	8-17	36-52	83-96	112-121
63_4	5242	glycerol ester hydrolase (P.	9-26	57-73	93-107	123-133
174_6	5243	ketopantoate hydroxymeth	71-80	203-212	242-254	265-274
206_16	5244	ornithine acetyltransferase	1-10	34-43	54-63	194-210
267_1	5245	NaH-antiporter protein (E. t	120-129	332-347	398-408	
322_1	5246	acriflavin resistance protein	58-75	153-164	203-231	264-284
415_2	5247	transport ATP-binding prote	108-126	218-227	298-308	315-334
214_3	5248	2-nitropropane dioxygenase	123-136	216-233	283-292	297-306
587_3	5249	clumping factor	5-14	43-54	59-68	76-95
685_1	5250	signal peptidase	59-68	72-81	86-95	99-108
54_3	5251	fibronectin binding protein I	23-32	37-46	50-59	89-98
54_4	5252	fibronectin binding protein I	43-52	66-75	95-104	147-156
54_5	5253	fibronectin binding protein I	49-60	81-90		
54_6	5254	fibronectin binding protein I	55-71	82-97	139-158	175-186
328_1	5255	lipoprotein (H. flu)	11-20	61-70	96-105	

Table 4

ORF	Antigenic Regions (cont)					
	Region 5	Region 6	Region 7	Region 8	Region 9	Region 10
46_1	215-242	333-352	376-385	416-432	471-487	
63_4	145-154	191-202	212-223	245-265	274-283	291-300
174_6						
206_16	239-259	275-284				
267_1						
322_1	298-319	350-359				
415_2	344-353	371-380	395-404	456-465	486-495	518-527
214_3	318-337	365-375				
587_3	106-115	142-151	156-166	173-182	186-198	204-213
685_1	113-122	130-145				
54_3	128-138	185-194	217-226	251-260	268-277	295-305
54_4	175-188	191-200	203-212	220-229		
54_5						
54_6	220-230	287-304	317-326	344-353	364-373	378-387
328_1						

Table 4

ORF	Antigenic Regions			(cont)		
	Region 11	Region 12	Region 13	Region 14	Region 15	Region 17
46_1						
63_4	306-315	319-328	366-376	395-420	453-462	467-476
174_6						
206_16						
267_1						
322_1						
415_2	539-555					
214_3						
587_3	217-226	278-287	318-327	332-342	351-360	377-386
685_1						
54_3	316-325	329-345	355-372	387-396	416-425	438-448
54_4						
54_5						
54_6	396-407	427-436	514-531	541-550	569-578	612-622
328_1						

Table 4

ORF	Antigenic Regions			(cont)		
	Region 18	Region 19	Region 20	Region 21	Region 22	Region 23
46_1						
63_4	485-500	513-525				
174_6						
206_16						
267_1						
322_1						
415_2						
214_3						
587_3	396-405	426-442	459-470	485-494	505-514	531-562
685_1						
54_3	455-462	472-491	517-536			
54_4						
54_5						
54_6	639-648	673-681	703-715	723-732	749-760	772-788
328_1						

Table 4

ORF	Antigenic Regions (cont)					
	Region 24	Region 25	Region 26	Region 27	Region 28	Region 29
46_1						
63_4						
174_6						
206_16						
267_1						
322_1						
415_2						
214_3						
587_3	567-578	584-601	607-840	844-854	858-870	877-886
685_1						
54_3						
54_4						
54_5						
54_6	793-802	811-826	834-848	866-876	893-903	907-918
328_1						

Table 4

ORF	Antigenic Regions (cont)	
	Region 30	Region 31
46_1		
63_4		
174_6		
206_16		
267_1		
322_1		
415_2		
214_3		
587_3	889-911	927-936
685_1		
54_3		
54_4		
54_5		
54_6	925-944	951-997
328_1		

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Human Genome Sciences, Inc.
- (B) STREET: 9410 Key West Avenue
- (C) CITY: Rockville
- (D) STATE: Maryland
- (E) COUNTRY: US
- (F) POSTAL CODE: 20850

(ii) TITLE OF INVENTION: Staphylococcus aureus Poly-nucleotides and Sequences

(iii) NUMBER OF SEQUENCES: 5255

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
- (B) COMPUTER: HP Vectra 486/33
- (C) OPERATING SYSTEM: MSDOS version 6.2
- (D) SOFTWARE: ASCII Text

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/009,861
- (B) FILING DATE: 05-JAN-1996

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5895 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

10 TCCATTATGA AGTCACAAGT ACTATAAGCT GCGATGTTAC CAATGTTTTT TAAAATCCCA 60
 GTAATAAAAT CAAAAAATAA GTTAAATAAT GTATTCAATT TAAGTCCTCC TTAATAAAGa 120
 15 aaataGGTAA TAATGTAATA GCTTCTATTA TGATGCCTAA TTGAATGAAT TGGGCAAATG 180
 GCTCTTTGAT GATAAGTGTG ATAATGAAAA GGGTTAAACT AACATAATC GCATAATATT 240
 TTTTTCGTTT AATAAGTCGC ACAGGAATGG GCTTCTTTTT AGTTGCTGCA GGAGCATATA 300
 20 CTGAGATTAC ACCTAAAGAA ATAAGTGTTA AAATAATCAT AATTAAAAAG TTAATATGAA 360
 AATTTACTAT TACTAAAGGT AAAAGTATAA ATAGTATAAT ACTTTCTACA TAACACCAAA 420
 AAGAAGAAGG TGCATGTGCa CCATGTGCAT GtCTTCTTAT TAAATAAAAT GTTAAATTCG 480
 25 TAATTAACGT AAACAGAAAA ATGTTTAAAA TATAGGCAAT AGTATACATA ACAATTAATT 540
 TACCTATATT TTTAGCTAAG ACCTGCATCC CTAATCGTAC TTGCAAAAAT TGAATATGAT 600
 30 CTAAGTTATT TCTCTTTTGA AGATACGTGG CAAACTGGTC AATTTTATTA TCAAAATAAT 660
 TCAATTTTAC ACCACTCTCC TCACTGTCAT TATACGATTT AGTACAATCT TTTATCATTa 720
 TATTGCCTAA CTGTAGGAAA TAAATACTTA ACTGTTAAAT GTAATTTGTA TTTAATATTT 780
 35 TAACATAAAA AAATTTACAG TTAAGAATAA AAAACGACTA GTTAAGAAAA ATTGGAAAAT 840
 AAATGCTTTT AGCATGTTTT AATATAACTA GATCACAGAG ATGTGATGGA AAATAGTTGA 900
 TGAGFTGTTT AATTTTAAGA ATTTTATCT TAATTAAGGA AGGAGTGATT TCAATGGCAC 960
 40 AAGATATCAT TTCAACAATC GGTGACTTAG TAAAATGGAT TATCGACACA GTGAACAAAT 1020
 TCACTAAAAA ATAAGATGAA TAATTAATTA CTTTCATTGT AAATTTGTTA TCTTCGTATA 1080
 45 GTACTAAAAG TATGAGTTAT TAAGCCATCC CAACTTAATA ACCATGTAAA ATTAGCAAGT 1140
 GAGTAACATT TGCTAGTAGA GTTAGTTTCC TTGGACTCAG TGCTATGTAT TTTTCTTAAT 1200
 TATCATTACA GATAATTATT TCTAGCATGT AAGCTATCGT AAACAACATC GATTTATCAT 1260
 50 TATTTGATAA ATAAAATTTT TTCATAATT AATAACATCC CCAAAAATAG ATTGAAAAAA 1320
 TAACTGTAAA ACATTCCCTT AATAATAAGT ATGGTCGTGA GCCCCTCCCA AGCTCGCGGC 1380
 CTTTTTTGTA ATGAAGAAGG GATGAGTTAA TCATCATTAT GAGACCCGCC GTTAAATAT 1440

	TCATTTGCAA AGGGCGAAAT GGGTTCCTTAC TGAGTTATCT ATTATAAAAA AATAAACATA	1560
	GACTTATGAA AAATCTCTCA TAAATCTATG TTTAGTCATG aCATGTGTTA AATATTATTT	1620
5	CGGGCGCTTC TTATTTATAC AAATCTAATT TAATACTTTT AAATACAGGT ATATTTTCgC	1680
	GTTGCTGTTT TACTTCATTT AAGTTTAAAT CTACAGTCAA AATATCTGCG GATTCATTTA	1740
10	ATTCTCCAAC TAAATCTCCA TTTGGGTTTA TAACTATCGA ATGACCAGCA TATTCTGTGT	1800
	TACCATCGAA TCCAGTGCTA TTAGTTCCAA TGACAAACAT ATTATTTTCA ATTGCACGTG	1860
	CCTTTAGTAA TGAATGCCAA TGTGAAGAC GTGACATAGG CCATTGCGCC ACATAAAATG	1920
15	CAATTTTAGC ACCACTACGA GCAGGATATC TTAATAATTC TGGAAAACGT AAATCATAAC	1980
	AGATAAGTTG GGTACACATAA GTACCGTCAG ACAATTGAAA GGGTTCAGCT ACGTATTCGC	2040
	CAGCGGTTAA AAATTCATGC TCTCTTAACA TAGGAACTAA ATGAACTTTG TCGTATTCaT	2100
20	TAATCAGCTG GCCACTTTTA TTCACACTAA AAGCTGTATT AAATATTTGA TTGTTTCTAA	2160
	TGTTAGAAAC TGACCCAGCT ACGATATCGA CTTTATATTT TTCAGCTAAA TGTTTAATAA	2220
	ATGAAAAACT TTGTCTAGA TTATTATCTG CTTTTTCATT TAAATGCTCT AAATCATAGC	2280
25	CATTATTCCA CATTTTCAGGT AAAACGACTA CATCTACTTC AGCATTCTATA TTTTTTTCGA	2340
	ACCATTGCGT TATTTGAGTT TCATTTTATG AACTATCTCC AAAACAATC GGTAATTGAT	2400
30	AAATTTGGAC TTTCATAACA TCACATCCTT GATAGATCTT ATATATAACT TACTAAAAGT	2460
	TATGTTGAAA CGCAAAAAAC GAGCACAAGA CATAAAATCA AAGTCCTAGG CTCTACAAAG	2520
	TTATATTGAC AGTAGTTGAT GGGGCCCAA CATAGAGAAA TTGGAACACC AATTTCTACA	2580
35	GACAATGCAA GTTGGGGTGG GCTCTAACAT AAAGAAATAC TTTTCTTTA GAAATTAGTA	2640
	TTTCTTATAC ATGAGTTTTA CTCATGTATT CCTATTCTTA AGTGCACATT AGCAGCGGCT	2700
	AATGTGTAAG AACTACTACA TAATGAATAA CTAATGATTC TTTATCATTT CTGTCCCATT	2760
40	CCTAACATA TATTGATTAT TTTTTTATTA CGAAACGATC TTCCACTGGA TTAAATGTTT	2820
	TTTCGCCAGC AGCTTCACGA ATATCACCAA ATGGCATTG AGCAATAAGT TTCCAACTTT	2880
45	TAGGAATATT AAATTCATTT GAAGTCATCT CATCAACAAG TGGATTATAG TGTGTGAATG	2940
	AAGCACCTAT GCCTTTAGTA GCTAATGCAG TCCAAATTGC AAATTGATGC ATGGCATTG	3000
	TTTGAGTTGA CCATATTGCA AAATTATCAT AGTAGTTGG CATTGTGTCT TGTAACCAC	3060
50	TTACAACATC TTGATCTTCA TAAAACAAAA TTGTACCGTA TGAATGTTTG AAGTTATCAA	3120
	TTTTTTGCTC AGTTGGCTCG AAATCAGAT TCTCTCCCAT GACTTCTTTT AAAATTGCTT	3180
55	TTGTGTTATC CCAAAATTTA TTATTGTTGT CATTTAACAA GAGAACAATT CTAGTTGATT	3240

	CATCGCTAAT	TGATATCGAA	TCTTTCAAAT	TATATATTGA	ACGTCTTTCT	TCCATTGCAT	3360
	TGTCAAAAGT	CATTGCTTTT	TTATCTTTTT	TAAATAAGCC	CATAATTATT	GCTCCTTCTT	3420
5	TAGTAAAGAA	TACTTAATAG	ACTAAGTATA	AAATTTATAC	TCGTACTTGT	AAAGCAATAT	3480
	TTACGAAAAT	TTCAAGAATA	TTAATATTCA	TTTTCAAATT	CCAAATATAA	ATGCATTTTC	3540
10	AACGCATATT	TATTATACTT	AGATTAATAC	TTACATGAAA	AAGGGAGGTG	TCTCGTGAAA	3600
	TGTCATATCA	TTGGTTTAAG	AAAATGTTAC	TTTCAACAAG	TATTTTAATT	TTAAGTAGTA	3660
	GTAGTTTAGG	GCTTGCAACG	CACACAGTTG	AAGCAAAGGA	TAACCTAAAT	GGAGAAAAAC	3720
15	CAACTACTAA	TTTGAATCAT	AATATAACTT	CACCATCAGT	AAATAGTGAA	ATGAATAATA	3780
	ATGAGACTGG	GACACCTCAC	GAATCAAATC	AAACGGGTAA	TGAAGGAACA	GGTTCGAATA	3840
	GTCGTGATGC	TAATCCTGAT	TCGAATAATG	TGAAGCCAGA	CTCAAACAAC	CAAAACCCAA	3900
20	GTACAGATTC	AAAACCAGAC	CCAAATAACC	AAAACCTCAAG	TCCGAATCCT	AAACCAGATC	3960
	CAGATAACCC	GAAACCAAAA	CCGGATCCAA	AACCAGACCC	AGATAAACCA	AAGCCAAATC	4020
	CGGATCCAAA	ACCAGATCCA	GATAACCCGA	AACCAAATCC	AGATCCAAAA	CCAGACCCAG	4080
25	ATAAACCAAA	GCCAAATCCG	GATCCAAAAC	CAGATCCAGA	TAAACCAAAG	CCAAATCCGA	4140
	ATCCAAAACC	AGACCCTAAT	AAGCCAAATC	CTAACCCGTC	ACCAGATCCC	GATCAACCTG	4200
30	GGGATTCCAA	TCATTCTGGT	GGCTCGAAAA	ATGGGGGGAC	ATGGAACCCA	AATGCTTCAG	4260
	ATGGATCTAA	TCAAGGTCAA	TGGCAACCAA	ATGGGAATCA	AGGAAACTCA	CAAAATCCTA	4320
	CTGGTAATGA	TTTTGTATCC	CAACGATTTT	TAGCCTTGGC	AAATGGGGCT	TACAAGTATA	4380
35	ATCCGTATAT	TTTAAATCAA	ATTAATAAGT	TGGGCAAAGA	TTATGGAGAA	GTTACTGATG	4440
	AAGACATTTA	TAATATTATT	CGAAAACAAa	ATTTTCAGCG	AAATGCATAT	TTAAATGGAT	4500
	TACAACAGCA	ATCGAATTAC	TTTAGATTCC	aATATTTCAA	TCCATTGAAA	TCAGAAAGGT	4560
40	ACTATCGTAA	TTTAGATGAA	CAAGTACTCG	CATTAATTAC	TGGTGAAATT	GGATCAATGC	4620
	CAGATTTGAA	AAAGCCCGAA	GATAAGCCGG	ATTCAAACA	ACGCTCATT	GAACCGCATG	4680
45	AAAAAGACGA	TTTACAGTA	GTTAAAAAAC	AAGAAGATAA	TAAGAAAAGT	GCGTCAACTG	4740
	CATATAGTAA	AAGTTGGCTA	GCAATTGTAT	GTTCTATGAT	GGTGGTATTT	TCAATCATGC	4800
	TATTCTTATT	TGTAAAGCGA	AATAAAAAGA	AAAATAAAAA	CGAATCACAG	CGACGATAAT	4860
50	CCGTGTGTGA	TTCGTTTTTT	TTATTATGGA	ATAAAAATGT	GATATATAAA	ATTGCTTGT	4920
	TCCGTGGCTT	TTTTCAAAGC	CTCAGGATTA	AGTAATTGGA	ATATAACGAC	AAATCCGTTT	4980
55	TGTAACATAT	GGATAATAAT	TGGAACAGCA	AGCCGTTTTG	TCCAAACATA	TGCTAATGAA	5040

AATATTAATG AACTTACTGT TGTAGCAATA ATAAATGCCA CGATACGATT ACCTTTAATC 5160
 GCATTAAATA ATTCTCCAAA GATTACTTTT CTGAATACAT ATTCTTCTAA TAAAGGACCA 5220
 5 ATAATAGATA CAAAGAAGAT AAATATAGGT ATTTTTCGAG CAATAATAAT TAGCTTTTCT 5280
 GTATTAGGAC TTAAGTGTG TCCACCATAA ATTTGCGTTA ATACAATGCT CACTACCATT 5340
 10 TGATAAATCA TTACCAATGC AAATCCAAGC AATGCCCATG GAATGATATA TTTTITAGGT 5400
 TCTTTAACTT CTAATTCTAA TTTTGTGGA TTTTAAATTT TTAAATTAAT TAAATAATC 5460
 GTCGTGGCGG CGATTAAAAA TAGAACAAGT TGTATGTAAA TGACTGCTTT AGTCAGTTCT 5520
 15 ATGCCACTAT ATTGTACAAA TGGTAATTTT TTTACAATGA GAAGCGGTAA AAATTGAGAC 5580
 AATATATAAA TAATAACAGT TAGCAATGAT GCCCATAATC tTGTCATAAT TTTCCTCCAA 5640
 ATATTTGTTT ATAATTTATT TTATCGTAAA TAACTTGAAG TTACAAAAC TAATTAAAAG 5700
 20 GTTATGACTT GAAATTTTGA CCAAATTTGA TTATTATAAA TGTATGTTAG CACTCTTTAA 5760
 TGTTAAGTGC TAACTTTTAG GTTTTTTAAG GAGGAACAAT CATGCTAAAA CCAATTGGAA 5820
 ATCGTGTGAT TATTGAGAAA AAAGAACAAG AACAAACAAC TAAAAGTGn ATTGTTTAAC 5880
 25 TGATAGTGCT AAAGA 5895

(2) INFORMATION FOR SEQ ID NO: 2:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6796 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TTTGAAAAA CAAGGTACGA TTGGTTTAAT AACATATATG AGAACCGATT CTACACGTAT 60
 40 TTCaGATACT GCCAAAGTTG AAGCAAAACA GTATATAACT GATAAATACG GTGAATCTTA 120
 CACTTCTAAA CGTAAAGCAT CAGGGAAACA AGGTGACCaa GATGCCCATG AGGCTATTAG 180
 45 ACCTTCAAGT ACTATGCGTA CGCCAGATGA TATGAAGTCA TTTTGTACGA AAGACCAATA 240
 CCGATTATAC AAATTAATTT GGGAACGATT TGTGCTAGT CAAATGGCTC CAGCAATACT 300
 TGATACAGTC TCATTAGACA TAACACAAGG TGACATTAAA TTTAGAGCGA ATGGTCAAAC 360
 50 AATCAAGTTT AAAGGATTTA TGACACTTTA TGTAGAACT AAAGATGATA GTGATAGCGA 420
 AAAGGAAAT AAAGTGCCTA AATTAGAGCA AGGTGATAAA GTCACAGCAA CTCAAATTGA 480
 ACCAGCTCAA CACTATACAC AACCACCTCC AAGATATACT GAGGCGAGAT TAGTAAAAAC 540
 55

	AAAGCGTAAC TATGTCAAAT TAGAAAGTAA GCGTTTTGTT CCTACTGAGT TGGGAGAAAT	660
	AGTTCATGAA CAAGTGAAAG AATACTTCCC AGAGATTATT GATGTGGAAT TCACAGTGAA	720
5	TATGGAAACG TTACTTGATA AGATTGCAGA AGGCGACATT ACATGGAGGA AAGTAATCGA	780
	CGGTTTCTTT AGTAGCTTTA AACAAAGATGT TGAACGTGCT GAAGAAGAGA TGGAAAAGAT	840
10	TGAAATCAAA GATGAGCCAG CCGGTGAAGA CTGTGAAATT TGTGGTTCTC CTATGGTTAT	900
	AAAAATGGGA CGCTATGGTA AGTTCATGGC TTGCTCAAAC TTCCCGGATT GTCGTAATAC	960
	AAAAGCGATA GTTAAGTCTA TTGGTGTTAA ATGTCCAAAA TGTAATGATG GTGACGTCGT	1020
15	AGAAAGAAAA TCTAAAAAGA ATCGTGTCTT TTATGGATGT TCGAAATATC CTGAATGCGA	1080
	CTTTATCTCT TGGGATAAGC CGATTGGAAG AGATTGTCCA AAATGTAACC AATATCTTGT	1140
	TGAAAATAAA AAAGGCAAGA CAACACAAGT AATATGTTCA AATTGCGATT ATAAAGAGGC	1200
20	AGCGCAGAAA TAATATTTTT ATTTCCTAGA TACATTTTAA GATTGTTAAA TAGAATCATT	1260
	AGTGAATCTT ATTTTAAAGA TAGTAAAGGA TTAATCTAAA TAAGTGCGGA TAATATAAAC	1320
25	ATAACAACAT AATTAAGAAGA CATAAATGAC aATAAAAGGA GTATAGAAAT GACTCAAAT	1380
	GTAAATGTAA TAGGTGCTGG TCTTGCCGGT TCAGAAGCGG CATATCAATT AGCTGAAAGA	1440
	GGAATTAAAG TTAATCTAAT AGAGATGAGA CCTGTAAAC AACACCAGC GCACCATACT	1500
30	GATAAATTG CGGAACCTGT ATGTTCCAAT TCATTACGCG GAAATGCTTT AACTAATGGT	1560
	GTGGGTGTTT TAAAAGAAGA AATGAGAAGA TTGAATTCTA TAATTATTGA AGCGGCTGAT	1620
	AAGGCACGAG TTCCAGCTGG TGGTGCAATTA GCAGTTGATA GACACGATTT TTCAGGTTAT	1680
35	ATTACTGAAA CACTTAAAAA TCATGAAAAT ATCACAGTTA TTAATGAAGA AATTAATGCC	1740
	ATTCCAGATG GATACACAAT TATCGCAACA GGACCACTTA CTACAGAAAC CCTTGCGCAA	1800
	GAAATAGTGG ACATTACTGG TAAAGATCAA CTTTATTTCT ATGATGCGGC TGCTCCAATT	1860
40	ATTGAAAAAG AATCTATTGA TATGGATAAA GTTTACTTAA AGTCCCCTTA TGATAAAGGT	1920
	GAAGCTGCAT ATTTAAACTG TCCTATGACT GAGGATGAAT TTAATCGCTT TTATGATGCA	1980
45	GTATTAGAAG CTGAAGTTGC GCCTGTAAAT TCATTGAAA AAGAAAAATA TTTGAGGGT	2040
	TGTATGCCTT TTGAAGTAAT GGCAGAACGC GGACGCAAGA CATTACTATT TGGACCAATG	2100
	AAACCAGTAG GATTAGAAGA TCCAAAGACT GGGAAACGTC CTTATGCGGT GGTTC AATTA	2160
50	AGACAAGATG ACGCTGCTGG TACACTCTAC AATATTGTTG GCTTCCAAAC GCATTTAAAA	2220
	TGGGGAGCTC AAAAAGAAGT CATTAAATTA ATTCCAGGCT TAGAAAATGT TGATATTGTT	2280
55	AGATATGGTG TGATGCATAG AAATACCTTC ATTAATTCAC CGGACGTATT AAACGAGAAA	2340

	TATGTAGAAA GCGCagcTAG CGGCTTAGTT GCAGGTATCA ATCTTGCGCA TAAATATTA	2460
5	GGCAAGGGTG AGGTAGTATT TCCGAGAGAA ACAATGATTG GAAGTATGGC TTAGTATATT	2520
	TCTCATGCTA AAAACAATAA GAATTTCCAA CCTATGAATG CTAAC TTCG GTTATTACCA	2580
	TCTTTAGAAA CTAGAATTAA AGATAAAAAA GAACGCTATG AAGCACAAGC TAATAGAGCT	2640
10	TTGGATTACT TAGAAAATTT CAAAAAACT TTATAAAATA GTTAGAAAGA CTAGATATGC	2700
	TATTCATTCT TAAGTCATCA ACGAGTAAGT AATGACTTTC TAAATGGAAA ATACTTATCC	2760
	TAGTCTTTTT AATTTTGGAA TTGTTACGTA TTTCTGACAA TTTAGAATTC GCATTCAAAA	2820
15	AATATCTAAA TAAATAACAC GCAATAAGTT GATTGATGTA ACATGTAAGA GAATGTTTTA	2880
	AATAAACTTT ATTTAAAAGG CAATGAAATA ATAAATGGCA AGGCTATTAA TAAAGACTTT	2940
	TAGTAATTAA TTTAAAAAAG AGGTATTCTA ATTAACAGGT TTTCCGATTA GTTACAATTA	3000
20	TTTAATTCTC AAAAGATTTA GAATTGATTA TCAAATTACT GTAAGCCCTT TGCTGTATAT	3060
	GCTACAATTC TTATTGATGG AGGGTAAATG TATTGAATCA TATTCAAGAT GCGTTTTTAA	3120
25	ATACATTGAA AGTTGAACGG AATTTTTTCG AACACACATT GAAATCATAT CAAGATGACT	3180
	TAATTCAGTT TAATCAATTT TTAGAACAAG AACATTTAGA GTTGAATACT TTTGAATACA	3240
	GAGATGCTAG AAATTATTG AGCTATTTAT ATTCAAATCA TTTGAAAAGA ACATCTGTTT	3300
30	CTCGTAAAT CTCAACGTTA AGAAGTTTCT ATGAATATTG GATGACGCTT GATGAGAACA	3360
	TTATTAATCC ATTTGTTCAA TTAGTACATC CGAAAAAGA AAAATATCTT CCGCAATTCT	3420
	TTTACGAAGA AGAAATGGAA GCGTTATTCA AACTGTAGA AGAGGACACT TCAAAAAATT	3480
35	TACGGGATCG AGTTATTCTT GAATTGTTGT ATGCTACAGG CATCCGTGTT TCGGAATTAG	3540
	TAAATATTAA AAAACAAGAT ATAGATTTTT ACGCGAATGG TGTTACCGTA TTAGGAAAAG	3600
40	GGAGCAAAGA GCGCTTTGTA CCGTTTGGTG CTTATTGTAG ACAAAGCATC GAAAATTATT	3660
	TAGAACATTT CAAACCAATT CAGTCATGCA ATCATGATTT TCTTATTGTA AATATGAAGG	3720
	GTGAAGCAAT CACTGAACGC GGTGTACGAT ATGTTTTTAA TGATATTGTT AAACGAACAG	3780
45	CAGGCGTAAG TGaGATT CAT CCCCACAAGC TCAGACATAC ATTTGCAACG CATTTATTGA	3840
	ATCAAGGTGC AGACCTAAGA ACAGTACAAT CGTTATTAGG TCATGTTAAT TTGTCAACAA	3900
	CTGGTAAATA TACACACGTA TCTAACCAAC AATTAAGAAA AGTGTATCTA AATGCACATC	3960
50	CTCGAGCGAA AAAGGAGAAT GAAACATGAG TAATACAACA TTACATGCAA CAACAATTTA	4020
	TGCTGTAAGA CATAATGGGA AAGCAGCTAT GGCTGGAGAT GGGCAAGTAA CGCTTGGTCA	4080
55	ACAAGTCATC ATGAAACAAA CGGCAAGAAA AGTGCGACGT TTATATGAAG GTAAAGTGTT	4140

	ATTACAACAG TTTAGTGGTA ACTTAGAAAG AGCTGCTGTT GAATTGGCAC AAGAATGGCG	4260
5	AGGCGATAAA CAATTACGTC AATTAGAAGC TATGCTAATT GTAATGGATA AAGATGCTAT	4320
	TTTAGTTGTC AGTGGAAGTC GCGAAGTTAT TGCTCCAGAT GATGACCTTA TCGCTATTGG	4380
	ATCAGGAGGC AACTACGCAT TAAGCGCAGG ACGTGCATTG AAACGCCATG CATCGCATT	4440
10	GTCTGCTGAA GAAATGGCAT ATGAGAGCTT GAAAGTAGCG GCTGATATTT GTGTCTTTAC	4500
	CAACGATAAT ATTGTTGTCG AAACACTATA ATAATCAGAG CACGATAAAT AATTACGAGC	4560
	AATTAATTTT AGTTAAAAGA CGGAGGAATG AAATTAATGG ATACAGCTGG AATAAGATTA	4620
15	ACTCCAAAAG AAATCGTATC TAAATTAAAT GAATACATCG TTGGACAAAA TGATGCTAAA	4680
	CGTAAAGTGG CAATTGCCCT ACCTAATCGA TACAGAAGAA GTTTATTAGA TGAGGAATCA	4740
	AAGCAAGAAA TTTCACCTAA AAATATTTTG ATGATTGGAC CAACCGGCGT TGGTAAAAC	4800
20	GAAATTGCAA GAAGAATGGC CAAAGTTGTC GCGCGCCAT TTATAAAAGT AGAAGCTACT	4860
	AAATTTACTG AGGTAGGTTA TGTAGGACGA GATGTTGAAA GTATGGTTAG AGATCTTGTT	4920
25	GATGTTTCAG TAAGATTAGT CAAGGCGCAG AAAAAATCAT TGGTACAAGA TGAAGCAACA	4980
	GCTAAGGCCA ATGAAAAACT TGTTAAGTTA TTAGTTCCAA GTATGAAAAA GAAAGCGTCT	5040
	CAAACGAATA ATCCTTTAGA GTCACTTTTT GGAGGTGCAA TTCCAAATTT CGGACAAAAT	5100
30	AACGAAGATG AAGAAGAACC ACCTACTGAG GAAATTAATA CAAAACGTTT TGAAATTAAG	5160
	AGACAGCTAG AAGAAGGCAA ACTTGAAAAA GAAAAGGTAA GAATTAAAGT CGAACAAGAT	5220
	CCTGGTGCTT TAGGTATGCT AGGTACAAAT CAAAATCAGC AAATGCAAGA GATGATGAAT	5280
35	CAATTAATGC CTAAAAAGAA AGTTGAGCGA GAAGTTGCTG TTGAGACGGC AAGGAAAATC	5340
	TTAGCTGATA GTTATGCGGA TGAACAAAT GATCAAGAAA GCGCTAACCA AGAAGCGCTT	5400
40	GAATTAGCAG AACAAATGGG TATCATCTTT ATAGATGAAA TCGACAAAGT TCGACGAAT	5460
	AATCATAATA GTGGTCAAGA TGTCTCAAGA CAAGGTGTTT AAAGAGATAT TTTACCTATA	5520
	CTTGAAGGTA GCGTTATTCA AACCAAATAT GGTACTGTGA ATACTGAACA TATGCTGTTT	5580
45	ATAGGTGCTG GAGCTTTCCA TGTATCTAAG CCGAGTGACT TGATACCAGA ATTGCAAGGT	5640
	CGTTTTCCGA TTAGAGTTGA ACTTGATAGT TTATCGGTAG AAGATTTTGT AAGAATTTTG	5700
	ACAGAACCAA AATTGTCATT AATTAAACAA TATGAAGCAT TGCTTCAAAC AGAAGAAGTT	5760
50	ACTGTAAACT TTACCGATGA AGCAATTACT CGCTTAGCTG AGATTGCTTA TCAAGTAAAT	5820
	CAAGATACAG ACAACATTGG TGCACGTCGA CTTACATACAA TTTTAGAAAA GATGCTAGAA	5880
55	GATTTATCAT TCGAAGCACC AAGTATGCCG AATGCAGTTG TAGATATTAC CCCACAATAT	5940

5
 10
 15
 20
 25

AAATATACAA	AAGGAGAAAA	ATTCATGAGC	TTATTATCTA	AAACGAGAGA	GTAAACACG	6060
TTACTTCAAA	AACACAAAGG	TATTGCGGTT	GATTTTAAAG	ATGTAGCACA	AACGATTAGT	6120
AGCGTAACTG	TAACAAATGT	ATTTATTGTA	TCGCGTCGAG	GTAAAATTTT	AGGATCGAGT	6180
CTAAATGAAT	TATTAAAAAG	TCAAAGAATT	ATTCAAATGT	TGGAAGAAAG	ACATATTCCA	6240
AGTGAATATA	CAGAACGATT	AATGGAAGTT	AAACAAACAG	AATCAAATAT	TGATATCGAC	6300
AATGTATTAA	CAGTATTCCC	ACCTGAAAAC	AGAGAATTAT	TCATAGATAG	TCGTACAACT	6360
ATCTTCCCAA	TTTTAGGTGG	AGGGGAAAGA	TTAGGTACAT	TAGTACTTGG	TcNAGTACAT	6420
GATGATTTTA	ATGaAAATGA	TTTGGTACTA	GGTGAATATG	CTGCTACAGT	TATTGGTATG	6480
GAAaTCTTAC	GTGAGAAGCA	TAGTGAAGTA	GAAAnAGAAG	CGCGCGATAA	AGCTGCTATT	6540
ACAATGGCAA	TTAATTCATT	ATCTTATTCT	GAAAAAGAAG	CGATTGAACA	TATCTTTGAA	6600
GAAC TTGGCG	GTACGGAAGG	CCTATTAATC	GCATCAAAAG	TTGCAGATAG	AGTTGGTATT	6660
ACTAGATCTG	TAATTGTAAA	TGCACTACGT	AAATTAGAAA	GTGCTGGTGT	AATTGAATCA	6720
CGTTCTTTAG	GAATGAAAGG	TACTTTTATT	AAAGTTAAAA	AAGAAAAAAT	CTTAGATGAA	6780
TTAGAAAAAA	GTAAAT					6796

(2) INFORMATION FOR SEQ ID NO: 3:

30
 35

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2073 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

40
 45
 50
 55

ATCCTAAAAAT	TnAAAATTAT	CACGCCTTTT	GaACAGCTTT	GTAACCaTct	GGACGATCAT	60
kAAATTCcAa	TGTAAATCCT	GGTTTAAaGT	TGATCTTTAA	CCTTATTTAA	AyCACCAATT	120
GTACGTATAT	TATGTTGTTT	AGCAAAATCA	CGTTTTACAG	CTAAAGCATA	CGTATTGTTA	180
TACTTCATTG	GTTTTAACAT	AGTCATTTGA	TATTTCTTTT	CAAGACTTTG	CTTAGCTTGT	240
TCATAAACTT	TTTTCTCTTC	TTTTGACTTC	AATGGTTCTT	TTGTTAATTC	ACCTAAAACT	300
GTTCCAGTAA	ATTCTAAATA	CCCATCTATA	TCGTCAGATT	TTAAAGCATT	AAATAAAAAT	360
GCTGTTTTGC	CCATACCATC	TTTCACTTCT	ACAGTATTTT	TGGTCTCTTC	TTCTATTAAA	420
ATTTTATACA	TATTTGTAAT	AATCGATGGC	TCGGAGCCAA	GCTTTCCAGC	TAACGTAATT	480
TTATCACCTT	TTTGTGCAAA	CATAGGAATA	GCGATAGCCA	GTATAATAAT	CATCACTATA	540

	TCAAATATAA TTGCCAATAA GGCTGCTGGA ATTGCACCTA ATAATATCAA CGATGCATTG	660
	TTACGGTCTA TACCTAATAA AATTAAATCT CCTAGTCCGC CTGCACCAAT TAATGCTGCT	720
5	AGTGTGCTG TACCTATAAT TAATACCATA GCCGTTCTTA CACCAGCCAT TATAACAGGC	780
	ATTGCTATCG GAAGTTCGAC TTTAGTTAAA CGTCTAAATG GTTTCATACC TATACCTTTA	840
	GCCGCTTCAA TGAGTGATGG ATCAACTTCT TTAATTCCAG TATACGTATT CCTTAAAT	900
10	GGTAACAACG CATACACTAC AAGTGCAATA ATTGCTGGCA CACGACCGAT ACCAAATAAA	960
	GGAATCATTAA AACCTAATAA TGCCAACGAT GGTATGGTTT GAAGAATTGC CGCAATATTC	1020
15	ATTACGATTT CAGATATCGT TTTAGTCTTC GTTAATAAAA TACCTAATGG TACCGCAATA	1080
	GCAGTTGCAA TCAATAATGC GATAAATGAT ATTTGAATAT GTTCTATCAT TGTCGAAAAG	1140
	AGTTGCCCCCT TACGTTCACT CAATATGTCg AAAAAAGTTAG TCATGTTGAG CTACCTCCTT	1200
20	TTTCTGGGAC AAATATTTGA AGATATCTTT CCTATCAATA ACATATTGAC CTACGCTATC	1260
	TTCTTGCA TG ACAATGACAC GCTCGCTCTC TGATAAAAGT TGATACAATA CTTCAATTGG	1320
	TTGATTGTCA TAAACAATTG GATAAGCGCT CATAGATGTA ACCTCATCGA TTGGTTTCAT	1380
25	AATATCCAAG TCACGGATAA TTGCGTTCTC TTCAACACAT GGCGCATCAT CTTCTAAATG	1440
	ACTACCCATA AATTGTTTAA CAAATTCAC T TGAGGATTA TTTTAAATC CTTCTGGTGT	1500
30	GTCAATTTGT TCAATATGCC CTTCAATCAA AAGACAAATC TTATCACCAA GTTTCATCGC	1560
	CTCTTGAATA TCATGTGTAA CAAATATGAT TGTCTTCTTA ATTTTAGTTT GTAATTCAAT	1620
	TAAATCATCT TGAAGTTTTT CTCGGCTGAT TGGGTCTAAT GCACTAAACG GTTCATCCAT	1680
35	TAAAATAACT GGTGGATCAG CTGCTAACGC ACGTATAACT CCTACACGTT GTCGTTGCCC	1740
	CCCTGACAAT TCATCAGGTT TTCTGTTTTT ATATTTTCA GGTCTAATC CAACCATTTT	1800
	AAGTAATTCA TCTACTCTTT TATCTATATC TTTTCTTTC CACTTTTTC TTTGTGGCAC	1860
40	TTGTGCAATa TTTTCTTTGa WTGTCaTATG TGGGAATAAT GCAATCTGCT GcAATACGTA	1920
	TCCAATATCC CAACKCATTT CGTATACTGG ATAATCACTT ATTGGTTTAT CTTTAAATA	1980
45	AATATAACCT TCACTTAAGT GAATGAGTCG ATTAATCATT TTTAATGTCG TAGTTTTTCC	2040
	ACAACCTGAA GGTCCAATTA GCACAAAAAA TTC	2073

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13321 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

	ACTATTCTAG CTTCATCAGT TATCATATAT TCTTTGAAAC ACTTGTAAGA AAATATAATG	60
5	AGTATTTACT ACATAATGAT ATTTCAAATT AGAAAAAGG AAGTTATGAT TTAATGGCCT	120
	TGAGCCTATC ATAACCTCCT TTTATCATTT TATTGTTGTG TTGATGTTTC GATAACGTGG	180
10	TACATCTTAT CAAACATCAA TTCGAAACCA TGCACCATGG CATCATGATA TTCTTTTTTC	240
	TTTTGCTTGT ATTCTAAATT AGTAAATCGT CTTTCTTTTT CAACTAATGA ACGATAATAA	300
	AATAGCATTG GGGTGCCACC TGTTTCACGT TCAAAAAATT CTACCTCAAT GACATCTTGC	360
15	GTTTCACTTA GTCCAGGCAT ACCGATAGTC ATCTTAACGT ATTCATCCAT AACTAAAGAT	420
	TCATAAATGC CTTCAATCAC ATTTACTTTG CCATTACGTT GTTGATCTAC AATACGATAT	480
	TTACCGCCTT CTTTAACGTC CGCTTCAATC TCTTTATTCG TTCTGGCTGA TGTCATAAAC	540
20	CATTGTTTCA ACAAATCTTT CTTTGTCCAA GCTTCGTATA CTAACCTCTGG AGAAAATTTA	600
	TAAAGCTTTT CAATTTCAAC TTCGACATGT TCATTCTCTA CATTAAATTT TGCCACTGTT	660
	GTCCACCCAC TTTGCTCTT ACTTTTATTT TAACGTATTT TTGCTCAGTT CCAAACATAG	720
25	ATGATCATCA TTTTAAAAG ATTAGCGTTA TACGGTGAGT ACAACATGAT CTGTTAATAT	780
	AACAAGCCAC CTTACTTGGC TACATCGATA TATTGTTAAG CATTAAATGTT TCATTTCTTG	840
30	ACTAGTGTTT TTTTTTAGCT TTGGAAAATT AAATAAAATC GCAATAAGTC CGCATAACCC	900
	TAATAATATA GGATAAATGC TGTATGGGAA TAACATTAAAC GGTGAAATAC CAGCTACACC	960
	AGCCGCTGaa ATGACTTGCG GGCTATATGG TAATAAACCT TGGAAGCAGC CTCCAAATAT	1020
35	ATCAAGAATA CTTGCTGATT TCCTTGAATC TACATCATAT TCATCTGCAA TATTTTTAGC	1080
	TAAAGGACCT GACATAATAA TAGAGATGGT GTTGTGTTGCC GTGGCAATAT CTGCGACACT	1140
	TACCAAACTA GCAATTCCTA ATTCTGCGCC ACGCTTTGAT TTCACTTTAG AGCGAACAAA	1200
40	TTGCAACAAC CATTCAATAC CACCATTGTG TTGAATAATA CCGACTAAAC CACCAATTAG	1260
	CAACGCAATC ATAGCAATAT CTTCCATGCT TATAATACCT TTGGACACTG CATCTAGTAG	1320
45	CCCCATCCAA CCGAATGAAC CATCTATGAG ACCAATGATT CCGGCTAATA ATGTTCCGCC	1380
	AATCAATACG ATAATGACAT TTACACCTAA TAATGCTAAT ACCAATACTA AGATATACGG	1440
	TACAACTTTA ATTAGATTAT AATCATAGTt TTTAGCATGA TTAAAGAAA TGCCATTTCGT	1500
50	TAAGAAATAC AGAATAATAA TCGTTAAAAT AGCACCTGGC AATACAATTT TAAAGTTTAC	1560
	TCTGAATTTA TCTTTCATT TCGTATGTTG TGTTCTAACC GCAGCAATTG TTGTATCTGA	1620
55	AATCATTGAT AGATTATCGC CGAACATTGC ACCTCCAACA ACTGTAGCCa tTGctAGCGC	1680

	TCCTACAGAC	GTCCCCATAG	ATATAGAAAC	AAACATACAA	ATCACAAACA	ATCCTACAAT	1800
	AATTAAATTT	TCTGGGATTA	ATGATAGTCC	TAAATTAACT	GTCGACTTTA	CGCCACCCAT	1860
5	TTTTTCAGCT	GTATTTGAAA	ATGCACCTGC	TAAAATAAAA	ATCAACATCA	TTAAAACAAT	1920
	GTTTGAATGG	CCTGCACCTT	TCGTGAAGAC	CTCAACTTTT	TTAGCAAATG	ATTCTTTTCG	1980
10	ATTCATTAAT	AACGCCACAA	TTACCGTTAT	CGTAATTGCA	ACATTTAATG	GCATTGAAGT	2040
	AAAATCACCT	GTGATAATAC	CTACGCCTAA	AAACAACGCC	ACAAATAATA	ACAAGGGGAA	2100
	TAATGCCCAA	GCATTGCTCT	TTTTATGTAC	TTCCATCCTT	TTACCTGCT	TTCCAATTAA	2160
15	AAATACCTCT	TTCTCACAAA	CGATGAAGAA	AGAGGTTTTTC	ATGTGCTTTA	CCTGCTTATC	2220
	TTCAAACCAT	TACGGTTACT	GGAATTGGCA	CATTTCGAGAT	GTTGCCGAGG	CTTCATAGGG	2280
	CCAGTCCCTC	CACCTCTCTA	GATAAGTGAT	GCTTATTTTAC	GTTTACGTTA	CAAGATAATC	2340
20	CTTAGTACGT	CAATCATAAA	TTAATCAGGA	GTCGTATAAT	ATTTTTCATA	AACAATCATT	2400
	GCTACTGTAA	TAATAATCAA	AACAATAATG	CTAATAACAA	GTAAAAGCCA	CCATTTAAGC	2460
25	ATTAATGCAA	TAAAAATGAA	CACGATAGAC	ACACTTACTA	ATATTAATGA	TATGACTTTA	2520
	AATTGCTGAA	CACGTTGCTT	GGAGATGACT	TTCAACTGTT	TGTTTGATAG	ACGCGTATTT	2580
	TTTATACTGA	TTCCCAGTAT	ATTTTCTAAT	ATTTGAACCA	ATACGATACT	TATTGCAAAAT	2640
30	ATAATAATTG	GTAAAACATC	ATAGCTCCCT	ATAGTTAATG	TATAAAATTAC	AAATCCAATG	2700
	TAAAGTAACC	CTGAGACAAA	GGATAAAAAG	TATGCGACGT	ATTGTGTTAA	CTTAATGATA	2760
	TGCTTTTTTAA	CGTTTTTGATG	TGTAAACCAT	ACATTCGAAA	CGATCGCAAC	TGCTACAAAT	2820
35	AATGTGAATA	CTATATATAA	TGGTAATTTT	TGTTTCAGGAA	AAACAGTCGC	TATTCCAAAA	2880
	GCTAATGCTA	AAATCAAAAA	TAATATAGCT	CTAGATACTA	TTAATGCCAT	AATAACAACC	2940
	CCTTTGTTTA	ATATCGAGTT	TGCAAATTTA	CGTTTATCAG	CGTTTCTATG	ATCAGTACTT	3000
40	CTACGGGTAG	CGTTTCTATG	TAATTTACAT	CATCTTAACA	TATAAATACT	TCGCTATTTA	3060
	ATTGAAAACA	TATCCTATTA	TTCTTTGTCC	GTTCTGACGT	TTAATATCTA	GCCTTAGGCA	3120
45	TTTCACTTGT	TAATGAATTT	AACTTTCTTC	CACTAACCGT	CCCTAAACCC	AATCCCGCAA	3180
	CAGTTTTTAA	CTTTTTCGTT	GTTGTCCTGA	CATCCTCATT	AAGAAAGTTT	ATTCTGCTTA	3240
	AAACTTATAA	TCCACACCCT	GAGCAAACGC	TCCTTATGAC	AGAGTATTAA	AATAAGCCGA	3300
50	TAAAGATACA	CACCTTTACC	GACTATTTAA	AATACACTTC	ACCAATTCAT	TTTAATTTAA	3360
	TGGATTGAAG	TAATAAAATT	AATATTATGT	TGTTCAATTA	AAAGCTTCAT	ACAAACCTAA	3420
55	TCTATTTGCA	CTCCACCGCT	AACACCGAAC	ACTTGTCCGG	TTGTATAACT	TGATTCTTCT	3480

EP 0 786 519 A2

	GT TTTT T T GAC CAAATG TTGG GATTTT ACTT TGAGGT T GTC CACCAGAAAT TTGTAATGGT	3600
	GACCAGAATG GACCAGGCGC TACACAGTTC ACTCTAATTC CTTT TGGTCC TAATTCTTCT	3660
5	GAAAAACTTT TAGTTAATGA AATAATTGCT GCTTTTGAAG CGGCATAATC ATGAAGAATA	3720
	GGACTAGGAT TATAACCTTG TACAGATGAT GTCGTTGTAA TTGACGCACC CGGTTTTTAAA	3780
	TATTCCAATG CTTTTTGAAC TGTCCAAAAT AGCGGATAGA CATTGTTTTT AAATGTTTTCT	3840
10	GTAAATGCCT CAGTTGTAAA TCCATGAATA TCATCATGAT ACTGTTGATG TCCAGCAACT	3900
	AAAGTAACAT TATCTAAGCC ACCTAATTGT TGATATGCTT GTTCAACAAG GTCATAGTTG	3960
15	AACTGTTTCAT CTCTTATATC ACCAGGAATT AACACTGCCT TTTGACCACT TTCTTCAATC	4020
	ACTTGGCGTA CTCTTGTGTC ATCTTGTTC TCACTCGGAA GATAGTTAAT CGCTACATCT	4080
	GCACCTTCTT TAGCATACGC AATTGCTGCT GCACGCCCTA TTGCTGAGTC ACCACCTGTG	4140
20	ACTAATATTT TATAGCCTTG TAAGCGTTGA TGACCTTGGT AAGACGTTTC GCCACAATCG	4200
	GGTGCTGGCG TCATTTT CAGA TTGTAAACCC GGTACCTCTT GTTCTTGTTC TTCATAATCC	4260
	GTTGTTTTTAA ATTTTGTTC AGGATCTTGA GCTGCCATTT TTTTACATCT CCTTATTCGC	4320
25	TTAATGGTTA TTATTTACCC AATCTTCCTA GGAACCTAAT CATGATTACA CTAAAAATTA	4380
	CTTTCTTCTT TATAAAAACA AGCTCGAATT ATTATGCAA TAGTCTCTTT ACAAATTCAA	4440
30	CAAAATACTC AGGTACTTTT TCCAGAATCC TTTTATCCGG TTTATATTGA GGATGATGTA	4500
	AATCATATTC ACTATGAGAA CCAATTAACG CAAATACACT TGGAAAATGT TGACTATAAC	4560
	CTGAAAAATC TTCTCCAATC GTAAGCGGCT GTTCCATCAT TCCCACCTTA TATCCAACAT	4620
35	GTTGGGCTAC TGCAATTGCT TTATGCGTCA ATGCCTCATC ATTCATCACA GCGCCAGGTA	4680
	AATGCGTATA ATTTAAATTA ATTTTCATAT TATATGCTTG AGCCAATCCG TCCGCAATAT	4740
	CTTGJAATCG TGTTTCTACA AGCTTTTCGTA CCACAGGATC AAAACTACGC ACTGTGCCTT	4800
40	GTACATACGC ATGATCAGCA ATGACATTCC AAGTATTACC ACATGATATT TGTCCAATTG	4860
	TTACTACCGC TTCATCAAAC GCAGATAGAT TTCTACTAAC TATGGATTGA ATACTATTAA	4920
	TCAATTGCGC CAACACAATA ACTGGATCGT TGCATTGTTT TGGCTTTGCA GCATGACCAC	4980
45	CCACGCCTTT AATATGAAAC TCAAAACGAT CTA CTGCTGA TGTAAATGCC CCTGTTTTGA	5040
	TTGCAAATGT ACCTACCGAA CGCGATGGGT CATTATGAAA ACCCAATACT GCTTGTTACAT	5100
50	CTTTTAATGC ATGTGTTTCA ATAATTTTAA AAGCGCCATG TCCTAGTTCT TCTGCTGATT	5160
	GAAAAATGAA TTAAACACGC CCAGTAAGAG TGCCCTCAAT TTCTTTTAAAT TTTACAGCTG	5220
	TAGCCAAAAT ACTAGCCATG TGAATATCAT GACCACACGC ATGCATAACA CCTTCATTTT	5280

55

	CAGCTATACA	ACTCAGACCT	TGTCCCACTT	CAGCAACAAG	CCCAGTCGCA	AGTGGTAAGT	5400
	CTAATATTCT	AATATGATGT	TCTGTTAAAA	TATCTTTAAT	TTTTTGTGTA	GTCTTAAATT	5460
5	CTTTATCGGA	TAGTTCTGGA	AATTGATGAA	AATACCTTCT	CCAGGTAACA	GCITGATCTT	5520
	TTAATCCCAT	CGGTCATTCC	CCTTCCTTAA	GTCAATGATA	TGTTGTCTAC	CCTACGATGA	5580
10	TCATCTTTGA	CTATTAAACG	ATGATTTTAC	AACAATGTAC	TCTTGTTAAT	TGCTTTTCGTT	5640
	AATGATAGAC	AGTTGTTTAA	TAATATCGTA	ACACTGTTGT	CAAACCTATTC	TAACCTTTTAT	5700
	AATTGAGACT	CTATACAAAA	ACGTGTTCTC	GAATATACTT	GTTTTTACAA	ACCACAAAAA	5760
15	GCTCTAAACA	TTAGTTTAAA	CCAATGCTTA	GAGCTTTCTA	ATTATTTTAT	GCITTTAAAAG	5820
	ATACTGTGTT	ATCTACGATG	ACCTTACCGT	CTTTAATAAC	TTTTTCTGCG	TGATTGATAC	5880
	CAAAATGATA	TGGAATATAT	TCATGATTTG	GTGCATCCCA	AATTACTAAA	TTAGCCTTAT	5940
20	CACCTGTGTT	AATTGTACCC	GCGTTAATGT	CTATTGCTTT	AGCAGCATTG	ACCGTAACAG	6000
	CATTCCAAAC	TTCATTAGGT	GATAGCTTTA	ATTTCAAGGC	TGCAATCGCC	ATAACAAGTT	6060
25	GTAAGTTGTT	TGTGACACTA	CTACCAGGGT	TATAATCAGT	TGCTAATGCA	ATCGCACCGT	6120
	TATTGTCAAG	CATGCCTCTT	GCATCTGCAT	AATCTTCTTT	ACCTAAATAG	AACGTCGTTG	6180
	CAGGTAAGAG	GACAGCTACA	GTATCACTAT	TTCGCAACTT	TTCTTTTCCT	TTATCACTAG	6240
30	AAGCTACTAA	GTGGTCTGCT	GATATTGCTT	GTTTCATCAAT	TGCTAATTCC	AGTCCGCCTA	6300
	ACGGATCAAT	TTCATCCGCA	TGTATTTTCA	CTTTAAAACC	TGCTTCTTTG	GCTTTTTCGA	6360
	TATAATGTTG	CGATTGTTCT	ATTGTAAATA	CACCTGTTTC	ACAGAAAATA	TCCGCAAAGT	6420
35	CTGCATATTG	TTTACTTCC	GGAAGTAACG	CAATCATTTC	TTCTAAAAAT	GCCTCATTTC	6480
	AACTTGCCTC	TTAGGTACA	GCATGAGGCC	CTAGGAAAGT	ATGTTTCATG	TCTAAATCAT	6540
	ATTTCTCAGC	TAAACGATTA	GACACTTTCA	ATTGCTTCAG	TTCATTTTCT	CTATCTAATC	6600
40	CATAACCACT	CTTACTTTCA	ACTGCAAGCA	CGCCGTGTTT	AATCATAGTA	AGCAAATCAT	6660
	GCTCTGCTTT	TTTAAACAAG	TCATCTTCGG	ATGTTTCTCT	AGTAGCATT	ACGGTAGATA	6720
45	ATATGCCACC	ACCCATTTCT	AATATTTCAA	GGTAAGACTT	ACCTTGACGT	TTTAATGACA	6780
	TCTCATGTTT	TCGAGATCCA	CCAAATGTTA	AATGGGTATG	TGCATCTACT	AATGCTGGGG	6840
	AACTACCTT	CCCACTAGCA	TCAATCGTCT	CAGTCGCATC	GTAGTCATCT	GTATGTGTTT	6900
50	CAGCATATAC	AATTTTGCCA	TCTTTAATGA	CAACTGTACC	ATTTTTCACA	ACATTTAATT	6960
	CATCTAATTC	CTTACCCTTC	AAAGGTTTAT	CTGTTGATCT	CGGTAAAATT	AATCTTGCTA	7020
55	TATGATTAAT	TATTAAATCA	TTCATTACTT	ATCACCTGCT	TTATCAATCA	TTGGAATATG	7080

	AACACCCATA	CCTGGGTCAG	TCGTCAATAC	ACGTTCCAAT	CTTCTTTCAG	CACGCTCTGA	7200
	TCCATCTGCT	ACAACAACCA	TACCCGCATG	AAGTGAATAT	CCCATGCCAA	CACCGCCACC	7260
5	GTGATGGAAT	GAAATCCATG	AACCACCTGC	AGCTGTGTTA	ATGAGTGCAT	TCAATACAGC	7320
	CCAATCACCA	ACCGCGTCAC	TACCATCTTT	CATACTTTCT	GTTTCACGGT	TAGGACTAGC	7380
10	AACTGAACCA	GCATCTAAAT	GGTCTCGTCC	AATAACAATT	GGTGCTGAAA	TTTCACCGTC	7440
	ACGTACAAGA	CGATTTAAAG	CTAAGCCCAT	TTTCGCTCTT	TCTCCATAGC	CTAACCAAGC	7500
	AATACGTGAT	GGTAGTCCTT	GATATGAAAT	TTTTTCTTCA	GCTAAATCAA	GCCATCTTAA	7560
15	TAACTTTTCA	TTTTCTGGGA	AAAGTTTGGC	CATTTCTTCA	TCCGCACGCT	CGATATCTTT	7620
	TGGATCACCA	CTCAACGCAG	CAAAGCGGAA	TGGCCCTTTA	CCTTCACAGA	ATAATGGTCT	7680
	AATGTAAGCT	GGTACAAAGC	CTGGGAAGTC	AAAAGCATTT	TTCACTCCGT	TATTGAAGGC	7740
20	TACTTGACGA	ATATTGTTAC	CATAATCAAA	TGCTACAGCG	CCACGTTTTT	GGAATTCAAG	7800
	CATTAATTCA	ACATGCTTTG	CCATTGAAGC	TTGTGACAGT	TCAACATATT	TTTTCGGATC	7860
	TTTTTCACGC	AATACTTTTCG	CTTCTTCTAC	AGAGTATCCT	TGTGGCACAT	ATCCATTTAG	7920
25	CGGATCATGT	GCACTTGTTT	GGTCAGTAAT	AATGTCAATT	TTAAATCCTT	TTTCTAGAAT	7980
	CGCTTGATGG	ATGTCTACAG	CATTTCCAAC	TAACCCGATT	GATAATCCTT	CTCCACGTTT	8040
30	TTTCGCCTCT	TCTGCTAATT	TTAATGCTTC	ATCTAAATCA	GCTGTTTTAA	CATCACAGTA	8100
	TTTCGTATCA	ATTGCTTAT	CAACACGTGT	TTTATCAACA	TCCACGCAA	TTGCTACCCC	8160
	ATGATTGATA	GTAATTGCTA	ACGGTTGCGC	ACCACCCATA	CCACCTAAAC	CTGCTGTCAG	8220
35	TGTAACAGTG	CCTGCTAAAT	CTCCATTAAA	GTGTTGATTA	CCTAGCTCGG	CAAATGTCTC	8280
	ATAAGTACCT	TGCACAATAC	CTTGAGAACC	AATATATATC	CAACTACCGG	CTGTGATCTG	8340
	TCCATACATG	ATTAAACCTT	TTTTATCTAA	TTTATTAAAA	TGATCCAGT	TTGCCCATTC	8400
40	AGGCACTAAT	ACTGAATTTG	AAATTAATAC	ACGTGGCGCT	TCTTCATGTG	TTTTAAATAC	8460
	AGCAACTGGC	TTTCCTGATT	GTAATAACAT	TGTCTCATCT	GATTCTAATT	CTCGTAACGT	8520
45	TTTCTCTATT	GCTTCAAAAG	CTTCCCAATT	ACGTGCTGCT	TTTCCAATAC	CACCATAAAC	8580
	AACTAAATCT	TCTGGTCTTT	CAGCAACTTC	TGGGTCTAAA	TTGTTGTATA	ACATTCTAAG	8640
	TACTGCTTCT	TGTTCCCAAC	CTTTACACTC	AATACTCAAA	CCTTTTTTTG	CTTGAATTTT	8700
50	TCTCATAAAA	TTGCTCCTG	TTCTTTTAAG	AAGTTAATTC	CACTAAATTT	AAAACGCTTA	8760
	CATTATTATC	TTCAATATTC	ATTATAGTAT	GTTAAAATAT	AGCCAACAAA	TATAAATAAA	8820
55	CTAATTATCC	ATAGCTTGAA	TCTATAAATA	AAAGGAGCAA	AACACATGAA	AATTATTTCAG	8880

	CATATTAGCC AGCCATCTTT AACTGCTACG ATTAAAAAAA TGGAAGCAGA TTTAGGTTAT	9000
5	GACTTATTTA CACGTTCAAC AAAAGACATC AAGATTACCG AAAAAGGAAT ACAGTTTTAT	9060
	CGTTATGCGA GCGAATTAGT TCAACAATAT CGATCCACGA TGGAAAAAAT GTATGATTTA	9120
	AGCGTTACAT CAGAACCAAG GATAAAAAAT GGGACTCTTG AATCTACGAA TCAATGGATT	9180
10	GCGAATTTAA TTCGAAAGCA CCATTCCGAC TACCCTGAAC AGCAATATCG TTTATATGAA	9240
	ATACATGATA AACATCAATC TATAGAGCAA TTAAGTGAAT TTAATATTCA TTTAGCTATA	9300
	ACAAATGAAA AAATAACCCA CGAAGATATA AGATCCATTC CTTTATATGA GGAATCTTAC	9360
15	ATTTTATTAG CACCCAAGGA AACATTTAAA AATCAAAATT GGGTAGATGT TGAAAAATTG	9420
	CCACTCATAT TACCAAACAA AAATTCTCAA GTGCGCAAAC ACTTAGATGA CTATTTTAAT	9480
	AGAAGAAATA TTCGTCCAAA TGTCGTTGTA GAAACAGATC GATTGGAATC AGCAGTTGGA	9540
20	TTTGTTTCATC TCGGCTTAGG TTACGCTATC ATTCCGAGAT TTTATTACCA ATCATTTCAC	9600
	ACGTCTAATT TAGAATATAA AAAAATTCGT CCAAACCTAG GCCGAAAAAT TTATATCAAT	9660
25	TACCATAAAA AACGCAAACA CTCCGAACAA GTACATACAT TCGTACAACA ATGCCAAGAT	9720
	TATTTATATG GACTTTTAGA GGCTCTTTAA CTTAAGTTAT TAGAGCCTCT TATGCAGTTG	9780
	CTCAGTCAAC TGTATACCTT TTGCCTTTAA CTTAAGTTAT TAGAGCCTCT TATGCAGTTG	9840
30	CTCAGTCAAC TGTATACCTT TTGCCTTTAA CTTAAGTTAT TAGAGCCTCT TATGCAGTTG	9900
	CTCAGTCAAC TGTATACCTT TTTCCTTTAA CTTAAGTTAT TAGAGCCTCT TATGCAGTTG	9960
	CTCAGTCAAC TGTATACCTT TTGCCTTTAA CTTAAGTTAT TAGTGCCTCT TATGTAGTTG	10020
35	CGTAGTCAaC TGTaTACCTT TTGCCTTTAA CTTAAGTTAT TAGAGCCTCT TATGCAGTTG	10080
	CGCAGATCAT CGTATAAAAA TTAATGACGT CATTTCAAAA ATCGATACAA AAATAATTTA	10140
	TTATAAAAAAT TCTAAGAAAG AAGTGAAGCA GATGTTAAAA TCTATTAAATC ATATATGCTT	10200
40	TTCAATCAGA AATTTAAACG ATTCAATACA TTTTATAGTA GATATTTTAC TTGGGAAATT	10260
	GCTATTGACT GGTAAAAAAA CTGCTTATTT TGAGCTTGCA GGCCTATGGA TTGCTTTAAA	10320
45	TGAAGAAAAA GATATACCAC GTAATGAAAT TCACCTTTTCA TATACACATA TAGCTTTCAC	10380
	TATAGATGAC AGCGAATTTA AATATTGGCA TCAGAGTTA AAAGATAATA ACGTGAATAT	10440
	TTTAGAAGGA AGAGTTAGAG ATATTAGAGA TAGACAATCA ATTTACTTTA CCGACCCTGA	10500
50	TGGTCATAAG CTAGAATTAC ATACTGGCAC ACTTGAGAAC AGATTAAATT ATTATAAAGA	10560
	GGCTAAACCA CATATGACAT TTTACAAATA AGGTGTCATT ATAAAAAGGC CTCTTGAAC	10620
55	CCGTTAAAT TTTAATTAAT TATTATATAA TAAGAGAACT TTTCAAACAA TACAGTTGTT	10680

	TTACTGCAAT TATTTTTC	AAATATCAAC GTTAATATAA	CTTCTATTAA GAAATACTCA	10800
	CATTCTGCCC TGCAATGCAA	ATCTCGTCAC ATATAAATAT	TTTTAATTAT TTTAAAAAAT	10860
5	GATGCACTAA ATTAGCAACG	AGCTTAGCAG TTCTATTGTC	AGCGTCATAT GTTGGATTCA	10920
	TCTCAGCAAT ACTAACTGAA	GACACCTTAT CACTTGGAAT	AATACGTTTT GCTAATTCAA	10980
10	GAACAGTATG TGGATACAAA	CCTAACACTG CCGGCGCACT	TACCCCAGGC GCAAACGCAC	11040
	TATCAATGAC ATCCATACAA	ATCGTAAACA TAATGACATC	ATGTTTCATGT ACAAACGTT	11100
	CAATCATATC TTTAATTGTT	GGTGATACGT GACTCAATAA	TTCATCTGCA AAGACATAAT	11160
15	CAATCTTTTT CTCTTTAGCA	TAATCAAATA AACTTTGCGT	ATTACCACCT TGAGCAATAC	11220
	CAAGCACTAA ATAATCTGTG	TTTTCATCTT CTTCTAAAAT	TGTCTAAAG CTCGTTCCAG	11280
	ATGTAGATTG TTGTTGACGA	CGTGTATCAA AATGCGCATC	AATATTTATC ACACCAATAG	11340
20	ATTGTGTTGG ATAGACTTTA	CGTGTGCTA AATATTGAGC	ATACGCAATA TCATGTCCAC	11400
	CACCTAATAA AAATGTTTGT	CTATGATTAG CAATTGACTT	CGCTGCAAGC ATAGCAAATT	11460
	CTTTTTGAGT ATCAATTAAT	TCCTCATGAT CATGATAAAC	ATTTCCGTAA TCGACTAAAG	11520
25	TTACATTGA TTCAAATCCG	GCAAACCTGC AAATGCTTGT	TTAATCGCAT CTGGTCCTTC	11580
	TTTTGCACCA ATGCGCCCTT	TGTTTAAAGC AACACCTTTG	TCAACAGCAT AGCCTAATAT	11640
30	ACCGACCCCT GATGGCATA	C TACTCTTTTC CAGCTTAGAC	AAATCTTCAA ATGTTACTGT	11700
	TTGAAATGT CTAAATTTTT	TCGGGTCTGT TTCACTATCT	AACCTTCCAG TCCATAAATT	11760
	TGGTTACCT TGCTTGTA	CAGCATTTC CCCTCTTATT	TATGTGGCTT ATTAACAATT	11820
35	AAAGTATAAC GTATAGGAAA	TTTTGAATTC AATTCATAGT	TAAATCCGTA TCTTAAAAAT	11880
	ACTTATCTAC ATTACTTTTA	CCCCTATTTT CTATGTAATA	ACGAATACTT AGCTGATTTA	11940
	TGTTAATAAA ATACGTCAAG	ACTATTACAT TTTCAATTAAT	ATTGACATAG ACAATTTATC	12000
40	TCTCGGCTTG TAATATGTAT	AATTGTTACT AAAAGATATT	TTGCTTGTTA CCTAATGGAG	12060
	GTTACATATA ATGAAGAACA	ATAAAATTTT TGGTTTTCAA	TGGGCAATGA CGATTTTCGT	12120
45	CTTCTTTGTC ATTACAATGG	CGTTATCCAT TATGCTCAGA	GATTTCCAGT CTATAATTGG	12180
	TGTCAAACAC TTTATATTTG	AAGTTACAGA TCTAGCACCA	TTAATTGCTG CAATCATTTG	12240
	TATACTCGTT TTCAAATATA	AAAAGGTCCA ACTTGCAGGT	TTAAAATTCT CAATCAGCCT	12300
50	GAAAGTAATT GAACGTCTAT	TGCTAGCTTT AATTTTACCT	TTAATTATTC TAATTATTGG	12360
	TATGTACAGC TTTAATACAT	TTGCAGATAG CTTTATTTTA	TTACAATCAA CAGGCTTATC	12420
55	AGTACCTATT ACACACATTC	TGATTGGACA TATTCTGATG	GCGTTCGTAG TAGAATTCGG	12480

5 TGTGTTGGT TTGATGTATT CAGTTTTCTC AGCAAATACA ACTTATGGTA CAGAATTTGC 12600
 TGCTTATAAC TTCCTTTATA CATTCTCATT CTCTATGATT CTTGGTGAAT TAATTAGAGC 12660
 GACTAAAGGA CGTACAATTT ATATTGCAAC GACATTCCAT GCTTCAATGA CATTCCGACT 12720
 TATTTTCTTG TTTAGCGAAG AAATCGGCGA TCTATTTTCA ATCAAAGTCA TCGCCATTTC 12780
 10 AACAGCAATC GTTGCAAGTAG GATACATTGG TTTAAGCTTA ATTATCCGAG GTATTGCATA 12840
 TTTAACAACA AGACGAAACC TTGAAGAACT TGAGCCTAAT AATTATTTAG ACCATGTCAA 12900
 TGACGATGAA GAACTAATC ATACTGAGGC TGAAAAATCT TCTTCAAATA TTAAAGATGC 12960
 15 TGAAAAAACA GGTGTAGCTA CTGCATCAAC GGTTGGTGTG GCTAAAAATG ATACTGAAAA 13020
 TACAGTGGCT GACGAACCAA GCATTCTATGA AGGTACTGAA AAAACAGAAC CTCAACATCA 13080
 CATAGGTAAT CAACTGAAT CTAATCATGA TGAAGATCAT GACATCACTT CGGAGTCAGT 13140
 20 AGAATCAGCm GaATCAGTTA AACAAGCACC ACmAAGTGAC gATTTaACAA ACGATTCAAA 13200
 TGAAGATGAA ATAGAGCAAT CATTAnAAGA ACCTGCGACT TATAAAGAAG ACAGACGTnC 13260
 ATCAGTTGTA ATTGATGCAG AAAACATAT CGAAAAAGCT GAAGAnCAAT CTTcAGATAA 13320
 25 A 13321

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8549 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

40 ATGTGTTGTA AACTTTTATG TTGAAAAAGC TACTTATCTC AATGAAAACA AGTAGCATTT 60
 AATAAATTAA TTAGTATACA GCTAGTTTTT CTAATTGTTC TTTAACTTGA ATTAAGTTTG 120
 ACCGTATTAG AGAGGCAGAT TGATCCATCG TTTGAATTGC TTGTCCTTCA TTTTCGTTCA 180
 45 AGCCATTACA AACCACTTCA AACTGTTGTG CCATTGATC AAGACGCGCA TGAGCTTGTG 240
 TGTTTAAAT AAACATATCG TCATAATGTG ATGGCGAATA GATAATTCGT CGTTGTATAC 300
 AAACGTATAA AAACCTTGTC ATATCAACGG TTTTGGCATT TTTAAACCTC TGTGTTTTCC 360
 50 ACGCATGTTT GCCCTTATTT AAATAATTTG CCCTTTTTTC GCCCGAAAA AAAAACACAA 420
 AAAAATAACC ACACTCCTAA ATTAATAGGT GGTGTGTTT TGTTGATTGT AGGGGTATAA 480
 AAATAACCGC ATTATTAAAG ATACGGTTAC TCTGTTATCT GTAAATATAA TAGTAGTTTA 540
 55

	AAACAGGACT CCACATAAAA ATCAACTCCT TTATATACCA TAATGATACT ATATTTTCTA	660
	GTTTATTTCA ATTTTTCAGT TTTTAAAAAT GAGTTTCTGT TTTTATTTAT ACGCTTTTCT	720
5	GTTTCTTTT TAAATTTTAT CTTTTTGTTA TTCCATTCAT TGTAAAATTC TATTAAATTA	780
	ACATAAAATT TTTTCATGCCC TATTTTATTT GTTGATGAGA TATCAATGTA AAGACTCAAT	840
10	ATTGTTTTTA AATAGATTG ATGCAACGAC TGATAAACCG TATTACTATC TGCTATGTTA	900
	TTGGTAAAT GCATAGAAAA ATATTCTAAT TTATTCATGC AATATATATG GGTTCATTA	960
	TACTTCTTAA TGAGTGATT TATACCTTGC AATACGTCAT TACTTTTAAT AACAAATTCT	1020
15	TTTTCACCTG TCGAAAAAGT CCACTGTTTA TCTCCTATAT TTTCTTTAAT TGTTTTCTTG	1080
	TTGTCAAATT CTAAAATTAT AGCCCGTAAA CACTCTTCTT TATAATTCTC GTTCTTGAAA	1140
	GTACGAAGCA AAATTTTAT AAATTCGGTA TTGGTGACTT TTTTATAAGT GTGATATTTT	1200
20	GCAATCTCTT TATCAGTAAA GACTGTTCTT AGTTCGTGAT TATCAAACT TAAATTCATC	1260
	TTATTCTCTA ATTCATTAAT TTTATCTTGC AAACCAACAT TTTCTAAAAT TTTCTTGTTT	1320
25	ATCTCCCCTA TATCAAACT CCTTTTCGAA ATTAATTTTG AAAACTCGTC TGCCATTTCa	1380
	ACAGCCTTTT CTTTCCTTTT ATACCTTTTG TTAAATTTAT GAACCACCGT TGCAGCATAA	1440
	TACGATATCC CACCAGATAA AATAGATGaT ATTATCGGTA TGTATATATC ACCTTTCATA	1500
30	TTTCCACCTC TTTTAACACA ATTAAGTATT ATGATACACA ACTTGCGCAA AAAGATGTAG	1560
	ACAGAACATA ATGGCGAACA AAAACAACCA CCCAGTAACT AGTATGGGTG GCGTAgACTA	1620
	TAACAACTCT ATGTTATCAA GATATATGTA TCGAGTGATG GCAAGGAAGA AGTCTCCTGC	1680
35	GGGACCAACA GTCAGATATA TGGCCTCTGC CGGGCTATAT AGTTCACTCC TACTATATAA	1740
	AAGTAAGTAT AACATAAAAA GCACCCCGTA AACTGTTATA CGGGAATGCT AAAGTCATAT	1800
40	ATACTACGGG GAGTAGTATG AAAACTATGC TCTCTATCGT AAGAAAAAAC ACCCAGTGAC	1860
	ATGCTTGGGT GAACAAGGAT AGATGTAAAT AGTTGATGCA TGTGTAcACA TCATAACAAA	1920
	AAACTAGCCC GAAGcTAGCT ATAACATAAA AAAATAGGCA AGTACCGAAG TACCTGCCAG	1980
45	TTACGCACAT TTAAATCTTG AGAGTAATGT TAAAAAGTGT ATAGGAATAT TAACATCCAT	2040
	CCAAATAGTT ATTTAATAAC TGTAAGATTc CCTATAATTA ATGTAGCaAA ATTTTATTTC	2100
	TAAGTAAATA CTAAATCGTG CTAAACTTAC CAAAACACT TATTCTATTA CCTGCCTTGT	2160
50	CTACCTCTCC TGTCGCTATA TAACGACGTT GTCCACTATT AGCAATATAA GTAATCCATC	2220
	TATAGCCATT GATGCAATAT GCGCCGTCAT ATTTAATTGT TCGGTTATTA GGTAATACAC	2280
55	CTGTAATTCT TGAATTAGTT GAATAGCCGT CCCTTACGTT ATTACCTTTA ACATTGGCAA	2340

	CTGGCACTGG TGGATTTTTT TGGTTTTTAG CTGATGTTTT AACATTACCA GCTACCAAAC	2460
	CACCTATAGG CTTACCATGA ATCGCACCGG CTATTAATTT AGAATACAAG TCATAGTTTT	2520
5	TCTTAATCCA ATCCATATCA TTTTATTAG TAATAAAACC TAATTCAGAT AAACGATAGT	2580
	TTATATTTAT TTCTGCTGAT ACATTAAACGT TTAGTAAATC ATTACGAGGT GTTACACCTC	2640
10	TTATTTGTCC TAAGTTATTT TTAATAACAT CTTGTATACT TTTATCAATA GTATCTGCAT	2700
	TGAATGACT TGAAATAATA ACATGCCCAC CACTTGCACT TTCTCCTGCT GCGTCTAAAT	2760
	GAATCTCTAG AACAATGTCA TACCCATGTG ATTTAACCCA ATATAAGCCA TAATCTTTAT	2820
15	TATTTCTTAC ATTAACACCG TAAGCAGTAT CTTGATACAT ATCTTGTGAT TGACTTGAGC	2880
	CACCATATAA TGCAACTTCG TGACCTGCAT GTCTTAAATA CTTAGCGATA TTTGGTGTTA	2940
	TATATTTACG GATAAAATCA CGTTCATTG TTCCGTTTCC GACTGCTCCA GGATCGTTAT	3000
20	AACCATGACC GGCTACAAGC ATAATTTTTT TAGGTTTAAT TACTGCTTGC TTTTGGCAG	3060
	TTGCTTGCTT AATAACGCTT TTAGCTTTAT CTCCAACACT TACTTTATCT GGGAAATTTA	3120
25	ATCTAATAAA ATACATTGGG TCATCGTAAT AATGAACATG TCTTGTAACG GTTTCGGGAC	3180
	CCCAACCAGG TTGCGCAACG CCATTTGTCC AACCTTTACC ATTCCAATTT TGGCCAAACG	3240
	ATGTGAAAGT GTTTAGATTA GCGCTCTCAA CAATTTCAAC ATGTCCaGct CCGCCACCAT	3300
30	ACTTTGACGG GAAAACGACA ATGTCCAAC TTTGCGGTAA AAAGCTATCA TAGTTTTTAA	3360
	TTATTTGCCC GTATTTTCA ATCCTTGCTT TATTATCAAA TGAATATTA TAAGCGTATA	3420
	AACCTTGTA CcTTTCGCTT GTTGCTATCA TAAAAACAT ATTTGCGTAA TCGTAACACT	3480
35	GAAATCCATA AAACAAATCA GGATTGAACT GCTTCCCTAA TGAATTATCA AACCATTTTT	3540
	CTGCTTGGTT TTTTGTATC AACATTGGTC AACACCTACC CTAAATCATT TGTGTCGTTT	3600
	ATAFTCGTAG GTGTCATTAC TTCTTTAATT GCGGCTTGCC CTGTTGCTTT TCTATACTTG	3660
40	TTTTCAGCTT TATATTTCTT TAGCTTTTGA TTTGCCCAT TACCTTCTTG AGATGTTGGA	3720
	TTATCTTTAT ATGTAGTATA TAAAGCAACA ACTGTTAAGA TAATCGATGA AACACTTTCT	3780
45	TCATCTACTG GTATCGGACT TATACCTTTA TTCGCTAAAA ACTGATTGAC TAATGCTAAG	3840
	ATCAATACGA TGTATCTTGT TATTACTTTT GCATCCATTT GTTTGCTCCT TTTATCCAAA	3900
	ATAAAAAGCC AGTGCCGAAG CACTGACTCT TAACTATTAC TTACACTTAC TAAACCAGAA	3960
50	ACACGACCAA AAGCTATATC CTAAATTTCC CTTAAGCATG GTAATCACCT CCTTTAAATG	4020
	CCAAAAATAG TTTTAAACAA GGCTATAACA AATGTACTTA GAATCGTCCC TATTAATCCT	4080
55	AGAATCCACA TCTTGATGTC TCTAATATTT TTAGCATTTT TCTCTTTATT TTTTTCATCT	4140

	TGCGTTCTCA GACTGTCTTC TATTCTGTCTG AATTTTTCAA ACATAGTCTT ATCATTCTCT	4260
	TCTAATCGCG TTAAACGCCA ATCTTGTTCTG TGTCGTTTGG TAAATCCAAA CATTACACCA	4320
5	CCCACTTTAT TCAAATTAAA AAGCCATAAG ATTATAACCT ATGACTCTAG ATTTTCTGGA	4380
	TACTTTTCTC CTGTAATAAT TGCATATTCC TCTTTATCTA TAACTTCCAT ATCTACATAC	4440
10	CACGCTATAT CTTCTTTACT ATATTCTTTC AATTGATACC ATGTTTAAAT ATCTTCGAAT	4500
	GTTGGTGAAA TTAATTTAAG CATTTTCAGT CTCTCCTTTA ACCTCTTCTA ATTTTTTATT	4560
	AAGTGTCA CA AGTTGTTTTG CCATTAGTGC ATTTTGCTTA TTAACCTGCA TCGATAACTT	4620
15	TGTACTTTGA ACAACTTGTT TCTGCATACT AGCAACCATT TTTCGTAAGA TGTCATCAGA	4680
	AGCGACTGTG TTTTGTTCTT CACTGTCAAT CTGTTGATGC AAGTCATCTT TTTCTTCTGA	4740
	ATAATCTTCG TTAATAACTA TTTCCCCATT TGAATATTTA AAGGCTTTAG GTCTAAAAAC	4800
20	TTGAGAGAAA TTTTCTGGTA AATTTTCAAT ATCAATACCT TCTTCAAAGC CACCAATGAT	4860
	AGCGTATGAA ATTATCTCAT TACGCTTGTT AACTAATATT TGCATTATTT TCTCACTCCT	4920
25	ATAATTTTGT TAATTGTCCC TCTATTTGCG TTCGCACCAG AGCCTCTTTG ACTTCCTAAG	4980
	TCGAAATAGA CATCGTTTGA TATAGTTAAA GATGTACGAC TAGATTTAGT TAATCCAAAC	5040
	TCATAAACAC CTCCACCATT TCCATCACCA TCTGGAAGAT TTGAGGGATT CAATGAAATC	5100
30	TTTCCTCCTC CAAAAGGACT GCCAAACTCT GTAAAGTCAC CACCTGGAAA AGTCCCATAA	5160
	AAAATTAATA AAATAAATTG GTCTAAACTC TCATTTAAGT ACAATGTAGA GCCCACACCA	5220
	TTTGCTGTTT CATCAAAAAT AACCGAATAC CTTTATTAA ACTTGTCATC TGCGTATAAT	5280
35	TTAGCGTTAC TTTCGGCCAT ATTAGCTTTT GATTGGGCAC TTTGAACAGT TTCAAAGGT	5340
	GTATTGTAAT CATTAATAGC TAATTCTGAC CACTCAGACC ATGAACCCGC TTCTTTTCTT	5400
40	TTAAACAAATA CTTTATTTGT ACCGTTCTGGT CGATAAGTCA TACGCTTGTA ATCTGAAGTT	5460
	ACTACTAAAT ATTCGACAGT ACCGTTAGTA CTAACACCTC TTGGATAATT TATAGCTTGC	5520
	GAAACATAAA TAAATTGGGT TGAATCACCT ATTCTTTGTT CTGGATTATT AAAATCAAAT	5580
45	CCAGTAATCT GCATTATCTT ACCATCATCT TTAGTAATCT TAGCTTTTTG CCAATTTGAA	5640
	GTAGAACCAC TTGTGACTAA ACCACCACTA TTCACTGACT GCTTGAAGGC TTCATGTTTC	5700
	TCATCCATAT ATCGCTTTTG CTCATCGAAT GTTCTTGAAT ATGCTTGCGC TTTATTTTCC	5760
50	AAATCAGATA TATGGCTATT AGCAAGTTGC TTTAATTCAT CTATACTTGA AGATTTTGCT	5820
	ATTTGAATAT CTGATAGACC TTTTCTTTA GCTTTTTCAA TCAGACTCGC ATAATCTTCA	5880
55	CCATTTTTTA TAGCCTCGTC CATTGCTTTC GCACGATCCA TAATAGTTTT TTCTAATTCC	5940

	TCAACGTTAA ATGTGATAGT TCTCTCGACA ACTACCACGT CTGAATTACC TAATTCTGCA	6060
	ACCGAAACTT GAGCTTGATA ACTTCCATCT CGTTTAAATTA CATCATTAGG TAATTGAAAT	6120
5	TTTAAAATAC CTTTAAATGG ATCTAATATT TCTAGTGGAG CAACTACCAT GACTCCTTTA	6180
	CCTCGAATCG CTATTCGTGC KTTGATATTT tCTTCACTCA ATAATAACGG TTGATTATTT	6240
10	TTAGTGATAT TAAAAAGAAG AACAGAAGAA TCACTCTCTC CTGTTCTAAA AGTTATATCT	6300
	AGATTTGAAA TATTTCCATA ATGCGCTGTG TTTTCTAAAT TTATAGCTAC AGATTCTCT	6360
	AAATTACTCA TTAACCTATA ATTCTCCCTT CGTGTAAGT CCATGGCCCT GAACCTGTTT	6420
15	TACTATCATA ATTTTTCAAT AGTATCTCAG CAGATGCTGT AACACTATTA CGAACTAGCC	6480
	TATGAACAAA GCCACCTGTG TTTGAAGCTT CTACATATAA GTTCCAACCA GCTACCCCTT	6540
	TACGTTCACT TGGAAAATCT GTAAAACGTT TTGTATCATC CGTAGTTAAA TAAAACGACA	6600
20	TGCCTACTAT GTTAATATCT GACATTTTTG TGATGAATGA AGGTACTCTC TCCCATTAC	6660
	CACTATTTTT AGGCACATAA TTCCAGTCCG AAATGTCTCC AGTTCCTCCA GAAAGCACCC	6720
25	TTTCAAAAGT CATCATATTC CTTGCATAAC TATTACGCGT CAATATCTGA ATTACATCAC	6780
	CGCCAGTTTG TGGTGGCTTA ACTTCCAAGA ACCAACCTGC ATCACGCCAT TCTCTGGTA	6840
	ATGGGAAATC ATCGATTGA ACTGTATGAT CAGTGTATAA ATAGTAAAGA CCTGGCTCTG	6900
30	TTAACATCCC AAGATTCTTA AGTTTATCAG GCCTCATTGG TAAAGGTTTA ACTCTACCAC	6960
	CTGTGTCACT CaTGATAAAA GGAACGCCTC TTGAGTGAAG TATTTCTAAA ATACCTCTTT	7020
	GCCCAATCAT GAAAATACGA TGTGTTCTAT TTCCaTCACC ACCGACAGTA ACACCTAGCA	7080
35	TCAAAGCTTT TTTACCACTA TCTTTGTCAT AGTATATTTG CAAACCTTtC TgCTTCCGCA	7140
	AATTGCGCAG GAAATGAATC tAgTGTTCa CCATAGTCAG CATTAACTG ATACGCTTCT	7200
	TCTCCTGTTT CTAAATCGAA AGCCGTTAAA TAGTTTCTAT TATTTGGATT ACTGTCTCCT	7260
40	GTATACCAAT ACAAGTATTT TTCATCAAAA GTCACACCCT GCATTGGTTG GGTTCGTTT	7320
	GTTAGTCTCA TAGGGATACT GATTTTATGC AAAACTTTAT CAATATTTTT ATCAACATCG	7380
45	TCTAAACTTC TTATCTCTAT ATAAATCATT GAGTTTTCAA GTTCCCACTG ACTTCTAGGT	7440
	CTCTCaATTC TGTATAGAAT TTTATTTTCT TTTTCATTTA TGACAGGGGT GATGTAGGGT	7500
	TTTTCTGGGT GTCCTGTAAA TACATCTTGC ATACCATACT TGCCATAGCT AATTTCCACA	7560
50	TTAGGCGTAT ACTTGAAACG AACTAATGTA TTCTCATTAT TACCATTAA GATAAACTA	7620
	TAAATCCATA ACTCATcATC AATATATCTA TAACCGTTAT GTGTACCATG ACCCCCACCT	7680
55	ACAATCAATG AGCTGTCTAT AAATTGACCA TTAGGTCTTA GACGACTTAG CATATAGCCA	7740

ATTACTGCAT TTGTAAGAGG TGCAAGTTCT GTCACAAATA AAAATTCTTG CTTATCAGGT 7860
 TCAAAACGAT ACTCGATATC AAGAATTTCT TGT TTGGTCT TATTTAATTC TCTTATAGTT 7920
 5 TCCTCTTTAT TAATTTGAGT TTTGGTTTCC CAATCGTCTA AATGTTCTTT TAATGTGTCA 7980
 AAGGTTTCGC CGTTTACATT AACTCGAGCT TGAACAATCT CATTAGCACT GTTATTACGT 8040
 10 GGTGCCACAA CAAGTGC GTT AATTTGACTT TGTAAGATT TGTTTACTGC TGCTTGCGAT 8100
 CTACCATTAT AATAAATTTG CTCAGCGAAG TGTGAATTG TTTAGCTyT CTGATGCAAC 8160
 TTAAACTCTG TTGTCAAGCC AAGCGCAAAT TGCTCTATTC TTTGTAAGTT TTGTATTTC 8220
 15 TTAGCTCTAT AATCTCGACC TGCTAAAGCT CCCAAATCCT TTATTAAATA CAAATTTTCC 8280
 ATAATGCACC TTCCTTTCTA ATAAATAGC ACTGTACCAA GTTCCCACT ATCGTCAACT 8340
 GTTATTTTCC ACAATTTACC GTTTGGGGAT TTCTGTACAA TGCTATTTTG AATAATTgcC 8400
 20 TGctTCGCCT ATTTTAAAT TATCTAATTT ATTTktATCA TTTACCGAAA TGATACCGTC 8460
 TTGAGGCAAT CCATCAATAN CACTACTGCC TGCATAAGGT ATCCCATTTA TAGCTTTCCA 8520
 25 ATGTGTAGCT GGAAAGTACT GTTTATCGT 8549

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3601 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AGGCGTGTAG TGACTTACGG nTAGGAACT ATGTATCCGA ATGATTTATT GAGACCAAAA 60
 AGGCATTAAA GTCCATTGAA ATATChGGTA GCGmGTTGGT ACgTGGACGT GGGGGCCCTA 120
 40 GATGTATGAG TCAACCATTA TTCAGAGAGG ACATTTAACG TAATAAATTA TAGAmACGAG 180
 GGTGAAAATA ATGACAGAAA TTCAAAAACC GTATGATTTA AAAGGCAGAT CATTATTAAA 240
 45 AGAAAGTGAT TTTACCAAAG CAGAATTCGA AGGACTTATT GATTTTGCAA TTACATTAAA 300
 AGAGTATAAG AAAAACGGTA TTAAGCATCA CTACTTATCT GGAAAAATA TTGCACTACT 360
 ATTCGAAAAG AATTCGACGA GAACGCGTGC TCGTTTTACA GTTGCGTCTA TTGATTTAGG 420
 50 TGCGCATCCA GAATTTT TAG GAAAAATGA TATTCAATTA GGCAAAAAG AATCTGTAGA 480
 GGATACTGCG AAAGTATTAG GTAGAATGTT CGATGGTATT GAATTCCGTG GTTTTTCACA 540
 55 ACAAGCTGTT GAAGATTTAG CGAAGTTCTC TGGTGTACCG GTGTGGAATG GATTAACAGA 600

	TCTAGAAGGA ATAAACTTAA CTTACGTTGG AGATGGACGT AATAATATTG CGCATTTCATT	720
	AATGGTAGCA GGTGCTATGT TAGGTGTTAA TGTAAGAATT TGTACACCTA AATCATTAATA	780
5	TCCAAAAGAG GCATATGTTG ATATTGCAAA rGAAAAaGCG AGTCAaTATG GTGGTyCAGT	840
	CATGATTACG GATAATATTG CAGArcCAGT TGAAaTwCm GATGCTATAT ATmCAGATGT	900
10	TTGGGTATCG ATGGGTGAAG AAAGTGAATT TGAACaCGTA TTAATTTATT AAAAGACTAT	960
	CAAGTGAATC AACAGATGTT TGATTTAACA GGTAAAGATT CAACGATATT CTTACATTGT	1020
	TTACCAGCAT TCCATGATAC AAATACACTT TATGGACAAG AAATTTATGA AAAATATGGA	1080
15	TTAGCTGAAA TGGAAGTTAC AGACCAAATC TTTAGAAGTG AACATTCAAA AGTGTtTGAT	1140
	CAAGCTGAAA ATAGAATGCA TACAATTAAG GCAGTAATGG CAGCAACATT GGGGAGTTAA	1200
	TCACTAAATG GAACGATATG AATATGATGT GTCTGATGAT ATAAGTGTCA TGTACAGACA	1260
20	CCTCATATTG GTATTAAAGG AGAAATGAAT ATGAACGAAT CAGGAGATAA CAAACTCAGT	1320
	AAATCTTCTT TAATTGGACT AGTTATAGGA TCCATGATTG GTGGCGGTGC GTTCAATATA	1380
25	ATGTCTGATA TGGGCGGTAA AGCCGGTGGA TTAGCCATTA TTATTGGTTG GATTATTACA	1440
	GCTATAGGAA TGATTTCATT AGCGTTCGTA TTTCAAAATT TAACCAATGA ACGGCCGGAG	1500
	CTAGACGGTG GTATTTATAG TTATGmTCAA GCAGGATTTG GCGATTTTGT AGGATTTATC	1560
30	AGTGmTTGGG GATATTGGTT CTCAGCGTTT TTAGGCAATG TTGCCTATGC AACACTATTG	1620
	ATGTCAGCAG TAGGTAACCT TTTCCCgATT TTTAAAGGAG GCAACACATT ACCAAGTGTT	1680
	ATTGTGCGCT CGTTACTACT CTGGGGTGTC CATTTCTTGA TTTTAAAGG CGTTGAAACA	1740
35	GCAGCATTTA TCAATAGTAT TGTTACTGTT GCAAAGTTAA TACCGATTTT ACTTGTAATC	1800
	ATATGCATGA TAATTGCATT CAATTTTGAC ACTTTTAAAA CAGGCTTTTT CAGTATGACG	1860
	TCAGAGGGTG TATTGCCATT TAGTTGGGCG AGCACAATGA GCCaaGTtAA AAGTACGrTG	1920
40	CTAGTGACAG TTTGGGTGTT TATCGGTATC GAAGGTGCAG TAATTTTTTC TAGTAGAGCT	1980
	nAAAATGAGA AAGATGTAGG TAGTGCCACG GTTATAGGAC TTATATCAGT TTTAATTATC	2040
45	TATyTCTTAT TAACTGTATT AGCTCAAGGC GTGATTTTGC AAAATCATAT TTCGCAATTA	2100
	GATTGCGCAA GTATGGCACA GGTGCTTGCA ACTATTGTAG GTGGTTGGGG ATCTACACTT	2160
	GTAAATATTG GTTTAATTAT TTCGGTACTA GGTGCATGGT TAGGATGGAC ACTGCTTGCT	2220
50	GGTGAATTAC CTTTCATTGT TGCAAAAGAT GGATTATTTC CAAAATGGTT TGCTAAAGAA	2280
	AATAAAAATG GAGCACCTGT AAATGCACTG CTTATTACCA ATATATTAGT ACAATTATTT	2340
55	TTAATAAGTA TGCTATTTAC ACAGAGTGCG TATCAATTG CATTTTCACT AGCATCAAGT	2400

	CGACAGCAAG CAACTACTAA ACAATGGACG ATTGGTATCA TAGCCTCAAT TTATGCTATA	2520
	TGGCTTATAT ATGCAGCAGG TATCAATTAC TTATTATTGA CGATGTTACT TTATATTCCA	2580
5	GCTCTTCTTG TTTATACaAT CGkTCmAAAG rATwATCAGa CACGTTTGAT TAAATCAGrC	2640
	TATATTCTtTT TTATGATTAT tATCGTACTT GCAGTTATCG GGTTAATTAA GTTATTGATG	2700
10	GGAACGATAA ATGTTTTTTTA AAAGGAGCGA CAAAAATATG AAAGAGAAAA TTGTCATTGC	2760
	ATTAGGCGGT AATGCGATAC AGACAACAGA AGCAACAGCT GAAGCACAAC AAACAGCTAT	2820
	TAGATGTGCG ATGCAAAACC TTAAACCTTT ATTTGATTCA CCAGCGCGTA TTGTCATTTC	2880
15	ACATGGTAAT GGTCCACAAA TTGGAAGTTT ATTAATCCAA CAAGCTAAAT CGAACAGTGA	2940
	CACAACGCCG GCAATGCCAT TGGATACTTG TGGTGCAATG TCACAGGGTA TGATAGGCTA	3000
	TTGGTTGGAA ACTGAAATCA ATCGCATTTT AACTGAAATG AATAGTGATA GAACTGTAGG	3060
20	CACAATCGTT ACACGTGTGG AAGTAGATAA AGATGATCCA CGATTTGATa ACCCAACTAA	3120
	AcCAaTTGGT CCTTTTTATA CGAAAGAAGA AGTTGAAGAA TTACAAAAAG AACAGCCAGA	3180
25	CTCAGTCTTT aAGAAGATG CAGGACGTGG TTATAGAAAA GTAGTTGcGT CACCACTACC	3240
	TCaATCTATA CTAGAACACC AGTTAATTCG AACTTTAGCA GACGGTAAAA ATATTGTCAT	3300
	TGCATGCGGT GGTGGCGGTA TTCCAGTTAT AAAAAAGAA AATACCTATG AAGGTGTTGA	3360
30	AGCGGTTATA GATAAAGATT TTGCTAGTGA GAAATTAGCA ACGCTGATTG AAGCAGATAC	3420
	CTTAATGATT CTTACGAATG TAGAAAATGT ATTTATTAAAC TTTAATGAAC CTAATCAACA	3480
	ACAAATCGAT GATATTGATG TAGCAACACT GAAAAAaTAC GCGGCACAAG GTAAGTTTGT	3540
35	GGAAGGATCG tGTTGCCAAA AATAGAAGCT GCGtACgtTT GTTGaAGtG GGGaAACCAA	3600
	A	3601

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 573 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

50	CGACACTATT AAATGAATTA GAGCACAATC TAACAAATCA AATTCATTTT TCAAAAGATG	60
	AACGACTCAC ACATATCGCT TTAAAGTTAT TCGAAACAAC CGATCCTGTT TCAACAAAGC	120
55	AACTTGCGCA AGATGTTAAT GTTTCGCGTC GGACAATTGC AGATGATATT AAAATGATTC	180

TTATTGGTGA GGAAGATCAT TATCGTAAAG CGTATGCACA CTTTATACAT CAATATATGA 300
 AACAAAGCTGC ACCTTTTATA GAGGCGGATA TCTTTAATTC AGAATCAATC GCATTGGTTC 360
 5 GCCGTGCCAT TATTAAGACA TTAAATAGTG AAAATTATCA TTTAGTTCAG TCGGCTATCG 420
 ATGGCTTAAT CTATCATATA CTCATTGCCA TTCAGCGTTT AAATGAAAAT TTTTCGTTCC 480
 ATATACCTAT CAATGAAATT GATAAATGGC GACATACTAA TCAGTATGCn ATTGCTTCAA 540
 10 AAATGATAGA AAACCTTAGAA CGCAGTGTA TGT 573

(2) INFORMATION FOR SEQ ID NO: 8:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1221 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTGATATTTA TAACGTTATA TTTTAATAGT TCACCTGGAT TATTAAATAA ATAGTCCGCC 60
 25 AAAATTTTCTT TTTCTTTATC AATCTGaTgK TAATTAACaC TTTCGaCTTC TGTAGGAATT 120
 CTAATGTCAA CAGAAGCATT GATATAAGCT TGATGTTGCA TGCAATCACA CTCCTAATCC 180
 30 TTCATmTmAA ACGGAGAAGT AAACCCGTCA CTATTCAAAT TCAATCCTTT TGCCCAATCA 240
 ACAGGCTTAT TCATGATAGT TTCGATTTCC TTAAGTCCAT TTGAACCTCT AGGTATTTCT 300
 ACAATTACTT CATCATGGAC ATGGCCAACT ATTTTAAAAC CTAATGCTTC AAGCCTTGCT 360
 35 ATAGAAATCG CAAGTAAATC CCTTGCAGTT GCTTGAACAA TATTCTCGAC TAACTTCCCA 420
 CCATACGTTT TTAACTTTGA CCATTTACGG TTAAGATCTA ACCCCATAAA TTCAACAACT 480
 TGACTACCCC AACTATTTTC ACCAACTAAA GCTTTTGGAT AAGCTAAAGC TCTTCCACTA 540
 40 GGCAGTTCAA TCATTAGAAA ACCTTTTTTTC ATATAAAATC TAAGTCCATG TGTATGATGC 600
 GTCTTTTCGGG ATTTTACAGT ATTAATTGCA GCCTCTTGGC AAGCCTTCCA AAAATTAAC 660
 45 ATGTTAGGAT TTGCGTTACG CCAACTATCA ACTAAACCTT GTAACGCTT TTCTTCAATG 720
 CCCATTTCCA ATGCACCCAT TGCTTTTAAA GCTCCAGCGC CACCTTGATA GCCTAAAGCT 780
 AATTCGGACA CTTTTCTTTT TTGTCTGAGA GGGTCGCCCT TAGTTATGCT TTCTACCGGT 840
 50 ACATTAAACA TTTGAGAAGC CGATGCTTCA TATATCTTTC CGTGTGTGTT GAATACATCT 900
 AAACGCCATT GTTCTTTTGC ATACCATGCT ATGACTCTTG CCTCTATTGC AGAAAAATCA 960
 CTTACTGCTA GTTCATTACC TTCTTCAGCA GTAAATGTCG TCCTAACTAA TTGACTTAAT 1020

55

AGATCTCTTG CTATTTCTAA TTCAGTATCT GAAATATAAT GCTTTGTTAA ATTCTGAAGT 1140
 TGTACACCTC TACCTGCCCC TCTTCCAGTA CCGGCACCGT AAAATTGAAA CAGACCTCTT 1200
 5 ACCCGTTCAT CACTGCACAT C 1221

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1090 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTTTGTTTGG TATGAGGTAG CAATGACGAC GTGTCATTGG TGGAGATTGT AAAAATACAT 60
 20 AATAAAAAGA AGCGGCAATG TATACCGCTC CTTTTTTATA CTACATACCG ATTTTCAACC 120
 ATCTCTTTCT ACTTAGTAAT AAGACAATAG TATTAAGTAT AAATAGAAGA ACGAAGAATG 180
 ATACTATATT TATAATTTCA GTAGGACACA TAAATGTTGA CTCGTTATTC AATATTTTTT 240
 25 CTACGGCAGC ATACATCGTA TTGCTCGCCT CAAATGGAGC AACGATACCA AATATATTTT 300
 TATTAATGGC AACTAAGATG ACTGAACCAA TCCAATATAC AATGCTGATA CCTAAGCTGA 360
 30 TTAAAATGTT AGGTGAAACC ATACTAATCG TTCCAACAAC TAAGATATAT TGTAAAGATA 420
 CGAGTGAAAA TAAGATTATT AATAGTAAGT AATGTGAGAA ATCCGAATAT ATAATTGAAA 480
 TAATAGTGAT ACTTAGAATT ATGAACACTA AACATTCAAA AAATAACACT GCTACCTTTT 540
 35 TATAGAAGAA GGTAAAGATA TTATCGCCAA TCAATTTATA AAACAGGATA TTTTATTTCG 600
 AATACTCTTT ATTAATAAAA TATGCAATAA CAAATGAAAA TAGTAAGAAC CCTAATTGCG 660
 TTGCAACAGT ATATGAACTG AAGAAAAACT GGCTATAGCT TAACTTTTAA ACTTTGTCTA 720
 40 TACCTATTGG TAAAAAATAC CCAAGTAAGA AAAGGAATGT GAATAGCACA ACAAGCGTGT 780
 AAATAATTTT ATTGGAAATA CTTTTTTTAA ATTCTAATTT CAAAGTGGAC ACCTCAATTA 840
 45 TAAATTAATG TAATCATTTA TGACTTCTTC TTTTGATTGG TACTCTTCTA TTTGAAGGTC 900
 TTTAAAAATA AAGTATTTAC CCGGCAAAGC ACTTAAATCG GATAAATTaT GTGTAATATT 960
 GATAATAGTT TTAGTTTGAT GGCTTTGAAT AAAATCATT AAAAATTCAT AAATTTTATT 1020
 50 AACTGTTTTT TTGTCTAAAG CGTTTGTAAC TTCATCTAAT ATGATTAAAT CATGATCTTC 1080
 CAATAAGAAA 1090

(2) INFORMATION FOR SEQ ID NO: 10:

(A) LENGTH: 904 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

10 TTAGGACTAT TTTATCATAT TCATTTAAAT TACGGCTAAA AATTTTAAAA ACGGGGATTA 60
 ATATATGGAA TTAAGCTATG AAAGTTAATT GATACTTGCA TTTTACGCTG ATTTATATAA 120
 GAATAACTAT TGTATAGTTT TAAAAACGAA CGTACGTTTG CAGGAGGCCA AATCATTGGC 180
 15 AATGAATAAA CAAAATAATT ATTCAGATGA TTCAATACAG GTTTTAGAGG GGTTAGAAGC 240
 AGTTCGTAAA AGACCTGGTA TGTATATTGG ATCAACTGAT AAACGGGGAT TACATCATCT 300
 AGTATATGAA ATTGTCGATA ACTCCGTCGA TGAAGTATTG AATGGTTACG GTAACGAAAT 360
 20 AGATGTAACA ATTAATAAAG ATGGTAGTAT TTCTATAGAA GATAATGGAC GTGGTATGCC 420
 AACAGGTATA CATAAATCAG GTAAACCGAC AGTCGAAGTT ATCTTTACTG TTTTACATGC 480
 25 AGGAGGTAAA TTTGGACAAG GCGGCTATAA AACTTCAGGT GGTCTTCACG GTGTTGGTGC 540
 TTCAGTTGTA AATGCATTGA GTGAATGGCT TGAAGTTGAA ATCCATCGAG ATGGTAATAT 600
 ATATCATCAA AGTTTTAAAA ACGGTGGTTC GCCATCTTCT GGTTTAGTGA AAAAAGGTAA 660
 30 AACTAAGAAA ACAGGTACCA AAGTAACATT TAAACCTGAT GACACAATTT TTAAAGCATC 720
 TACATCATTT AATTTTGATG TTTTAAGTGA ACGACTACAA GAGTCTGCGT TCTTATTGAA 780
 AAATTTAAAA ATAACGCTTA ATGATTTACG CnwGGgTAAA GAGCGTCAAG AGCATTACCA 840
 35 TTATGAAGAA GGGATCaAAG rGTTgTTAGT atGTCCAaTG ArGGAAAAGA AGTTTTGCCT 900
 GACG 904

40 (2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11271 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

50 GATTTCTAAA TCAAGATCTG TTTTACGATA ACCATTCAAA CCTTGACGTT CATCTTCTTC 60
 AGGTTGATTT TGTGCTGTG TGTCTTTGTT GTCAGAAGTC GCTACTGTTT TTTTATTATC 120
 55 TGTTTCTTTA GTCATAACAA ACGCCTCCGT TATAAAACGC TATATTTAAT GATATGTGAT 180

TTAATAAGAC GATTCAGCAA GTTTTAAAGT ATTATTTGAC TATGTTGGAT TAGGCATCTA 300
 GTCCTATAAT ATCACTGACA TTGTCAAAAT GATGATCTTT TAAGTAACGT GCGATGCCTT 360
 5 TGTTCATTTT CTTAGTTAAA CCTGGGCCTT CAATAACAAG TGATGAATAA ATTTGAATAA 420
 GTGACGCACC GTGACGCATC ATTTTGATTG CATCTTCAGT ACTGAATACG CCGCCTGTAC 480
 10 CTATAATTAA AAATTCACCA TTTGTTTGCT GATAAgCATa CTTAATCAAT TTTAAATTAC 540
 GTTCAAATAA TGGACGACCA CTCAAACCGC CTTCTTCGAC TTTATTAGCA GAAGTTAAAC 600
 CATCTCGTTG TCGCGTTGTG TTTGCTAAGA TGATACCGTC AAATGTCTCA GTAATCGCTG 660
 15 GTAATAGTGC TTTTAAGCCA TCGAAATCCA TATCAGACGT TAGTTTTAAA TAAATTGGCA 720
 CTGTTACATC ATGTTGTTTT TTAAATGCTG TTAAAGCTTG GCATAACATT GAAAATTCAT 780
 CTTTATCATG GAAGTTTTGA AGATTTTCAG TATTTGGAGA ACTGATGTTG ACTGTGAAAA 840
 20 ATGAAACGTC GTGTTTAAAC GTATCAATAA CCTTTATATA ATCTTGATAA CGCGCTTCAT 900
 AAGGTGTCAT TTTATTCACA CCAACATTGA TACCAACAGG TACTTGATAA GCATTTTTTAC 960
 GCAAATGACT TAGTGCTTTG TTCATACCAA TATTATTGAA GCCCATTGCA TTTATCAAGG 1020
 25 CGTCATCTTC TAATAATCTA AACATGCGTG GTTGAGGGTT ACCCGGTTGA GGTTTAGGTG 1080
 TGATACCACC TAATTCTAAA GCACCGAATC CAAGGTGTTT CAATGCTTTT GGTACTTCGC 1140
 AAGATTTGTC GAAACCAGCT GCTAAgCCAA TTGGATTGTC GTACGTATTA CCTTGTATCG 1200
 30 TTTGTGATAA CGTTGGATTG TTATAAGTAA ATAGTTTATC GACGACTGGG AATAAAACCG 1260
 GaAACTTTTG TaACGTTTTT AATGCATCGA TAGTTAGTCC GTGTGCTTTT TCGGGTTTCGA 1320
 35 TTTTGAATAA GAAAGGTTTA ATTAATTTGT ACATGAGTAT GCTCCTATTT CATTATATTT 1380
 GAGGCTTACT ATCCTCAACT TAATATATGT GAAATATATT CTTTTAATAG ACTAGCATTT 1440
 CCATACATAA TTTCTAGTT AAAACTAAAA AGTTTGAAGA ATTGACGCAA gTTTGAATAA 1500
 40 CGTTTTTAAG ATTAAATCAT CCTAATTAGG CAATATTATA GTATAAAGTA AGTAGATTGG 1560
 AAGGTGTTTG TATGAATGAA CAATGGTTAG AGCATTTACC TTTAAAAGAT ATTAAAGAGA 1620
 TTTACCAGT GAGTGGTGGT GATGTAAACG AAGCATATCG AGTCGAAACA GATACGGATA 1680
 45 CATTTTTCTT ACTTGTCCAA CGTGGACGTA AAGAATCATT TTATGCTGCA GAAATTGCAG 1740
 GTTTAAATGA ATTTGAACGT GCAGGTATCA CGGCACCTAG AGTAATTGCA AGTGGCGAGG 1800
 50 TTAACGGTGA TCGGTATTTA GTGATGACGT ATTTAGAAGA AGGGGCTTCA GGGAGTCAAC 1860
 GCCAATTAGG GCAACTCGTA GCTCAATTAC ACAGTCAGCA ACAAGAAGAA GGCAAATTTG 1920
 GCTTCTCATT ACCTTATGAA GGTGGCGATA TTTCTTTTGA TAATCATTGG CAAGACGATT 1980

55

	GGCTATGGGA TGCCAACGAT ATCAAAGTAT ATGACAAAGT GCGACGTCAA ATTGTGGCGG	2100
	AATTAGAAAA GCATCAAAGT AAACCGTCTT TATTACATGG TGACCTATGG GGTGGTAATT	2160
5	ATATGTTCTT ACAAGATGGT CGTCCGGCGT TATTTGATCC AGCGCCATTA TATGGTGACA	2220
	GAGAATTCGA TATCGGTATT ACAACGGTAT TTGGTGGTTT TACGAGCGAA TTTTATGATG	2280
10	CGTATAATAA ACATTATCCA CTCGCAAAAG GTGCATCCTA TAGACTTGAA TTTTATCGTT	2340
	TATATTTATT GATGGTCCAT TTATTGAAAT TTGGTGAGAT GTACCGTGAT AGTGTGCGC	2400
	ATTCTATGGA TAAGATTTTA CAAGATACAA CAAGTTAGTT AAGACGTTAG ATTGAGATAA	2460
15	ATAGATAATA TGCACAGATA TTTTACAAT GAGAAGCGAT ACAGCTGCCT CAATAAAAAAT	2520
	ATTGTGCGT TTTTATTGTT GGAAAAATAA ATTTTAATCG CTATTGTTAA TTTCTGTAAT	2580
	GTAAAAAAG GTTGAGTTAC AATAAAAGTG ATTTTATAAC TTTTGTTC AATAAATTCT	2640
20	AGGAATGATA CATATTTATT GATACAATAA TTTTGAATAT AATCATAAAA CAATATTTAA	2700
	GTATAATTGA ATGTTTGAAT ATCATATATT GATACAGTTT CTAATAATTT TAAAAATAAT	2760
	TAAATGGAGA GAGGTGTAAA TGATGAGTAC AGTTCAAAGT GATATTTTTTA AGACCAATAG	2820
25	TGCATCATCA TCTATTAAAA GCGCTGTTGA AACATGTAAT AATGTGTCGA AACCGGATAA	2880
	AGATGAAAGT ACAACAGTAA GTGGAAATAA TAATGCTCAT AGTGTGATAG ATGATTTGAT	2940
30	GAGTAAGAAT CAATCTGTTG CTGAAGCAAT ACGAACTGCG AGCGATAATA TACAAAAAGT	3000
	TGGTGAGGCT TTTGACCAAA CTGACGTAAT GATTGGTAAT GAAATTGGTA AAAATTAAAA	3060
	CGTGGTGAAA TGATGTCGAA TAACTGGAT GAAATCAATA AAATAATCAC AGCGAAACAT	3120
35	GAGCAAATGG ATGACTTATA TGATGAAAAG CGAGAGGTTA AAGCATTGAT AGATGAAAGT	3180
	GATGCGCTTA ATCATTGAT AGATCAATTA TATCAACATT TAGGTGAGCG TTATTATAGT	3240
	AGCAATATGG CTAGTCGTAT GGAACAGTTC CGCGATGAAT TTCATTTTGC GAAACGACGT	3300
40	TCAACGGAAG CGTTATACGA GCAGCAACAG CAAATTCAAC ATGGCATTG TAAAGTGGAA	3360
	GAAGAGATGA TTGACTTGGA AATGCGAAGG AATGTGAAA TTGAGACGGT GACAAAGGAG	3420
	GAAAAATAAT GGAAACAATA GGAAGCATT TTTATTTAAA AGAAGGTTG CAAAAGTTAA	3480
45	TGATTATTAA TAGAGGmCCA aTTGTAGAAA TTGAAATCA AAAGTATATG TTTGACTATT	3540
	CTGCATGTAA ATATCCGATT GGTGTTGTAG AAGATGAAAT TTATTATTTT AACGAGGAAA	3600
50	ATATAGATTC AGTTATTTTT AAAGGTIATT CTGATCAAGA TGAGGTTAGA TTTCAAGAGT	3660
	TGTTTTGAAA TATGAAACAA AATTTGGATA GTGAAATACA ACGTGGAGAA GTTACACAAC	3720
	AATAAAGAAA TACTTTTTCT TTATTGGGGT GGGACGACGA AATAAATTTT GTAAAAATAT	3780

	ATGTCATTCA TAATCATTTG AACTAAACGT AGCAGCCTTA AATTTTAAAA AAAGACACAT	3900
	ACCAACTTCC GAAATGTAGA TGAATTCTCT ACAATAACGG AAGTTTTTCT TTTAATATTG	3960
5	AAATTTCTCA AGGATAGGTC TATACTTTAT AAATCGTAAT TATTACGATT TATAATCAAA	4020
	AACAATAACT TGAAATAGAT CATTGAGGGA GTGTTAATAT GCAACATCAT AAAGTGGCTA	4080
10	TTATcGGTGC CGGTGCTGCA GGTATAGGTA TGGCCATTAC CTTAAAAGAT TTCGGTATAA	4140
	CAGATGTCAT TATTTTAGAA AAAGGAACAG TAGGACATTC ATTTAAACAT TGGCCGAAAT	4200
	CGACCCGTAC GATCACGCCA TCATTTACGT CTAATGGATT TGGCATGCCT GATATGAATG	4260
15	CAATTTCCAT GGATACTTCA CCAGCATTTA CATTTAATGA AGAACATATT TCCGGAGAAA	4320
	CATATGCTGA ATATTTACAA GTGGTTGCCA ACCATTACGA GCTGAATATC TTTGAAAATA	4380
	CAGTTGTCAC AAATATATCT GTAGATGATG CATATTATAC GATTGCAACG ACAACAGAGA	4440
20	TATATCACGC GGATTATATC TTTGTGCGAA CAGGTGATTA TAATTTCCCT AAAAagCCAT	4500
	TTAAATATGG TATTCATTAT AGTGAAATTG AAGACTTTGA TAACTTTAAT AAGGGGCaAT	4560
	ATGTGGTTAT CGGAGGTAAT GAAAGTGGCT TTGATGCTGC ATATCAACTT GCAAAAAATG	4620
25	GCTCTGACAT CGCACTTTAT ACTAGCACAA CCGGTTTAAA TGATCCGGAT GCTGATCCTA	4680
	GTGTTAGATT GTCACCTTAT ACACGTCAGC GACTAGGTAA TGTCATTAAG CAAGGTGCTC	4740
30	GCATCGAAAT GAATGTACAT TATACAGTTA AAGATATTGA TTTAACAAT GGACAGTATC	4800
	ATATCAGTTT TGATAGCGGA CAAAGTGTGC TTACACCTCA TGAACCAATA CTAGCAACTG	4860
	GCTTTGATGC AACAAAAAAT CCAATCGTTC AACAAATTATT TGTGACAACA AATCAAGATA	4920
35	TTAAATTAAC AACACATGAT GAATCGACAC GTTATCCGAA TATTTTTATG ATTGGTGCAA	4980
	CAGTTGAAAA TGATAATGCC AAATTATGCT ATATCTATAA ATTTAGAGCG CGATTGTCAG	5040
	TACTTGACACA TCTTTTAAACA CAGCGGGAAG GcTTACCAGC TAAACAAGAT GTCATTGAAA	5100
40	ATTATCAAAA AAATCAAATG TATTTAGATG ATTATTCATG TTGTGAAGTG TCATGCACAT	5160
	GTTAGAAGTG AAATATGATA TGAGAACTGG GCATTATACG CCCATACCTA ATGAACCTCA	5220
	TTATTTGGTT ATTAGTCATG CGGATAAACT TACCGCAACA GAAAAAGCGA AATTAAGATT	5280
45	ATTAATCATA AAACAGAAAT TAGATATTTT ATTGGCAGAA AGTGTAGTTT CTTcGCCTAT	5340
	AGCGAGTGAA CATGTGATAG AACAAATTGAC ACTATTTTCAA CATGAGCGAC GACATTTAAG	5400
50	ACCTAAAAA AGTGCACAT TTTTAGCCTG GTTGTGATA TTTTAAATGT TTGCATTGCC	5460
	AATCGGTATC GCTTATCAAT TTTCAGATTG GTTTCAAAAT CAGTATGTGT CAGCATGGAT	5520
	AGAATATTTA ACTCAAACAA CATTGCTCAA TCACGATATA TTACAGCATA TATTATTTGG	5580

55

	ATTGATTAGT TTATCAACTG CTATAATTGA TCAAACAGGA CTCAAATCAT GGATGATATG	5700
	GGCAATTGAA CCGTCAATGT TATGGATAGG ATTACAAGGT AATGATATCG TGCCACTATT	5760
5	AGAAGGGTTT GGATGTAATG CAGCAGCTAT TTCACAAGCA GCACACCAAT GCCATACCTG	5820
	CACGAAGACA CAGTGTATGA GTTTAATAAG CTTTGGTAGT TCTTGTAGTT ATCAAATAGG	5880
10	TGCGACATTA TCTATTTTAA GTGTAGCTGG AAAGTCATGG CTATTTTATGC CGTACTTAAT	5940
	ATTAGTACTT TTAGGTGGCA TCTTACATAA AGGATATGGT TGAAAAAGAA TGATCAACAA	6000
	CTTAGCGTTC CGCTACCTTA TGATAGGCAA TTACATATGC CAAATATACG TCAAATGTTG	6060
15	CTACAAATGT GGCAAAATAT ACAAATGTTT ATCGTTCAAG CGCTACCTAT TTTTATCACA	6120
	ATCTGTCTTA TTGTTAGTAT TTTATCACTA ACGCCAATTT TGAATGTTTT ATCACAAATA	6180
	TTTACACCTA TATTATCGTT ATTAGGCATC TCGTCAGAAT TGTCAACCAGG GATTTTATTT	6240
20	TCAATGATTC GAAAAGACGG CATGCTCTTG TTTAATTTGC ATCAGGGCGC CTTATTACAA	6300
	GGAATGACAG CAACACAGTT ACTACTACTT GTGTTTTTTA GTTCAACATT TACAGCGTGC	6360
25	TCGGTCACAA TGACGATGCT TTTGAAACAT TTAGGTGGTC AGTCAGCACT AAAATTAATT	6420
	GGAAAGCAAA TGGTGACATC ATTGTCTTTA GTTATTGGTG TAGGCATCAT TGTTAAATA	6480
	GTAATGCTGA TTATTTAAAA AAAATGAACT ATAAGTGAAT ATAGAGTCAT GTCAGTCAAT	6540
30	AGGAGATCTA TCTTGAATA TGCTATTCAT ATGAAGTATA AGAGGAGAGT CGCAGATGAA	6600
	AATAGTTATT ATAGGTGGGT TTTTAGGTGG CGGTAAAACG ACTGTCTTAA ATCATTTGCT	6660
	CGCTGAATCA TTAAAGGAAT CGCTGAAACC AGCAGTCATC ATGAATGAAT TTGGGAAAAT	6720
35	GAGTGTGAT GGTGCCTTAG TATCTGAAGA CATACTTTA AGTGAAGTGA CAGAGGGGTG	6780
	TATCTGTTGT GCAATGAAAG CAGATGTATC AGAACAGTTA CATCAATTAT ATTTAAAAGA	6840
	GCAAC̄CAGAC ATTGTATTTA TTGAATGTAG TGGGATTGCA GAACCGGTCT CTGTCTTAGA	6900
40	TGCTTGTTTA ACGCCTATTT TAGCTCCGTT TACAACAATT ACACATATGA TTGGTGTAAT	6960
	AGACGCAAGC ATGTATAAAC ACATTAAATC ATTCCCTAAA GACATCCAAG GCTTATTTTA	7020
45	TGAGCAATTA GCATATTGTT CTGTCTTATT TGTTAATAAA ATAGATTCAG CAGATGTTGA	7080
	AACAACGAGC AACTATTGA AAGATTTAGA AGTTATTAAC CCAGAGGCCG ATATACAAGT	7140
	CGGTATGCAT GGCAGCGTCA CTTTGCCAAT ATCAGTTAGA CAAATGACAG CAACTTCTGA	7200
50	CAATAAACAT AAGTCTTTAC ATCAAATGAT TAATCATCAA TTTGTGCAAT CACCAGTCAA	7260
	ATGTACTAAA GCAGAGTTTA TAAACGTTT AGCATGCCTT CCGTCTCATA TTTATAGGTT	7320
	GAAAGGGTTT ATGACATTTG AAGACACCGC ACATACGTAT CTCATTCAAT TTACACAAGG	7380

55

	CGGAAAGGGT ATTTCAAAAG AAGACTATCA ATGTTTGGAA CAGTAGTGTT TTCAGTGGAA	7500
	GAGAATGGTT AACATGCCTT CATGTATAAT AACGAGTTGA TTTGAACGTT TAAGCGTAAA	7560
5	TAAAAATAAG CTTGGTCAGC CATCAAATAT AATTTGAAAA CTGTCCAAGC TGTTTTATTA	7620
	GAGAACAAATC AATTAAACCCC ACATATTAA TAATACATCA GCAAAGCCTT CAGGTTTTTG	7680
10	AATATAACCT AAGTGACCGC CTGGAATATC TACAATAGGT ATGCCAGTTT CTTTATTTAT	7740
	ATAAAAGTTA ACATCTGTG GGAAGGAGCC TCTAGAATCT GTCCCATTTA GTAGGGTGAT	7800
	TTTATCGCTG TATTTTGTGA AATCATCCAA AGTAATATCT GAATGCGTAT ATTGTCTAAT	7860
15	TTCAAATTCT GACCAGAACA TCGTACGTTT GTACTGTCTT ATACGTCCTT CTTCAGTATC	7920
	AGCAGGTTGA GACATCATTT TTGCATCAAT TGGTGCGATA TTTAATGTTT CGCCAAATGT	7980
	TTTCATGCCT TTTTCTAAGC CTTCTGTAA AATTTGATGC ACAATGTCAT CATTTTTATC	8040
20	TTTCCAATAA GTACTGTCTG GTAAAAATGT ATTAATTGGT GGTTCGTGAA ATGCAATCTT	8100
	TTTAACGACT TCAGGGTAAT CTTTAAACAC ATGCATCGCA ACGATTGAAC CTGAACTTGA	8160
	ACCTAATATA TAGACAGGTT CATCACTTAA TGACTTTGCA AGTTCGGCAA TGTCTGTGC	8220
25	GTCCGCGTTG ACACGATAAT CACTGTCAGG GTTTGAAGCG GAATCAGGGA GTGGTTCAGT	8280
	TAACTCGCTT TCTCCATAAT CACGACGATC AACGGCTACA ACAGTAAAT GGTCTTTTAA	8340
30	CTGTTCTGCA AGAGGCAGAA AAATGTCTCC GGTACCGTTT GCACCAGGAA TAAAGATGAG	8400
	CACGGGTCCT TGTCGCACTT GGTGGTATCG TAATTTAGCG CCTTGTAATT CTAAAGTTTC	8460
	CATATTCAAT GACCTCCATT TGTTAATTGT TAGGTGATAA ACCTAATAAT TTAGCACCAT	8520
35	TTGTATAACT TATTTTCTCT TTTTCTTCAT CTGTTAAACC CAGTTCATCT AAAAATACAC	8580
	CTAATTTTTTC AGGCTCAATA TATGGATAAT CAGCAGCATA AAGAATTCTA TCAATACCTA	8640
	CTTCTTTCTT GACTAAATCA AACTGTGGCT TCGTTAACAT GCCACTCGGT GTGATATAAA	8700
40	AATTATTTTT AAAGTAATAG CTTACAGGGT GGTTCAAATG TTCAGCGAAT AAAGCTTCAT	8760
	CCATACGTTT TAAGAAGAAT GGGATAAACT CACCCCAATG TCCAATAATC ATATTTAACT	8820
	TTGGATAACG ATCAAAAATA CCAGATAATA CTAGATGTAT TGTATGAATG CCGACATCAA	8880
45	TGTGCCAACC ATAACCAAAA CAAGCAAATG TTGCCGCGAGT TACTTCAGGA TAATTTCCCTT	8940
	TATAGTATGA TTGATAAATG TCACTGTTAA CTGGCGCGGG ATGTAGATAA ATCGGTACGT	9000
50	CTAAATTTTC AGCTGTTTTG AAAATAATGT CATATTTGTC TTGATCAAGA AAACCATCTT	9060
	GTGCACGTCC CATAATGAGC GCACCTTTGA ATCCTAAATC ATTGATGCAA CGTTCGAATT	9120
	CTCGCGCTGC GGCTTCAGGC TCATTGATAG GTAAAGTTGC AAAGCCTACA AAGCGATTGG	9180

55

	TCTGACCAAC CAAATTTGAA GGAGAACCAT TTCCATAAGA TAAGACTTGA ATTTGAACGT	9300
	CTTGATTATT CATAAATTGG ATACGTTTCAT CATGATGTGA TAATTCGTCTG GCATTTGTAA	9360
5	AACCTGTCTT TTTTcAAGG CCTTCTAACA TTACTTTTCAT CGGTACACCT TTAGGATCTG	9420
	CTGATATCGC ATTCATCGTT TCTTTTTGAA TATCTTCAAT GACATAATGT TCTTCAAACG	9480
10	TAATACTTTT CATTTACTTC GCCTCCATAT TGTATTGCAT GTTTATTGCA TCTATTGCAG	9540
	AAGCATTTTT TATATACCTC TAATTTCAAT GTTTGTAACA TAAAATTGAT CTACCAAGGC	9600
	ATCTCTCCAT CGCCATTAAT AAATGTACCT GTTGGGCCAT CTGCACCAAT CGTTGCTAAT	9660
15	TGAATGATTG GCTTGATTCC TTCAGAAACG TGTTTGAAT TATTACTAAA ATCACCACCT	9720
	AAATCAGTAT TTGTAGCGCC TGGATCAGCA GCATTGATTT GCATGTTAGG TAATCCTTTA	9780
	GCGTATTGTA GCGTTAGCAT TGTTACTGCC GATTTAGACG AACAATAAGC TAATGAATTC	9840
20	ACTTTAGATT CAGCTGTTTC GGGGTTTGTA ACCATTCCAA ATGAACCTAA ACCACTTGAT	9900
	ACGTTGACGA CAACAGGTTG TTCAGATTTT TCTAAGAGAG GGACGAATGT ATTCATCATT	9960
	CGTACGATAC CGAATACATT CGTTTGATAT ACTTCTTCAA CGTCACGAGG TGTCAATTTG	10020
25	GAAGGTGCTG AAAATTGACC AGATATACCT GCATTGTTAA TGAGGATATC AAGACGGCCT	10080
	TCTTTTTCAG CAATCATGTT ATAAGCATTT TTGACTGAGT AGTCACITGT AACATCTAAT	10140
30	TGTACATAAT GAACACCTAA TTTTGTGAT GCTTGTGTC CTCTTACATC ATTCCGAGAA	10200
	CCTATATAAA CTTTGTAACC CAATGCTTTA AGTGCCTCTG CACTTGCATA GCCTAACCTT	10260
	TTATTGCCTC CTGTGATTAA CACAATTTTA GTCATTACGT CCCACCTCAT CTAAATAAAT	10320
35	GTTTAATAAA TAATTTCTGT ACGCTTCAAT TGAAATATGG CGATGCTCTA TTTGGAAGGC	10380
	AAATACACTA GTTGATAATG ATTGCAACAG CATATCTGTT TTGAATCGT GTAAGTGTCTG	10440
	TCATGCTTTT TAAATAAGTC ATAATAAAAA TCAAATAATT CTTGATAAAA TCGCCTTTGG	10500
40	TAAAAACGTA ATTTATTGTT GCCTGCTTCA ATACATTGCA GTAGTGCCTT ATTATCGATT	10560
	TTAAATTGTA AAAGATAATC TAACGACACT TGCATAACCT CATAATTAGA ATGATAGTCA	10620
45	TCTTTAATTT GCTTAAATG AGTGATAAAA ATATCAAGGT CTCTTTGTAT GACGTAGTAG	10680
	CATAAATCGC TTTTATCTTT GAAATGTCTG TACAATGTCC CCATACCGAT ACCTAGTTCT	10740
	TTAGCAATAC GATTCATACT AATGTTTTCA ACGCCTTCTT CATCAAAAAG TTTGTGCGCT	10800
50	ATTTCTTCAA TTCGTTGCCT ATTCTCTTTT GCATCTTTTC GCATGATTAC ACCTACTTAA	10860
	AATTCTCTAA AATTGACAAA CGGATAACTC TCCGTTTATT ATAAAACGTG TTAAGAAAGT	10920
55	TAGCAATGAA TTTGCAATAA CTATTAAATA TCATAAAAGA AAAGAGTGTT GATAATGTCT	10980

ACCTTATCGG TTCAAATGAT TGCTGAAAAA CTGAATGTCA CTACAGAAGA TGTGGAAAAA 11100
 GTATTAGCTA TGACAGCGCC ACTAGGCATT TTTAGTCATC AATTACAACG ATTTATTCAT 11160
 5 TTAGTATGGG ATGTCAGAGA TGTAATAAAC GACAATATTA AAGGAAATGG ACAAACACCA 11220
 GAACCATATA CGTATTTAAA AGGTGAAAAA GAGGACTATT GGTTTTTAAG A 11271

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6261 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

20 CAACCCGTTT AGAACAAAAT AAAAACCGTA CAATTTTATC ATCTTAATGA TTATTGTACG 60
 GAAAACTTT TTTACATCAT ATCTGCATGT GCATAATCGA TATCGGTAAA TTTATTATAT 120
 TGTTCATAA AATGTAAGTT AACTGTGCCT GTTGGACCGT TACGTTGCTT AGCAATGATA 180
 25 ATTTCAATTT CACCGTTTTT ATCATTCGTT TGTGGCTCGA AACCACCATC ATCGTCATCA 240
 TCTTCATCGC CGCCACGGTT ATAGTAATCA TCACGGTATA AGAATGCAAC GATATCGGCA 300
 TCTTGCTCAA TCGAACCAGA TTCACGAATA TCACTCATCA TTGGACGTTT ATCTTGTCGT 360
 TGTTCAACAC CACGAGATAA CTGACTTAAT GCGATACTG GACATTTTAA TTCACGGGCT 420
 AATGCTTTTA ATGTACGAGA GATTTTCAGAA ACTTCCTGTT GTCTGTTATC GGACGCACGT 480
 35 GAACCACTAC CTTGAATCAA CTGTAAGTAG TCAATCACAA TCATGTCTAA GCCATGTTCT 540
 TGCTTTAATC GACGACATTT AGAACGTAAA TCATTAATTC GAATACCCGG TGTATCATCA 600
 ATAAAAATCT TCGTACGTGA TAATTTACCT ACCGCTATAG TAAAACGACT CCAATCTTCC 660
 40 TCAGTCATAG TACCCGTTCT TAAGCGGTTT GAGTCAACAT TTCCAGAACT ACAAATCATA 720
 CGTGTGGCTA ACTGATCAGC ACCCATCTCT AGCGAGAAAA TACCAACTGT ATACATATCT 780
 TCATGCGTTG CAACTTTTTG TGCAATATTA AGTGCGAACG CAGTCTTACC TACAGATGGA 840
 45 CGCGCTGCAA GGATAATTAA ATCATTTCCG TTGAACCCTG CTGTCATTG GTCTAAATCT 900
 CGATATCCTG TAGGTATACC TGGTGTTTGA CCACTATTTT GATCAAGCTC TTCAGCTGTT 960
 TCATACACTT GTCCTAAGAC GTCTCGAATG TCTTTAAAGC CATCGCTTTC ACGAGAAGAT 1020
 50 GATAGCTCTA AAATTCGACG TTCTGCATCA CTTAAAATCG CATCTAGTTC AAGTTCATCA 1080
 TTATATCCAT CATTTGGCAAT ACTATCTGCA GTTTGAATCA ATCTACGTTT TAATGCATGC 1140

	TCTGCAAGAT ATTGCGGGCC ACCCGCTTcA TTCAACGTAC CTTCCGTCGA TAATTGATCC	1260
	ATCAATGTTA CAACATCAAT TTCTTTATTA TCTTCATTTA AGTGCATCAT TGCACGGAAA	1320
5	ATATGTTGAT GGGCACCCT ATAAAACGAC TCAGGAAGCA AAACCTCCTG AGTAGTATTA	1380
	ATCAATTCTG GATCTATAAT AATTGAACCT AAGACAGACT GTTCAGCTTC ATTGTTATGC	1440
10	GGCATTGAT TTTGCTCATA CATTCTATCC ATGAATGGTT ACACCTCTTA TTCAATCCA	1500
	ACTTTATTGT TCAACTGTGT GTACGCGAAT TGTACCTTCA ACTTCTTTAT CTAATTTAAC	1560
	AGGTACATTC GTATATCCTA GGAATGAAT TCCATTTGGT AAATCCATTT TACGTTTATC	1620
15	AATTTTAATA TCATGTTGTG CTTTTAGTGC TTCGGCAATT TGTGTTGTAC TTACTGACCC	1680
	AAACAATTTA CCACCTTCAC CAGTTTTTGC TGaTACTTCA ACTTCAATGT TTGATAACGT	1740
	TTCTTTTAAT GCTTTAgCAT CTTCAATTC TTGTTGGCGT TCTGTTTTG CACGTTTTTT	1800
20	CTGTAACCTC AATTGTTTAA GGTTACCTGG TGTGCTTCT ACAGCATAAT TCTTTTTCAA	1860
	TAAGAAGTTA TTTGCATAAC CTACTGGTAC TTCTTTAACT TCACCTTTTT TACCTTTACC	1920
	TTTACCTTTA ACATCTGTG TAAAAATTAC TTTCATGCAT CTTCACTCCT ACTTAATTGT	1980
25	TCTGTAATTG CTTGTTGTAA TTGTGCTATC GCCTCTTCGA CTGTCACACC TTTAAGTTGT	2040
	GTTGCCGCAT TGGTTAAATG TCCACCGCCA CCAAGTGCTT CCATTGTTAA CTGGACATTT	2100
30	ACTGAACCGA GTGAACGCGC AGATATACCA ATCAGATTAT CTTACGTCT CGCAACAACA	2160
	TATGATGCTT CAATACCTTC TAAACTTAAC AGTTCATCTG CTGCTTGTGC AACTGTTACT	2220
	GGATGATAAA TTTTATCGTC TGAACCATGC GcAATGGCTA TGCCATTATC TTCAACTTTT	2280
35	ACAGTTCGAA TTAATTCAGA TCGATTAATG TAAGTATCCA CATCATCTTT TAAGAAATGT	2340
	TGCGTTAAAA TCGTATCTGC ACCATGTGCA CGTAAATAAC TCGCTGCATC GAATGTTCTT	2400
	GATCCTGTTC GTAATGTAAA GTTTCTTGTA TCTACAATAA TACCTGCATA CATCACTGTT	2460
40	GATTCAAGAC GTGTTAAACG TTGTTCTGTT GGTGATATT CCAGTAACTC TGTTACCAAT	2520
	TCAGTGTGCG AACTTGCGTA TGGTTCCATA TATATCAACA ATGGATTAGA GATGAAGCTT	2580
45	TCACCACGTC TATGATGATC GATAACAACCT TTACGGTTTG CTTTATTTAA GACATTTTCA	2640
	TCTAAAACCA GTTCCGTTT ATGCGTATCA ACAATCACTA CGGTTGTCTT AGATGTCATC	2700
	ATATCCCAAG CATCATCTGA TGTAAATAAT CGCTCTCTTA ACTCTGGCTT TTTATCTATT	2760
50	TCGTTTCATCA CGCGTCGTAA TGTGGATCA ATGTCAGTCT CATTTAATAC GATGTATGCT	2820
	TCTAAATTAT TCATCATGTC AAATCTAGAC ACACCGATTG CTGCACCAAT TGCATCTAAG	2880
55	TCAGGACGTT TATGTCCCAT GATAATGACT TTGTACCCT CTGCAAGGAT ATCTTTTAAC	2940

	CCATAGAAAC GCACATTACC ATTAATACTT TTAATTGCAA CTTGGTCGCC ACCGCGTCCT	3060
	AATGCTAAGT CTAGGCCTGA TTGTGATAAT TCACCTAAGT CGATTAAATT TTCAGTACCT	3120
5	TCACCAACAC CGATACTTAA TGTTAATTGG GCACGATAAC CAACACTTTT TTCACGTAAT	3180
	TGACTCAAGA TATCAAATTT AGATTCTTCT AAGTCAGCTA ATATTTTTTG ATTTAAATAG	3240
10	GCTACGAATT GATCGGAACT GTATCTTTTG AAAAATATAT TATACTCAGT TGCCCATCGA	3300
	CTAATGACAC GCGTTACCAT TGAGTTGATT TCCGAACGCT GCGTATCATT CATATTTTGC	3360
	GTAATCTCAT CGTAGTTATC TAAAAATAAT GTCGCAATGA TTGGTTTAGA ATTTTCATAT	3420
15	AGTTCATTTG TTTGTACTTG TTCAGTTATA TCAAAGAAAT AGAGGCAGTG ATCATTCTCA	3480
	GAATAACGTA CTTGGAAATG ATACTGATTA TATTCTATTT CAACGGATTT CACTCTATCT	3540
	AATTGCTTTA AAATGTTTGG AAATACTTCA TTTACAGATT CAGAAATGAC ATTCGCTTCC	3600
20	ATATGATCTG TCATAAATTG GTTAACCCAT TCGATGTGAT CATTTTCATC TAAAACAATG	3660
	ATACCAATTG GTAAATGTTT GATTGCTTTA TTATTTGTTG TTGAAATTTG AGCACTCAAA	3720
	CCATCTACAT AACTATCCAT TTTCAATAAA GCTTGTCTGA ATAAATGAT GCTAACAATA	3780
25	ATCATCACGA CAAGAACGAT AGATGCAATT AGTGCTATAA GACTATTAAA GATAAACCAT	3840
	ACACCCATTA AAACAATTGC TGTGATGATC ATGATGACAA ATGGTATTAG TAAAGCTTTC	3900
30	TTAGTGGA CTGCGATT CAT TATTCCACCT CTATTCCTT TTTAGAATTA TTTTTCATGA	3960
	TTGCTTCAA ATTCAAATTT AAATCGATAA CACCAAGTAG TCCTACAATA TGTGTCGTAG	4020
	GTGTCAGTAT TGTACCGATA ACCAATAGTA AAATCGTTAC TGCATTCCGC AAACCTTTTCG	4080
35	CTTTACCAAA GAAATGAATA ACACTTAAAC CTTGAATATA CATTACTAAT GATAACACAA	4140
	GTTGGAAGTT TAAAAGAATG CTCTGGAACA CACTCGGTTG ACCTGTAAAT AATAAACATA	4200
	TGATAACAAT AATGTATATC CATAATAAAA TACCGCTCAT TTGCCACGCG AAAAGTGGCT	4260
40	TAAATACAGG TGTAGCGATT TTAAATTTTC GTAAATCGG AAATGTAACG ATTAAGTTAA	4320
	TTAAGACGAT TAAAAATGTA ATGATAATGA TGAAACCTGG TAATTGAACG GTCGCTTGTC	4380
	TAAACCCTTC TTCTAATATT TGGGTCATAT TCGCATCGGC ACCGCTCATC GTAATCGCTT	4440
45	CATGTAATGT TTGCTTGAAA GGTTTTACTA TGCTCGCTGA TGGTGGAATC CTTCCGAATG	4500
	TTTGTAGTAA CATAAAAGCG ATTAATGAAA TTTArCTCAT CGCTACTGTT GTTACGTATA	4560
50	ATATTCTTTC TTTAGACGTT CTTTCTTTGA GCAATTGACC AATAATTAAA CTTGCAATTA	4620
	AGACTAATAT GATGGCACTT AAAACGAAAG TATTACCTAA AACAGTTGTT ATAATTACTG	4680
	TAATAAGTGC ACTAATCCCG AAAGATTGTA TTGATTTATT CCATAAAACG ATACCTGGTA	4740

55

	CAAATACCAA CGCAATCGTT GCAATTATTG TTGCTTTAGG TTGTATTTT GAAAACACAT	4860
	AAGCCACTCC CATATTTTAA ACTATAGCTA TTATTTTAACT CTCTTTAATG AAAATTAACA	4920
5	ATTTATAGAT TGTATGCTTC TATTTTCATTT AATTGAATAA TAACTTTCAT GTTTTATAAG	4980
	TAATTAACAT ACTCATTGTA ATCGCTTTTG TGTGCTTTCA TTTTCAACAT GATTATTTAA	5040
	TCCCCTACA TAGCAATCAA GCTTGATTGA GATTTACAAT ACATTTCCAC TCTCATGTAC	5100
10	TCTAGATGTT TTTGAATATG ATAAGTGTGA TTTAGTGGCT TCATTCTTTG AAAATATATA	5160
	TTATTACTTA CGCTTAAAT GCTTTAAATT TAAGAAATGA TATAAGTTAG GTGCCCAGGT	5220
15	ACTAAAGTTT AGTAGGATC CATCATGCCC AACATTATCA GGCACGAAGA AATGACGATG	5280
	ATATTTAAAA CGTTCACCTA ATGCACGAAC TTGATCATCC GGATATAGCA AATCATCTAT	5340
	GAACCCCATC GTTAACACTT TTGTTTCTAA ATTTTAAAA ACATGCGTTA CGTCTGTGCG	5400
20	ACCTCGGTCA ATGTTGTGAC TATCCAATAC ATCTAGCAGT GTCAGATAAC AATTCAAATC	5460
	AAAATGTTCT TTAAATTTAT TACCTTGATG TTGTTGGTAT GCGACTACTT CATCCGGCGT	5520
	AAAACGTTCA TCATACTTT TTGATGATCG ATATGTCAA AAACCTAATT GGCGTGCAAT	5580
25	ACTTAGACCT TCCTTACCAC CAAGATGAAT GGCTTGCCTT GCAATTTTCAT TGAAAGCTCT	5640
	ACTATAAGAT GATGTTGAC TTGTTGCAGC AAGGATAATG GCTTTATCTA CTTCAAACCTG	5700
	TTGATTGTAG AGTAGTTCCA TTGCTTGCAT ACCTCCAAGA CTTCCCCCTA TTAAATATT	5760
30	AATCTTATCA TAACCAAGGG CTTGTATACC TCGTTCATTC GCTCTGACTA TATCTCTTAA	5820
	TGTTAATTTT TTAGGAAAT GAGGGTCGTT TAAAGGTGAA CTTGAACCGA AAGGACTACC	5880
35	AATAACATCA AATGTTAAAA ATTGATAATC GTGAATGGGT ATATATCCCC CATCAATAAT	5940
	TTCTCGCCAC CAACCCGGAT AATCATCTGT TCCATATGTT AAATGATTGC CAGTTAATGC	6000
	ATGACAACT ACAACTAATG GTTGTCCATG ATAACCGACA TGCTCATATC TCAAACGCAA	6060
40	GTAATCTATG ACTTCCCCAG ATTCTGTAAT AAATCCCCCT AAATTTAAAG TATCTACTGT	6120
	GTAATTTGTC ATTGTTCTTT CCTCCTTAAA CAAAAAACT TCTCACCCTA TTGAAAAGTA	6180
	AGAAGTCTTT ATACTTATCA TTCGAGTAAC TCGTTGGTTT TAGCACCGTG CTATAAAGTC	6240
45	GGTTGCTGAA GTATCACAGG G	6261

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- | | |
|----|-----------------------------|
| 50 | (A) LENGTH: 1222 base pairs |
| | (B) TYPE: nucleic acid |
| | (C) STRANDEDNESS: double |
| | (D) TOPOLOGY: linear |

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

5 ATGCGATTAA CTCTGGAAAT ATCTTTTCCA TATTTACGTh TTAAATTATT CAGCAAATTC 60
 ATACGAGaTT CATACTCGTT yAACTTGT TCGTCGAATT CTGTATTAGC CATTTCATCA 120
 TATAACTCAT GTTTTGCATC TTCTAAAATG TAGTAAAATT GATCAATATC TTCTTTTAAT 180
 TTGTCATATT TGTITGGAAC TATATCGTTT ATTGTTAACA AATGGTTGCT TAGTTCATAT 240
 10 AAACGATCAG TGATAGCATT TTCATCCGTT AATGTCATAT ATGCGTTATT AAGCGCTAAG 300
 CTTAATTTTT CAGAGTTTTG AATGCGTTTA ATATCTATTT CAAGTTGCTC TATTCGCCT 360
 15 TCTTTTAGAT GTGCTTCAGA CAATTCTTCT AATTGGAATT TCATTAAATC TAAACGCTGT 420
 AGCAATGCTT GGTCTGCTGA TTCTAAATCT TCTAACTCTT GCTTTTTGGC TTTATAATTT 480
 TGAAAAGTTT GGTGATATTT ATCCAACAAA TCTTGATAAC GTGATTCTGC GTAATTATCC 540
 20 AATAATGTTA AATGGTATTT TTGTTTCAAC AAAGACTGCG TTTTCATGTTG GCCATGAATA 600
 TCTAATAATT CTTCGATAAC TTTTCGTAAA TCTTGTAAG TAACTGTTTG ATTATTAATT 660
 TTACAAAGAC TTTTACCAGA GCTGAAAATT TCCCGTTTAA CTAATAAAAA ATCTTCATCT 720
 25 ACATCAATAT CCATATTTTT CAATATATGT ATAGCATCTT TACTCTCGTC AATATCAAAT 780
 ATACCTTCGA TGACAGCCTT TTTTTCACCA TGTCTTACAA AATCAGATGA AGCTCTCATT 840
 CCAATTAATT GTCCAATTGC ATCTATAATA ATTGACTTAC CTGAACCCGT TTCACCACTT 900
 30 AAAACAGTTA AACCATCAGA AAATTGAATT TCTAATTCTT CAATAATAGC AAATTGCTTG 960
 ATTGATAAGG TTTGTAACAT AAATCATCG CATCCTTATA ACAAATTGAA AATTCTTGAC 1020
 TTGATTTTCA CACTTGCCCTC TTTGCTTCGA CAAATAATTA AACAAGTATC ATCACCACAA 1080
 35 ATTGTGCCTA GTACTTCTTC CCAATTGATT TGGTCTAATA TAGCTCCAAT AGATTGTGCA 1140
 TTACeAGGTA TGTTTTTAGA ACAAGTAAAT TATCAGTACC ATCTATATTA ACAAAGGAAT 1200
 40 CCATTAAATA ACGTCCCAAT TT 1222

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1021 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

55 TTTGTTATTA TTACnTnAAA TAATTGCATT ACTTTTTACT GATGGTACAA CTTTCCATCC 60

TTCTTTTGGC ACGACATAAT TGTCTTTATC TTGAACTAAA TATCCGCCAG ATACTGAAAC 180
 AAACTCTTCT TCGTTACTGT CTATAGTCAT ATCAATTTCT AATAATCTTA CATTCTTCTT 240
 5 TTGTTTTTAAA ATATCTAATG CTTTCATCTGT AAATTTTGGT GCAATAATGA CTTCCAAAAA 300
 GATACTATGC AATTGCTCTG CTAATCAGG TGTTACAGCT CGGTTTAATG CAACAATTCC 360
 ACCAAATATT GATTGACTAT CCGCTTCATA CGCATGTTGA AATGCTTGTT CTATCGTGTC 420
 10 ACCGATACCA ACACCACATG GATTCATGTG TTAAACCGCA ACTGTAGCAG GTGTATCAAA 480
 CTTTTTAACT AAAGCTAGTG TAGCATCTGC ATCTTTAATA TTGTTATAGC TTAATTGTTT 540
 CCCATGTAAT TGTTTAGCGC CTGCAATCGT GTGCTTAGCA TTCGAAGTTC TCACAAAATA 600
 CGCTGATTGT TGTGGATTTT CTCCATATCT TAAAGTTTCT TTATCCCCTT TAAAGAAACG 660
 TACAATCGCT TCATCATATT CTGCAGTATG CTCAAAACT TTAATCATT AATGATTGTCT 720
 20 ATATGACTCA TCTAACGAAT CGTTTCTTAA TCGCGTCAAT ACTTCTTGAT AATCTGCCGG 780
 ATGTACAATT GTTGTTACAT GTTTATAGTT TTTAGCTGCA GCACGTAACA TTGTTGGACC 840
 ACCAATATCA ATATTTTCAA TTGCTTCGTC CATCGTCACA TCAGGGTTTG CAACAGTTTG 900
 25 TTGGAATGGA TATAAATTAA CTACTACCAT ATCAATTAAA TCTATATGTT GTTCTGATAA 960
 TTCATTTAAA TGCTGCGGTT TATTTGATC AGCTAAAATG CCACCATGAA CAGCCGGATG 1020
 T 1021

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3759 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TCATTCCTC CTAAATTGTT ATTACACTAT TACACaTAGC TAATCATCAA TGTGAAATCA 60
 CCTTCAAAGA CACTATCCAA ATCTTCAGAA GTCAAAATAA AGTTTGTTACC AGTAGTCAGT 120
 45 TTGAAAATTT CACCATCGAC AATCATTGTC CCTTCGCCCTT CCAACACTGT AACTAAACAG 180
 AACTCTCTAG GCTTCATATA ATTTAACGTG CCAGAAATTT CCCATTTAAC CAATGTAAAG 240
 AAATCATTG ATACAATGTG TGTACACTTA TGGTTTTCAA TAATTTGCT TTCAGGCAAA 300
 50 ATATTAGGTA ATGGTGCAAT GTACTGAATA ACGTCTAAAAG CTTTTTCAAT ATTTAACGGT 360
 CTATCATTAT ATTGATTATC TTGACGATTG AAATCATAAA GTCTATATGT AATGTCTGAC 420

	ATAAAAtAGa ATTCyCCAGG kTTTACTtTA AtatATCyAA gTatCGaCtC tATCGTTCCG	540
	TGTTGAACAT GATTGCAAC TTCTTCTCTA GACTCTGCTA ATGTCCctAT AACTATTTCT	600
5	GCATCTTCTT CTGCATCTAT AATATACCAA CATTcAGATT TGCCATATTG CCCgTTTTCA	660
	TGCTCATAAG CATAAGAATT ATCAGGGTGC ACATGAATAG AAAGTGATTc TCTTGcATCC	720
	ACTATTTTAG TTAGAAGCGG AAAATCTTTG CTTGGGAAAT CACCAAACAA TTCACGATGT	780
10	TCTGACCAAA TACGGTCTAA TGTTTGACCT TGATATGGTC CATTaATAAT CTCGCTCGTA	840
	CCATTtGGAT GTGCTGACAC ACACCAACAT TCCCCAGTT GTATCATTGT CTAATTGATA	900
15	TCCAAACTCA CTTAGACGTT GACCGCCCCA TAATTTTGTT TTTAAAATTG GTTGTAaaaa	960
	TAATGGCATT GTTGcACCTC CATTGTGATT AAGTAAGCAA TAGAACTCTG ATGTTGTGTG	1020
	TCCATTATAT TTTGATTTTG TTCTCATTTA CATCGTATTA TTAACtTCCA CATTTCAAAT	1080
20	TAActATTAG TGATTGTACC ATATTtACTA ACATTGCAGT ACTGCCAATT AAAAGnGCTT	1140
	CACTTAAATT TACAGTACTT TAACATTTTC AAAAAttTAT AGCATAGAGA TTATATCTCT	1200
	CTTACATTTG TACATATTTc CCTTTAAATT TACTCGCCCA TTATACCAAT TAATAaACAA	1260
25	CTTTAAAGT TGtGCCATAC ATTGTTCAAA TTCttTGTAa AACGCATAGA CAATACGTAC	1320
	TTATTCATAC TTATAATTCA TCATTTTCAA AAAATAACGA GTTACGAAAA AGTAACCCGC	1380
30	TTCAAATCAT ATTTACTATC CTTATTaATC CGTTTCATTT TCAAATTGAG TTAAAGCATC	1440
	TTTAATGTCC TGATCACCAC TAATAATTTG AAActCTTGG TGATTAAAAT GATTGGATGT	1500
	GACAATTTCT TTTAATACTG TCGCAACATC TTCTCTAGGA ATTTcACCTT TACCATCAAA	1560
35	ATATTGTGCA GCTTCTATCT TTCCAGATCC TGCTGCATTT GTAAGTGCCC CTGGATGTAA	1620
	AATTGTATAA TTCAAACCTG nAACGTCTTA AATAGTCATC AGCGTAATGT TTAGCTATTG	1680
	TATATGGCTT TAAATCACCG CTATCATCAA AAGCCTGACG TCTCGAATCA TATGTTGAAA	1740
40	CCATGACATA GTGTTTAATA TTGGCCTCTT TACTCGCAAT CATTGATTTA ACAGCACCAT	1800
	CTAAATCGAC AATAATTGTT TTATCTGCAC CCGTGTtCCC TCCAGAACCT ACTGAAAAGA	1860
	TAActTTATC GAATGGTTTA AACGTCTCAG TTAAAGTCTC TATTGAATCA TTTTCAACAT	1920
45	CAACAAGAAT TGCTTTcATA CCTGTGATT TTAACGCATT AAGTTGATCT GATTGCCTAA	1980
	CACCAGCAGT AAATGGTACA TTTTCTTTTG CTAATTGTTG CACTAGTAAC GAACCTACAC	2040
50	CGCCATTAGC ACCTATAACC AAAATATTCA TTTACAACAC TCTCCTATkT ATTATTCTCT	2100
	ATGCCATACC ACTTTATGAG ATATGTAAAA CTTGTTACAA CTATAAAAAAT CAATTGACAT	2160
	ACTACTGGGA ACGTATTAAA TTAATATATG AACAAATATT CATATGAAAG GATTGTCATA	2220

55

tCaAGGCATT AGcGATTACA ATCGAATACG TATCaTGGAA TTGTTATCaG TCAGCGAAgC 2340
 AAGTGTGGT CACATTtCAC ATCAATTGAA TTTATCTCAA TCAAATGTCT CGCACCAATT 2400
 5 AAAATTACTT AAAAGTGTGC ATCTTGTGAA AGCAAAACGA CAAGGCCAAT CAATGATTTA 2460
 TTCATTAGAT GACATCCACG TAGCAACTAT GTTAAAGCAA GCCATACATC ACGCGAATCA 2520
 10 TCCTAAAGAA AGTGGGTTAT AATATGTCTC ATTACATCA TCATCATGAC CATATGCATA 2580
 GTCATGTAAC TACAAATAAT AAGAAAGTAT TGTTTATATC GTTTTAAATA ATCGGTCTAT 2640
 ATATGTTTAT CGAAATCATC GGCGGTCTCC TTGCTAACAG CTTGGCATTa CTATCTGACG 2700
 15 GTATCCATAT GTTTAGCGAC ACATTCTCAT TAGGTGTTGC ACTTGTCGCA TTTATTTATG 2760
 CTGAAAAGAA TGCCACAACt ACAAAAACAT TTGGTTATAA ACGTTTCGAA GTACTCGCAG 2820
 CGTTATTTAA CGGTGTAACG CTTTTTGTA TAAGTATTTT GATTGTTTTT GAAGCGATTA 2880
 20 AACGTTTCTT TGTTCTTCT GAAGTTCAAT CAAAAGAAAT GTTAATCATT AGTATTATCG 2940
 GTTTAATTGT CAATATCGTT GTTGCAATTCT TTATGTTTAA AGGCGGCGAC ACTTCACACA 3000
 ATTTAAATAT GCGTGGTGCT TTTCTACATG TTATCGGAGA CTTATTAGGT TCAGTTGGCG 3060
 25 CCATTACTGC AGCTAkTTTA ATTTGGGCAT TTGGATGGAC AATCGCCGAT CCTATCGCAA 3120
 GTATTTTAGT TTCCGTTATT ATTTTAAAA GTGCTTGGGG TATCACAAAA TCTTCAATTA 3180
 ACATTTTAAT GGaAGGCACA CCAAGTGATG TTGATATAGA TGAAGTTATA ACTACTATTA 3240
 30 AAAAGGATTC ACGAATACAA AGTGTGCATG ATTGCCATGT TTGGACAATT TCAAATGATA 3300
 TGAATGCATT GAGTTGTCAT GTTGTGTAG ACCATACATT GACAAAGAAA GAATGTGAAT 3360
 35 TATTATTAGA AAaCATTGAG CATGATTTAT TACATTTAAA TATTCACCAT ATGACTATTC 3420
 AATTAGAAAC GCCTAATCAC AAACATGATG AATCGATTAT ATGTTCAGGA ACACATAGTC 3480
 ATTCACATAA CCATCATGCT CATCATCAG CGCATGTACA TTAATAATTT TAACCTACTG 3540
 40 CCATTGCATC GATTAACTT TTCAATGGCA GTAGGTTTTT TATGCTTTA TGGCGACTTG 3600
 TTTGGTCTTT GATGATGCAA TGTTTATTAA CAAATTTCA ACTATTATTT CTTACATTAG 3660
 TCATATTTTT GACAATTTAC TATTATAATT CTCTAACTTT AGTCACTTTA ATTAATTTTT 3720
 45 ATTAGATATT AATATGAAA TAACGTGTTT TTTGTTATT 3759

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13086 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

	TAATTATCGC GCATAACAAA ACATTAGCAG GACAATTATA TAGTGAGTTT AAAGAATTTT	60
5	TTCTTGAAAA CAGGGTGGAA TACTTTGTAA GTtACTATGA TTATTATCAn CCAGAGGCAT	120
	ACGTACCGTC TACTGACACT TTTATTGAAA nAGATGCCTC AATCAnTGAT GAAATTGATC	180
	AACTACGACA TTCTGCTACA AGTGCATTAT TTGAACGCGA TGATGTAATT ATTATTGCTA	240
10	GTGTAAGTTG TATATATGGT TTAGGTAATC CTGAAGAATA TAAAGATTTA GTAGTAAGTG	300
	TTCGAGTTGG TATGGAAATG GATAGAAGTG AATTACTTAG AAAACTTGTC AGATGTGCAA	360
	TATACACGAA ATGACATCgA TTTcCAACGA GGAACGTTTC GAGTGCCTGG TGATGTAGTG	420
15	GAAATATTCC CAGCCTCTAA AGAAGAACTT TGTATAAGGG TTGAGTTTTT CGGCGATGAG	480
	ATTGACCGTA TCCGAGAAGT TAACTACCTA ACAGGTGAAG TGTGAAAGA AAGAGAACAT	540
20	TTTGCGATAT TCCCAGCTTC TCACTTCGTA ACACGTGAAG AAAAGTTGAA AGTTGCGATT	600
	GAACGTATTG AAAAAGAATT GGAAGAACGA TTGAAAGAAT TACGAGATGA GAATAAATTA	660
	CTAGAAGCGC AAAGGTTAGA ACAGCGTACC AACTATGATT TAGAAATGAT GCGAGAGATG	720
25	GGATTCTGTT CAGGAATTGA AACTATTCC GTACATTTAA CTTTGCACC ACTGGGTTCC	780
	ACACCATATA CTTTATTGGA TTACTTTGGC GATGATTGGT TAGTAATGAT TGATGAATCA	840
	CATGTGACAT TACCGCAAGT TCGAGGCATG TATAACGGAG ACAGAGCGCG TAAACAAGTT	900
30	TTGGTGGATC ATGGGTTTAG ATTACCGAGT GCATTAGATA ACCGTCCACT TAAATTTGAA	960
	GAATTTGAAG mAAAGACAAA ACAACTTGTC TATGTATCTG CAACGCCTGG ACCATACGAA	1020
	ATTGAACATA CGGATAAGAT GGTGAACAA ATTATTGTC CTACTGGTTT ACTGGATCCT	1080
35	AAGATTGAGG TTAGACCTAC TGAAAATCAA ATTGACGATT TATTAAGTGA AATTCAAACA	1140
	AGAGTgAGCG TAATGAACGC GTACTTGTTA CAACGCTCAC TAAAAAGATG AGTGAAGATT	1200
	aACCACATAC ATGAAAGaAG CGGGTATTAA aGTLAATTAT CTGCATTGAG AAATCAAGAC	1260
40	ATTAGAACGA ATTGAAATAA TTAGAGACTT ACGAATGGGT ACATATGATG TTATCGTAGG	1320
	TATTAATTTA TTAAGAGAGG GTATTGATAT ACCAGAAGTT TCTCTAGTTG TCATATTAGA	1380
45	TGCAGATAAA GAAGGGTTTT TACGTTCTAA CCGCTCATT ATTCAAaCAA TAGGTAGAgC	1440
	TGCGCGTAAC GATAAaGGTG AAGTCATTAT GTATGCCGAT AAAATGACTG ATTCGATGAA	1500
	GTATGCAATT GATGAGACAC AACGTCGTCG AGAAATACAG ATGAAACATA ATGAAAAACA	1560
50	TGGTATTACA CCTAAACAA TTAATAAAAA AATACATGAT TTAATTAGTG CTACTGTTGA	1620
	AAATGACGAA AATAATGACA AAGCACAAAC TGTGATACCT AAGAAGATGA CGAAAAAAGA	1680

55

	TTTCGAGAAA GCTACAGAAT TAAGAGATAT GTTATTTGAA TTAAAAGCAG AAGGGTGACA	1800
	AGTAAATGAA AGAACCATCC ATAGTAGTAA AAGGTGCTCG TGCGCATAAC TTGAAAGATA	1860
5	TTGATATCGA ACTACCTAAA AaTAAATTAA TTGTTATGAC AGGTTTATCT GGGTCAGGTA	1920
	AATCGTCATT AGCATTGAT ACTATATATG CTGAAGGACA ACGACGTTAT GTTGAATCAT	1980
	TAAGTGCCTA TGCGCGTCAA TTTTtagGCC AAATGGACAA ACCAGATGTT GATACAATTG	2040
10	AAGGATTATC GCCAGCAATT TCAATAGATC AAAAAACAAC AAGTAAAAAT CCAAGATCAA	2100
	CTGTAGCAAC AGTAACAGAA ATATATGATT ATATACGTTT GTTATATGCA CGTGTTGGTA	2160
15	AACCTTACTG TCCAAATCAC AATATAGAAA TTGAATCGCA AACAGTACAA CAAATGGTTG	2220
	ACCGCATTAT GGAATTAGAG GCACGTACAA AGATTCAATT ATTAGCACCT GTCATCGCTC	2280
	ATCGTAAAGG TAGTCATGAA AAGCTAATCG AAGATATTGG TAAAAAAGGT TATGTACGTT	2340
20	TAAGAATCGA TGGCGAAATT GTTGATGTAA ATGATGTACC TACTTTAGAT AAGAACAAGA	2400
	ATCATACAAT AGAAGTTGTT GTAGACCGAT TAGTTGTTAA AGATGGAATT GAAACACGAC	2460
	TAGCTGACTC TATAGAAACT GCCTTAGAGC TTTCAGAAGG ACAATTAACA GTCGATGTCA	2520
25	TTGACGGGGA AGACCTTAAG TTTTCAGAAA GCCATGCTTG TCCTATATGT GGATTTTCAA	2580
	TCGGAGAGTT AGAACCAAGA ATGTTTAGCT TTAACAGTCC TTTTGGTGCT TGTCCGACAT	2640
	GTGATGGCTT AGGCCAAAAG TTAACAGTCG ATGTAGACTT GGTGTGTTCCC GACAAAGATA	2700
30	AGACGCTAAA CGAAGGTGCA ATAGAACCTT GGATACCGAC GAGTTCTGAT TTTTATCCAA	2760
	CATTGTAAAA ACGTGTGTTGT GAAGTTTATA AAATCAATAT GGATAAACCT TTTAAAAAGT	2820
	TAACAGAACG TCAACGTGAT ATTTTATTGT ATGGTTCTGG TGACAAAGAA ATTGAATTTA	2880
35	CATTTACACA ACGTCAAGGT GGTACTAGAA AACGAACAAT GGTTTTCGAG GGTGTAGTTC	2940
	CTAATATAAG TAGACGATTC CATGAATCTC CTTCAGAATA TACACGTGAA ATGATGAGTA	3000
40	AATATATGAC TGAACCTACCT TGCGAAACTT GTCATGGAAA GCGATTGAGT CGTGAAGCKT	3060
	TATCTGTTTA TGtagGTGGT TTAAATATTG GTGAAGTAGT CGAATATTCA ATCAGTCAAG	3120
	CGCTGAAC TAATAAAAAC ATTGATTGTT CAGAACAAGA TCAAGCGATT GCAAATCAAA	3180
45	TATTGAAAGA AATTATTTCC CGACTCACTT TTTTAAATAA TGTGGGACTT GAATATTTAA	3240
	CGTTAAACAG AGCTTCAGGT ACACTTTTCAG GTGGTGAAGC ACAACGTATT CGATTAGCAA	3300
	CGCAAATTGG GTCGCGTTTG ACTGGTGTCT TATATGTATT AGATGAGCCA TCAATTGGAC	3360
50	TGCATCAAAG AGATAATGAT CGATTAATTA ATACACTTAA AGAAATGAGA GATTTAGGAA	3420
	ATACTTTAAT TGtagTTGAA CACGATGATG ATACAATGCC TGCGGCTGAT TACTTAGTGG	3480

55

	AGGTAATGAA AGATAAAAAA TCATTAAACAG GACAATACTT GAGTGGTAAG AAACGTATTG	3600
	AAGTACCTGA ATATCGCAGA CCGGCTTCAG ATCGTAAAT TTCTATACGT GGAGCTAGAA	3660
5	GCAACAATCT TAAAGGGGTT GATGTGGACA TACCACTATC AATCATGACG GTTGTTACAG	3720
	GTGTATCAGG TTCTGGTAAA AGCTCATTAG TAAATGAAGT ATTATACAAA TCATTAGCTC	3780
	AAAAAATTAA TAAATCTAAA GTAAAGCCAG GATTGTACGA TAAGATTGAA GGTATTGATC	3840
10	AACCTGATAA AATTATTGAT ATTGATCAAT CACCAATAGG TAGAACGCCA CGCTCTAATC	3900
	CAGCAACATA TACTGGTGTG TTTGATGATA TACGTGATGT GTTTGCGCAA ACAAATGAAG	3960
	CTAAAATTCG AGGATATCAA AAAGGGCGTT TTAGTTTTAA TGTAAGAGGT GGACGCTGTG	4020
15	AAGcTTGTAA AGGTGACGGT ATTATTAAAA TTGAAATGCA TTTTTTACCT GATGTTTATG	4080
	TTCCCTGTGA AGTGTGTGAT GGTAAACGAT ATAATCGTGA GACACTAGAG GTTACTTACA	4140
20	AAGGTAAAAA TATTGCTGAC ATTTTAGAAA TGACTGTTGA AGAAGCAACA CAATTTTTTG	4200
	AAAATATTCC TAAGATTAAG CGCAAGTTAC AAACACTAGT TGATGTTGGT CTTGGATACG	4260
	TCACATTAGG TCAACAAGCT ACAACGTTAT CAGGTGGTGA GGCTCAACGT GTGAaACTTG	4320
25	CATCTGAACT TCATAAACGT TCAACTGGTA AATCTATTTA TATCCTAGAT GAACCGACAA	4380
	CAGGGTTACA TGTGACGAT ATTAGTAGAT TATTAAAAGT ATTAAACCGA TTAGTTGAAA	4440
	ATGGTGATAC TGTGTGAATT ATTGAACATA ACCTAGATGT TATCAAAACA GCAGACTATA	4500
30	TTATAGACTT AGGTCCTGAA GGTGGTAGTG GCGGTGGTAC TATTGTTGCG ACTGGCACAC	4560
	CCGAAGATAT TGCTCAGACA AAGTCATCAT ATACAGGAAA GTATTTAAAA GAAGTACTTG	4620
	AACGAGATAA ACAAATACT GAAGATAAAT AAGATTAAAA GAAGTGAAGG ATGTTATAAA	4680
35	TTTATCCTTC GCTTCTTTTT ATTAATTTAG TAATGAATAG TAGAAAGAAA AGATGCGTAA	4740
	AAAGaATTAT GTTAAGATAG GGTCAATCTA GAGTAGTTAA ACATAAATCG AACTGGGAGT	4800
	GGGACAGAAA TGATAAAGAA TCACTAATGA TTTATTATGT AGTGGTTCCT TGTCATTAGC	4860
40	CACAGCTATT GTGTACTTAA AAATAGGaat GCaTgAGTGC AACTCATGCA TAAGaAATAC	4920
	TAATTTCTAA AGAAAAAGTA TTTCTTTATG TTGGGGCCCC GCCAACTTGC ATTGTTTGTA	4980
45	GAATTTCTTT TCGAAATTCT TTATGTTGGG GCCCCGCCAA CTTGCATTGT TTGTAGAATT	5040
	TCTTTTCGAA ATTCTTTATG TTGGGGCCCC GCCAACTAAT TCCAATATAT CATTGTAGAG	5100
	CTTAGGTCAT TGATTTTTGG CTCGGACTTT TATGGCGATA TGAACCATGT AAATTAAGCA	5160
50	AGCAATAAAT TAATGATTGA TATTGACTTG TAAATAATA ACAATAATGA ACAATTAATA	5220
	TTTATTTTAG CTTTCAATG TAGATTGGTG TTATATTTTT GATATGATAA GAAGAGATGT	5280

	ACATTAAAGT TAGATTTAAT CGCTGGTGAA GAAGGACTAT CGAAGCCAAT TAAAAATGCT	5400
	GATATATCAA GACCGGGCTT AGAGATGGCA GGTtATTTTt CACATTATGC GTCAGATAGA	5460
5	ATACAACTAT TAGGAACAAC GGAACtATCG TTTTACAATT TATTACCAGA TAAGGATCGC	5520
	GCAGGTGTA TGCCTAAACT ATGCAGACCA GAAACGCCTG CAATTATTGT GACACGTGGA	5580
	TTGCAGCCAC CAGAAGAATT AGTTGAAGCT GCAAAAGAAT TAAATACCCC ACTTATAGTT	5640
10	GCTAAAGATG CgACTACAAG TTtAATGAGT CGCTTAACAA CGTTTTTtAGA GCATGCACTT	5700
	GCAAAGACGA CATCTTTTACA TGGTGTtTTTA GTAGATGTtT ACGGTGTtTG TGTACTAATT	5760
	ACCGGTGATT CAGGAATAGG TAAAAGTGAG ACTGCGTtTG AATTAGTTAA ACGTGGGCAT	5820
15	AGATTAGTAG CAGATGATAA TGtAGAAATA CGTCAAATTA ATAAAGATGA ACTAATAGGG	5880
	AAACCACCAA AGTTAATAGA ACATCTATTA GAAATACGTG GACTAGGTAT TATCAATGTT	5940
20	ATGACTTTAT TTGGCGCGGG TTCAATATTA ACTGAAAAAC GAATTAGATT AAATATTAAT	6000
	TTGGAAAACT GGAACAAGCA AAAGTTATAT GACCGCGTAG GTCTTAATGA AGAGACGCTA	6060
	AGTATTTTAG ATACTGAAAT CACTAAAAAA ACAATACCTG TAAGACCTGG TAGAAATGTT	6120
25	GCGGTAAATTA TTGAGGTCGC TGCAATGAAC TATCGATTAA ATATCATGGG CATTACACG	6180
	GCCGAAGAAT TTAGTGAAAG ATTAAATGAA GAAATTATCA AGAACAGTCA TAAGAGTGAG	6240
	GAGTAGGTTG AATGGGTATT GTATTTAACT ATATAGATCC TGTGGCATTt AACTTAGGAC	6300
30	CACTGAGTGT ACGATGGTAT GGAATTATCA TTGCTGTGCG AATATTACTT GGTtACTTTG	6360
	TTgCACAACG TGCACTAGTT AAAGCAGGAT TACATAAAGA TACTTTAGTA GATATtATTT	6420
	TTTATAGTGC ACTATTTGGA TTTATCGCGG CACGAATCTA TTTTGTGATT TTCCAATGGC	6480
35	CATATTACGC GGAAAATCCA AGTGAAATTA TTAAATATG GCATGGTGGA ATAGCAATAC	6540
	ATGGtGGTTT AATAGGTGGC TTTATTGCTG GTGTtATTGT ATGTAAAGTG AAAAATTtAA	6600
40	ACCCATTtCA AATTGGTGAT ATCGTTGCGC CAAGTATAAT TTTAGCGCAA GGAATTGGAC	6660
	GCTGGGGTAA CTTTATGAAT CACGAGGCAC ATGGTGGATC GGTGTCACGC GCTTTTTTAG	6720
	AACAATTACA TTTGCCTAAT TTTATAATAG AAAATATGTA TATTACGGC CAATATTATC	6780
45	ATCCAACATT CTTATATGAA TCCATTtGGG ATGTCGCTGG ATTTATTATC TTAGTTAATA	6840
	TTCGTAAACA TTTAAAATTA GGAGAAACAT TCTTTTTATA TTAACTTGG TATTCAATTG	6900
	GTCGATTCTT TATAGAAGGA TTACGTACAG ATAGCTTAAT GCTCACAAGT AATATTAGAG	6960
50	TTGCACAATt AGTATCAATt CTTTTAATTT TAATAAGTAT AAGTTTAATT GTATATAGAA	7020
	GGATTAAGTA TAATCCACCG TTGTATAGCA AAGTTGGGGC GCTTCCATGG CCAACAAAAA	7080

55

	TTATGGCGTG TATACCGTCT TGTTAAATTT TCGAAAGTTT TTAAGAATGT AATTATCATT	7200
	GAATTTTTCGA AATTTATTCC AAGTATGGTA CTGAAAAGAC ATATATATAA ACAACTTTTA	7260
5	AATATTAATA TCGGTAATCA ATCGTCGATA GCTTATAAAG TAATGTTAGA TATTTTITAC	7320
	CCAGAACTGA TTACGATTGG TAGTAACAGT GTTATTGGTT ACAATGTAAC AATTTTGACG	7380
	CATGAAGCAT TAGTTGATGA ATTTTCGTTAT GGACCAGTGA CGATAGGATC TAACACTTTG	7440
10	ATTGGTGCAA ATGCTACCAT TTTACCCGGT ATAACGATTG GTGACAATGT AAAAGTTGCA	7500
	GCTGGTACGG TTGTTTCAAA AGATATACCG GATAATGGAT TTGCATATGG CAACCCTATG	7560
	TATATAAAAA TGATTAGGAG GTGACAATTT TATGGCGCAA AAGAATAATA ATGTAATTCC	7620
15	AATGACTTTT GATGATGCAT TTTATCGTAA AATGGCTAAA CAGAAGTTTA AACAAAGAGA	7680
	ATATAAACGA GCTGCTGAAT ACTTTGAAAA AGTGTTAGAA TTGTCACCTG ATGATCTGGA	7740
20	AATTCAAATT GATTATGCAC AATGTCTAGT GCAACTTGGT ATTGCTAAAA AAGCAGAACA	7800
	TTTATTTTAT GACAATATTA TTTATAATAG GCATCTAGAA GATAGCTTTT ATGAATTGAG	7860
	TCAGCTCAAC ATTGAAGTTA ACGAACCAAA CAAGGCATTG TTGTTTGGTA TTAATTATGT	7920
25	TATTGTTAGC GACGACCAAG ATTATAGAGA TGAATTAGAT CAAATGTTTG ATGTGAAATA	7980
	TCAAAGTGAA GAACAAATTG AACTTGAAGC TCAATTGTTT GTAGTTCAA TACTATTCCA	8040
	ATATCTTTTT TCTCAAGGTC GATTAAAAGA TGCAAAGAAT TATGTCTTAC ATCAACCACA	8100
30	AGAAGTTCAA GATCATCGTG TAGTACGTAA TTTATTGGCA ATGTGTTATT TATATCTCGG	8160
	TGAATATGAT ACgGCTAAAG CATTGTACGA aGCACtATTA CAAGAGGATA GTACaGATAT	8220
	ATATGCATTA TGCCATTATA CTTTGCTACT TTATAACACT AAGGAAAATG AACAAATATCA	8280
35	AAAATATTTA AAAATATTAA ACAAAGTTGT ACCTATGAAT GACGATGAAA GTTTTAAATT	8340
	AGGTATTGTA TTAAGTTATT TAAAGCAGTA TCGTGCATCA CAACAATTGT TGTACCCTTT	8400
	ATATAAAAAA GGGAAATTTT TATCAATTCA AATGTACAAT GCTTTAGCAT ATAATTATTA	8460
40	TTATTAGGT GAAGAAGACG AAAGTCATTA CTACTGGGAT AAATTGAAGC AAATTTCTAA	8520
	AGTGGAATTT GGACATGCGC CTTGGGTAAT TGAAAATAGC AAAGAAGTTT TTGACCAACA	8580
45	TATTTTGCCA TTAATTCAA GTGATGACAG TCATTATCGT TTATATGGTA TTTTTTTATT	8640
	GGATCAATTA AATGGTAAAG AAATTGTGAT GACGGAAAGT ATTTGGCAGG TTTTGAAAA	8700
	TCTAAATAAT TATGAGAAAT TGTATTTAAC GTATTTAGTT CAAGGTTTAA CGCTCAATAA	8760
50	ATTAGACTTC ATTCATCGCG GCTTATTAAAC GCTTTACCAT AATGAATTAT TTGTAAGTGA	8820
	AAATGATGTA ATGTTTGCAT GGATTAATCA AGGTGAACTC ATAATTGCTG AAAAAGTAGA	8880

55

	TCGAAACGTT	ACAAAGAAGC	AAATTACAAC	ATGGTTAGGC	ATAACACAAT	ATAAACTGAA	9000
	CAAAATGATT	GAATTTCTCT	TGAGCATATA	GATTTATGAA	AAGTTAGATT	TATTATATAA	9060
5	TGCGCATAAT	GATTAATAAT	GAGGAGGCGT	TAATAAAATG	ACTGAAATAG	ATTTTGATAT	9120
	AGCAATTATC	GGTGCAGGTC	CAGCTGGTAT	GACTGCTGCA	GTATACGCAT	CACGTGCTAA	9180
	TTTAAAAACA	GTTATGATTG	AAAGAGGTAT	TCCAGGCGGT	CAAAATGGCTA	ATACAGAAGA	9240
10	AGTAGAGAAC	TTCCCTGGTT	TCGAAATGAT	TACAGGTCCA	GATTTATCTA	CAAAAATGTT	9300
	TGAACACGCT	AAAAAGTTTG	GTGCAGTTTA	TCAATATGGA	GATATTAAAT	CTGTAGAAGA	9360
	TAAAGGCGAA	TATAAAGTGA	TTAACTTTGG	TAATAAAGAA	TTAACAGCGA	AAGCGGTTAT	9420
15	TATTGCTACA	GGTGCAGAAT	ACAAGAAAAT	TGGTGTTCGG	GGTGAACAAG	AACCTGGTGG	9480
	ACGCGGTGTA	AGTTATTGTG	CAGTATGTGA	TGGTGCATTC	TTTAAAAATA	AACGCCTATT	9540
20	CGTTATCGGT	GGTGGTGATT	CAGCAGTAGA	AGAGGGAACA	TTCTTAACTA	AATTTGCTGA	9600
	CAAAGTAACA	ATCGTTCACC	GTCGTGATGA	GTTACGTGCA	CAGCGTATTT	TACAAGATAG	9660
	AGCATTCAAA	AATGATAAAA	TCGACTTTAT	TTGGAGTCAT	ACTTTGAAAT	CAATTAATGA	9720
25	AAAAGACGGC	AAAGTGGGTT	CTGTGACATT	AACGTCTACA	AAAGATGGTT	CAGAAGAAAC	9780
	ACACGAGGCT	GATGGTGTAT	TCATCTATAT	TGGTATGAAA	CCATTAACAG	CGCCATTAA	9840
	AGACTTAGGT	ATTACAAATG	ATGTTGGTTA	TATTGTAACA	AAAGATGATA	TGACAACATC	9900
30	AGTACCAGGT	ATTTTTGCAG	CAGGAGATGT	TCGCGACAAA	GGTTTACGCC	AAATTGTCAC	9960
	TGCTACTGGC	GATGGTAGTA	TTGCAGCGCA	AAGTGCAGCG	GAATATATTG	AACATTTAAA	10020
	CGATCAAGCT	TAATTCGAAG	TCGAATTAAG	ATGTTGAGCT	GTAAATTATT	TGGATATTTA	10080
35	TTTTAATAGT	GTCATCACAG	CGTTAAAATA	ATGTCTTACT	TTTAAATTAA	AGCAAATTAT	10140
	ATAGEAAACT	AGAACTTAGT	ACGTATCATT	TGTGCGTTTC	AATGAGTTCT	AGTTTTTTTA	10200
	TATGTTATAT	TAAACTTATA	ACTTTATGGG	AGTGGGACAG	AAATGATAAA	GAGCCACTAA	10260
40	TGATTTATTA	TGTAGTGGTT	CTTAAACATT	AGCCACAGCT	AATGTGTACT	TAAAAATAGG	10320
	AATACATGAG	TAAAACTCAT	GCATAAGAAA	TACTAATTTT	TATAGAAAAA	GTATTACTTT	10380
45	ATCGTTGTCC	CACCCCAACT	TGCACATTAT	TGTAAGCTGA	CTTCCGCCA	GCTTCTGTGT	10440
	TGGGGCCCCG	CCAACTTGCA	CATTATTGTA	AGCTGACTTT	TCGTCAGCTT	CTGTGTTGGG	10500
	GCCCCGCCAA	CTTGACACATT	ATTGTAAGCT	GACTTTTCGT	CAGCTTCTGT	GTTGGGGCCC	10560
50	CGCCAACTTG	CATTGTCTGT	AGAAATTGGG	AATCCAATTT	CTCTATGTTG	GGGCCCACAC	10620
	CCCAACTCGC	ATTGCCTGTA	GAATTTCTTT	TCGAAATTCT	CTGTGTTGGG	GCCCACACCC	10680

55

	ACTCGCATTG CCTGTAGAAT TTCTTTTCGA AATTCTCTGT GTTGGGGCCC CTGACTAGAG	10800
	TTGAAAAAAG CTTGTTGCAA GCGCATT TTCAGTCAA CTACTAGCAA TATAATATTA	10860
5	TAGACCCTAG GACATTGATT TATGTCCCAA GCTCCTTTTA AATGATGTAT ATTTT TAGAA	10920
	ATTTAATCTA GACATAGTTG GAAATAAATA TAAACATCG TTGCTTAATT TTGTCATAGA	10980
	ACATTTAAAT TAACATCATG AAATTCGTTT TGGCGGTGAA AAAATAATGG ATAATAATGA	11040
10	AAAAGAAAAA AGTAAAAGTG AACTATTAGT TGTAACAGGT TTATCTGGCG CAGGTAAATC	11100
	TTTGTTTATT CAATGTTTAG AAGACATGGG ATATTTTTGT GTAGATAATC TACCACCAGT	11160
	GTTATTGCCT AAATTTGTAG AGTTGATGGA ACAAGGAAAT CCATCCTTAA GAAAAGTGGC	11220
15	AATTGCAATT GATTTAAGAG GTAAGGAACT ATTTAATTCA TTAGTTGCAG TAGTGGATAA	11280
	AGTCAAAAGT GAAAGTGACG TCATCATTGA TGTTATGTTT TTAGAAGCAA GTACTGAAAA	11340
20	ATTAATTTCA AGATATAAGG AAACGCGTCG TGCACATCCT TTGATGGAAC AAGGTAAAAG	11400
	ATCGTTAATC AATGCAATTA ATGATGAGCG AGAGCATTG TCTCAAATTA GAAGTATAGC	11460
	TAATTTTGTT ATAGATACTA CAAAGTTATC ACCTAAAGAA TTAAAAGAAC GCATTTCGTCG	11520
25	ATACTATGAA GATGAAGAGT TTGAACTTT TACAATTAAT GTCACAAGTT TCGGTTTTAA	11580
	ACATGGGATT CAGATGGATG CAGATTTAGT ATTTGATGTA CGATTTTAC CAAATCCATA	11640
	TTATGTAGTA GATTTAAGAC CTTTAACAGG ATTAGATAAA GACGTTTATA ATTATGTTAT	11700
30	GAAATGGAAA GAGACGGAGA TTTTCTTTGA AAAATTAATC GATTTGTTAG ATTTTATGAT	11760
	ACCCGGGTAT AAAAAAGAAG GGAAATCTCA ATTAGTAATT GCCATCGGTT GTACGGGTGG	11820
	ACAACATCGA TCTGTAGCAT TAGCAGAACG ACTAGGTAAT TATCTAAATG AAGTATTTGA	11880
35	ATATAATGTT TATGTGCATC ATAGGGACGC ACATATTGAA AGTGGCGAGA AAAAATGAGA	11940
	CAAATAAAAG TTGTACTTAT CGGTGGTGGC ACTGGCTTAT CAGTTATGGC TAGGGGATTA	12000
40	AGAGAATTCC CAATTGATAT TACGGCGATT GTAACAGTTG CTGATAATGG TGGGAGTACA	12060
	GGGAAAATCa GAGATGAAAT GGATATACCA GCACCAGGAG ACATCAGAAA TGTGATTGCA	12120
	GCTTTAAGTG ATTCTGAGTC AGTTTTAAGC CAACTTTTTC AGTATCGCTT TGAAGAAAAT	12180
45	CAAATTAGCG GTCATCATT AGGTAATTTA TTAATCGCAG GTATGACTAA TATTACGAAT	12240
	GATTTTCGGAC ATGCCATTAA AGCATTAAGT AAAATTTTAA ATATTAAAGG TAGAGTCATT	12300
	CCATCTACAA ATACAAGTGT GCAATTAAAT GCTGTTATGG AAGATGGAGA AATTGTTTTT	12360
50	GGAGAAACAA ATATTCCTAA AAAACATAAA AAAATTGATC GTGTGTTTTT AGAACCTAAC	12420
	GATGTGCAAC CAATGGAAGA AGCAATCGAT GCTTTAAGGG AAGCAGATTT AATCGTTCTT	12480

55

GCGTTAATTC ATTCTGATGC GCCTAAGCTA TATGTTTCTA ATGTGATGAC GCAACCTGGG 12600
 GAAACAGATG GTTATAGCGT GAAAGATyAT ATCGATGCGA TTCATAGACA AGCTGGACAA 12660
 5 CCGTTTATTG ATTATGTCAT TTGTAGTACA CAAACTTTCA ATGCTCAAGT TTTGAAAAAA 12720
 TATGAAGAAA AACATTCTAA ACCAGTTGAA GTTAATAAGG CTGAACTTGA AAAAGAAAGC 12780
 ATAAATGTAA AAACATCTTC AAATTTAGTT GAAATTTCTG AAAATCATTT AGTAAGACAT 12840
 10 AATACTAAAG TGTTCGAC AATGATTTAT GACATAGCTT TAGAATTAAT TAGTACTATT 12900
 CCTTTCGTAC CAAGTGATAA ACGTAAATAA TATAGAACGT AATCATATTA TGATATGATA 12960
 ATAGAGCTGT GAAAAAATG AATATAGACA GTGGTTCTAA GGTGAATCAT GTTTTAAATA 13020
 15 AGAAAGGAAT GACTGTACGA TGAGCTTTGC ATCAGAAATG AAAAATGAAT TAACTAGAAT 13080
 AGACGT 13086

20 (2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1350 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

30 CATTAGTCAT GAAAATAGCC GACAACTTCA TCTGTGAAAT CACCGGCCCTT TTATTTTAGC 60
 TAACTTTATT TCTGATTTTA CGATTTTAAT TGATCATACA GAGAAAGTGA TCTTTTACA 120
 ATTTCTAAAA ACTCATGATC TATATTGGAC ATTTGATGAA AATAAGACAA AATGTTTTCT 180
 35 GTTAGCTTCT CTGTTTTTGG GAATGAATCA TCTTCTTTAA TCCAAATCGC TAATTCGCCT 240
 AATGGTGTTC TATCATCTTT AAATGTTTGT ATATATTCGT AAAAGCTCAT AGTATTCCTT 300
 CTCTCAATTT ACTTATATAA ATCCTACCAC GAAAGCTTTC AAGAAAACAC AATTAAATGT 360
 40 CTATTTAGTG AACTTTTTTA GGTGTGTCAC TCTTTTAAATG TCTGCCAATT AGGTCAATTA 420
 ATCATCACA TGTACAATTA ACTCTATTTT CAGTTCATAT ACTCACACAC CGTTTTTGAA 480
 45 CAACACATTA ACTTCTCATT TAGATAAAAC GCAAAAAGC CTGGCACCAA TACAATAGAT 540
 GCCAGACTAA GAGTCTACTA TATAAATTTA TTTAGCGTAT GGTTTTACTT CGATTGCACC 600
 TTCATTTTCA TCATGAACAC CATGCTTATA ATAATCAATA TATTGTGGCT CTAAAGGCTT 660
 50 TCTGCCACGT ATAATGTCTG CTGCTTTTTT AGCTAACATT AAAACAGGTG CGTGTATATT 720
 GCCATTTGTC GTACGTGGCA TAGCTGATGC ATCAACTACA CGTAAATTTT CCATACCGTG 780

55

ACTACAAGAT GGGTGTAAATG CTGTTTCACC ATCTCTACGA ACCCAATCAA GAATTTCTTC 900
 GTCTGTTTGC ACTTCTGGTC CTGGTGAAAT TTCTCCACCA TTGAATGGAT CCATTGCTTT 960
 5 TTGAGATAAG ATATTTCTTG CTACACGAAT TGCTTCTACC CATTCTTTT TATCTTCTTC 1020
 TGTGATAAA TAATTAAAGC GGATACTTGG TTTTTCGAAT GGATCTTTAG ATTTGATTTT 1080
 CAAGCTACCA CGAGAGTTTG AATACATTGG TCCTACGTGA ACTTGATAAC CATGTGCGAC 1140
 10 CGCTGCCTTT TGACCATCAT ATCTTACAGC TATTGGTAAG AAATGGAACA TTAAGTTAGG 1200
 ATAATCAACT TCGTTATTTG AACGTACAAA TCCGCCACCT TCAAAATGGT TAGATGCTGC 1260
 15 TGCACCTGTA CGTGTGAAAA TCCATTGTAA ACCAATAAAT GGcATGCGCT TGATATCTAA 1320
 GCTTGGCtGt AATGATACAG GTTCCTTACA 1350

(2) INFORMATION FOR SEQ ID NO: 18:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1376 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TAATGCTATT GGCAACACCA TATATGAAAn CTCCAAACGA TCCTAAACCG ACTATAGATT 60
 30 CACCAAATTT nACAATCCAT GAATAAGTA GTGGCCATAA GAATAACAAT ATGACAACTA 120
 AAAATGTACA GTAAAATGCA GTCATAATTG GAACTAGACG TTTACCACTA AAAAATGATA 180
 ATGCTAATGG TAATTCTGTT TCACTAAACT TATTGTATGC ATAAGCTGCT ATTAAACCTA 240
 35 TTACAATACC AACAAAGACA TTGCCATTAT TCATCTTTTC AAAAGCTGAA TTTATTTCCG 300
 ATGCTTTTCAT TCCTAATAAA GCGCTAATT TCATTGGTGA TAATACAACT GTAACATAAA 360
 40 AATATCCTAA CGTrGCTGCA rGCGsGACTG CACCATCATT TTTCTTTGCC ATTCCTATAG 420
 CTACACCAAT TGCAAATAAA ATACCTAATT GCTCTAAAAT CGTAGTACCT ACCGTAGTAA 480
 AGAACATTGC GATTTTCGGC GTCGCATGAA GTGCATTAA CGTATTACCA ATTCCGGCAA 540
 45 TAATTGCTGC AGCCGGTAAA ATGGCAACTG GTAACATTAA CGAACGCCCT AAATTTTGGA 600
 AAAATTTATA CATTGAATGT CATCCTTCTT AAAATAATGT AGAAATATAA AGATTACTAA 660
 TGTAAGTAGA ATAAGTACTT CGATACTCCG TTATAGTCAC CTAGGCTTAC TAACCAGCTA 720
 50 TATTTCTACC TCAAGTTATT TTATAAACTT TTTACAATTT CATGCAATTC TTGTGTAAAC 780
 TTTGCTGTTT GTGTTTCAAT CTCTTTTGTA ATATAATCGA TACGCTCGTT TCGTTTAAAA 840

55

AAAGACCGTG AATCTTAGTA GGACCAACAT AAGCAACAGG TAATATTGGT GACTTACTTA 960
 ACATTGCAAT TGTGGAAGCA CCACGTTTCA AAGGTGCACC TTCTTGCGAT GTGCGAGAAC 1020
 5 CTGTTGGGAA GATACCAACT GTCTTATTAT CTTTCAACAA ATTGATTGGG CGTTTTAAAG 1080
 TACTAGGTCC TGGATTTTCA CGATCTACAG GAAATGCATT TAAAGACGTT AAAAATTIAC 1140
 CAATCCATTT ATTTTGAAT AATTCTTTTT TAGCCATATA ATGAATTTGA TTAGGATATA 1200
 10 ATGCCATACC TAGCATAATG ACTTCGTTAT AACTTTCATG CGTACAAGTT ACGACATATT 1260
 TACTATCCTT AGGAATATTA TCTTTACCGA TTACGTATAA TGATTTTGAC ATTTTAACTA 1320
 AAATGAAATT CAAAATCTTA CTAATCACTG AATACATTGT GCCACCTACT TAACTT 1376
 15

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7363 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TTGTCATACC AATATTTTGT AAAATATGGA ACACAAGTAA AGTGACGAAA CCAACGATAA 60
 AGATTTTGTT AAATTGATCT TCAATTTTCG CAGCTAATCT TATTAGATGG AAGATTAAAA 120
 30 ATAAAAATAT TAAGATCAAT ATGACAGAAC CGATAAAGCC AAGTTCCTCT CCAATCACTG 180
 AAAAGATAAA GTCAGTATGA TTTTCAGGTA TATAAACTTC ACCGTGATTG TATCCTTTAC 240
 CTAGTAACTG TCCAGAACCG ATAGCTTTAA GTGATTCACT TAAATGaTAG CCATCACCAC 300
 35 TACTATATGT ATAGGGGTCA AGCCATGAAT TGATTGCTCC CATTGATAC AGTTGGaCAC 360
 CTAATAAATT TTCAATTAAT GCGGGTGCAT ATAGaATACC TAAATGACT GTCATTGCAC 420
 CAACaATACC TGTAATAAAG ATAGGTGCTA AGATACGCCA TGTTATACCA CTTACTAACA 480
 40 TCACACCTGC AATAATAGCA GCTAATACTA ATGTAGTTCC TAGGTCATTT TGCAGTAATA 540
 TTAATAACTT TGGTACTAAC GAGACACCAA TAATTTTGAA AAATAATAAC AAATCACTTT 600
 GGAATGATTT ATTGAATGTG AATTGATTAT GTCTAGAAAC GACACGCGCT AATGCTAAAA 660
 45 TTAATAAAT TTTCAATGAAT TCAGATGGCT GAATACTGAT AGGGCCAAAC GTGTACCAAC 720
 TTTTGGCACC ATTGATAATA GGTGTAATAG GTGACTCAGG AATAACGAGC AAGCCTATTA 780
 50 ATAATAGACA GATTAAGAAA TACAATAAAT ATGTATAATG TTTAATCTTT TTAGGTGAAA 840
 TAAACATGAT GATACCTGCA AAAATTGCAC CTAAAATGTA ATAAAAAATT TGTCTGATAC 900

55

	TTGCTAAAAC AGCTATAGTG GCTACTAATA CCCAGTCTAC TTTGCGAAnC aATGCTTATC	1020
	CGGCTGTTGA CGAGATGAAT AATTCATTGC AAACCTCCTT TATACTCACT AATGTTTATA	1080
5	TCAATTTTAC ATGACTTTTT AAAAATTAGC TAGAATATCA CAGTGATATC AGCTATAGAT	1140
	TTCAATTTGA ATTAGGAATA AAATAGAAGG GAATATTGTT CTGATTATAA ATGAATCAAC	1200
10	ATAGATAcAG ACACATAAGT CCTCGTTTTT AAAATGCAAA ATAGCATTAA AATGTGATAC	1260
	TATTAAGATT CAAAGATGCG AATAAATCAA TTAACAATAG GACyAAATCA ATATTAATTT	1320
	ATATTAAGGT AGCAAACCTT GATATATCAT TGGAGGAAAA CGAAATGACA AAAGAAAATA	1380
15	TTTGTATCGT TTTTGGAGGG AAAAGTGCAG AACACGAAGT ATCGATTCTG ACAGCACAAA	1440
	ATGTATTAAA TGCAATAGAT AAAGACAAAT ATCATGTTGA TATCATTTAT ATTACCAATG	1500
	ATGGTGATTG GAGAAAGCAA AATAATATTA CAGCTGAAAT TAAATCTACT GATGAGCTTC	1560
20	ATTTAGAAAA TGGAGAGGCG CTTGAGATTT CACAGCTATT GAAAGAAAGT AGTTCAGGAC	1620
	AACCATACGA TGCAGTATTC CCATTATTAC ATGGTCCTAA TGGTGAAGAT GGCACGATTC	1680
	AAGGGCTTTT TGAAGTTTTG GATGTACCAT ATGTAGGAAA TGGTGTATTG TCAGCTGCAA	1740
25	GTTCTATGGA CAAACTTGTA ATGAAACAAT TATTTGAACA TCGAGGGTTA CCACAGTTAC	1800
	CTTATATTAG TTTCTTACGT TCTGAATATG AAAAATATGA ACATAACATT TTAATAATTAG	1860
	TAAATGATAA ATTAAATTAC CCAGTCTTTG TTAAACCTGC TAACTTAGGG TCAAGTGTAG	1920
30	GTATCAGTAA ATGTAATAAT GAAGCGGAAC TTAAGAAGG TATTAAAGAA GCATTCCAAT	1980
	TTGACCGTAA GCTTGTATTA GAACAAGGCG TTAACGCACG TGAAATTGAA GTAGCAGTTT	2040
	TAGGAAATGA CTATCCTGAA GCGACATGGC CAGGTGAAGT CGTAAAAGAT GTCGCGTTTT	2100
35	ACGATTACAA ATCAAAATAT AAAGATGGTA AGGTTCAATT ACAAATTCCA GCTGACTTAG	2160
	ACGAAGATGT TCAATTAACG CTTAGAAATA TGGCATTAGA GGCATTCAAA GCGACAGATT	2220
40	GTTCTGGTTT AGTCCGTGCT GATTTCTTTG TAACAGAAGA CAACCAAATA TATATTAATG	2280
	AAACAAATGC AATGCCTGGA TTTACGGCTT TCAGTATGTA TCCAAAGTTA TGGGAAAATA	2340
	TGGGCTTATC TTATCCAGAA TTGATTACAA AACTTATCGA GCTTGCTAAA GAACGTCACC	2400
45	AGGATAAACA GAAAAATAA TACAAAATTG ACTAAGTCTG GTTGTATTATA TGATTAATGT	2460
	TACATTAAAG CAAATTCAAT CATGGATTCC TTGTGAAATT GAAGATCAAT TTTTAAATCA	2520
	AGAGATAAAT GGAGTCACAA TTGATTCACG AGCAATTTCT AAAAATATGT TATTTATACC	2580
50	ATTTAAAGGT GAAATGTTG ACGGTCATCG CTTTGTCTCT AAAGCATTAC AAGATGGTGC	2640
	TGGGGCTGCT TTTTATCAAA GAGGGACACC TATAGATGAA AATGTAAGCG GGCCTATTAT	2700

55

	AAACCCTAAA	GTAATTGCCG	TCACAGGGTC	TAATGGTAAA	ACAACGACTA	AAGATATGAT	2820
	TGAAAGTGTA	TTGCATACCG	AATTTAAAGT	TAAGAAAACG	CAAGGTAATT	ACAATAATGA	2880
5	AATTGGTTTA	CCTTTAACTA	TTTTGGAATT	AGATAATGAT	ACTGAAATAT	CAATATTGGA	2940
	GATGGGGATG	TCAGGTTTCC	ATGAAATTGA	ATTTCTGTCA	AACCTCGCTC	AACCAGATAT	3000
	TGCAGTTATA	ACTAATATTG	GTGAGTCACA	TATGCAAGAT	TTAGGTTTCG	GCGAGGGGAT	3060
10	TGCTAAAGCT	AAATCTGAAA	TTACAATAGG	TCTAAAAGAT	AATGGTACGT	TTATATATGA	3120
	TGGCGATGAA	CCATTATTGA	AACCACATGT	TAAAGAAGTT	GAAAATGCAA	AATGTATTAG	3180
	TATTGGTGTT	GCTACTGATA	ATGCATTAGT	TTGTTCTGTT	GATGATAGAG	ATACTACAGG	3240
15	TATTTCAATTT	ACGATTAATA	ATAAAGAACA	TTACGATCTG	CCAATATTAG	GAAAGCATAA	3300
	TATGAAAAAT	GCGACGATTG	CCATTGCGGT	TGGTCATGAA	TTAGGTTTGA	CATATAACAC	3360
	AATCTATCAA	AATTTAAAAA	ATGTCAGCTT	AACTGGTATG	CGTATGGAAC	AACATACATT	3420
20	AGAAAATGAT	ATTACTGTGA	TAAATGATGC	CTATAATGCA	AGTCCTACAA	GTATGAGAGC	3480
	AGCTATTGAT	ACACTGAGTA	CTTTGACAGG	GCGTCGCATT	CTAATTTTAG	GAGATGTTTT	3540
25	AGAATTAGGT	GAAAATAGCA	AAGAAATGCA	TATCGGTGTA	GGTAATTATT	TAGAAGAAAA	3600
	GCATATAGAT	GTGTTGTATA	CGTTTGGTAA	TGAAGCGAAG	TATATTTATG	ATTCGGGCCA	3660
	GCAACATGTC	GAAAAAGCAC	AACACTTCAA	TTCTAAAGAC	GATATGATAG	AAGTTTTAAT	3720
30	AAACGATTTA	AAAGCGCATG	ACCGTGTATT	AGTTAAAGGA	TCACGTGGTA	TGAAATTAGA	3780
	AGAAGTGGTA	AATGCTTTAA	TTTCATAGAG	ATTAGTCGAG	GGACCTTTTA	CTTATAAAAA	3840
	TGATTTGAAT	TAATACTAAA	AGATTACAAA	GAAGAGGTGG	TTTTGTGTGT	AAATACAAAA	3900
35	TTGCCTTTTT	CTTTTTATGT	TAAATCTATA	AATTTGAAAC	TAAATCAAGG	TTAATCTAT	3960
	GTACACACTT	TATATAGGAA	GTAGTTTGAA	TGTTTATATA	ATGTTTACAA	AAAAGATGTA	4020
	GTATTATAAT	GTCTAATTTT	ACATGTGTTT	CAGTAAAATT	TGTTGTGGAA	TGTTAACGAT	4080
40	ATACGTATTT	TATAAAAaAT	TTTTTATAAT	GATTATTCCA	ATGATGCGTA	ACGCTTACAT	4140
	CTTATCTAAT	GCTAGCTTTT	TGACAAAAAT	ATGACAATCA	ATTAATGTGA	TTCTAATAAA	4200
45	TATTCGCAAA	TTGCTTTTAT	GCGATTAAAT	TTTTTTGGTG	GTACTATATA	GAAGTTGATG	4260
	AAATATTAAT	GAACCTTATAT	GCAAAAGTAT	ATTGAGAAAT	AAACAGGTAA	AAAGGAGAAT	4320
	TATTTTGCAA	AATTTTAAAG	AACTAGGGAT	TTCGGATAAT	ACGGTTCAGT	CACTTGAATC	4380
50	AATGGGATTT	AAAGAGCCGA	CACCTATCCA	AAAAGACAGT	ATCCCTTATG	CGTTACAAGG	4440
	AATTGATATC	CTTGGGCAAG	CTCAAACCGG	TACAGGTAAA	ACAGGAGCAT	TCGGTATTCC	4500

55

EP 0 786 519 A2

	AGAATTGGCA ATGCAGGTAG CTGAACAATT AAGAGAATTT AGCCGTGGAC AAGGTGTCCA	4620
	AGTTGTTACT GTATTCGGTG GTATGCCTAT CGAACGCCAA ATTAAAGCCT TGAAAAAAGG	4680
5	CCCACAAATC GTAGTCGGAA CACCTGGGCG TGTTCATCGAC CATTTAAATC GTCGCACATT	4740
	AAAAACGGAC GGAATTCATA CTTTGATTTT AGATGAAGCT GATGAAATGA TGAATATGGG	4800
	ATTCATCGAT GATATGAGAT TTATTATGGA TAAAATTCCA GCAGTACAAC GTCAAACAAT	4860
10	GTTGTTCTCA GCTACAATGC CTAAAGCAAT CCAAGCTTTA GTACAACAAT TTATGAAATC	4920
	ACCAAAAATC ATTAAGACAA TGAATAATGA AATGTCTGAT CCACAAATCG AAGAATTCTA	4980
	TACAATTGTT AAAGAATTAG AGAAATTTGA TACATTTACA AATTTCTAG ATGTTTCATCA	5040
15	ACCTGAATTG GCAATCGTAT TCGGACGTAC AAAACGTCGT GTTGATGAAT TAACAAGTGC	5100
	TTTGATTTCT AAAGGATATA AAGCTGAAGG TTTACATGGT GATATTACAC AAGCGAAACg	5160
20	TTtAGAAGTA TtanAGAAAT TTAAAAATGA CCAAATTAAT ATTTTAGTCG CTACTGATGT	5220
	AGCAGCaAGA GGACTAGATA TTTCTGGTGT GAGTCATGTT TATAACTTTG ATATACCTCA	5280
	AGATACTGAA AGCTATACAC ACCGTATTGG TCGTACGGGT CGTGCTGGTA AAGAAGGTAT	5340
25	CGCTGTAACG TTTGTTAATC CAATCGAAAT GGATTATATC AGACAAATTG AAGATGCAAA	5400
	CGGTAGAAAA ATGAGTGCAy TcGTCCACCA CATCGTAAAG AAGTACTTCA AGCACGTGAA	5460
	GATGACATCA AAGAAAAAGT TGAAACTGG ATGTCTAAAG AGTCAGAATC ACGCTTGAAA	5520
30	CGCATTTCTA CAGAGTTGTT AAATGAATAT AACGATGTTG ATTTAGTTGC TGCACTTTTA	5580
	CAAGAGTTAG TAGAAGCAAA CGATGAAGTT GAAGTTCAAT TAACTTTTGA AAAACCATTA	5640
	TCTCGCAAAG GCCGTAACGG TAAACCAAGT GGTTCCTGTA ACAGAAATAG TAAGCGTGGT	5700
35	AATCCTAAAT TTGACAGTAA GAGTAAACGT TCAAAAGGAT ACTCAAGTAA GAAGAAAAGT	5760
	ACAâAAAAAAT TCGACCGTAA AGAGAAGAGC AGCGGTGGAA GCAGACCTAT GAAAGGTGCG	5820
	ACATTTGCTG ACCATCAAAA ATAATTTATA GATTAAGAGC TTAAAGATGT AATGTCTTGA	5880
40	GCTCTTTTTT GTTTTCAATA ATTGATTCTC TGTAATATC aAGTaCTAA CGTTTTAAAG	5940
	GTAAATATT TAATTGGATT GAGATCTGTA TGCGTTTATA TCaTTCTGTG TAAATATGGT	6000
	TCTCCACCAA ATGTGGTGAG TATATAATTT AAAGAACTAT TTTTAAATTA AGAATAATCG	6060
45	AACATAAATA AACTTTATGA AATTTTCAGTA TCATGTTCTT ATAAAAAACA ATAGGGCTTT	6120
	TTGctGACGC TAGTGCGCGA TAAATAATAA GTTGAATATA AAAAAGATCA CTGCCAATCA	6180
50	TTCGTTTAAT GGCAGCGATC TTTTTTATTT AATTATTTCT CTTTCCACTG CAACATTTGA	6240
	TAACCAATGC GTGGATGTGT TTTAATAATA TCTTTGCGT CCTCATGACA TTGTGAAAGT	6300

55

CCATATATTC GTTTTAATAT CATCTCATAA GTGAGTACTT TTCCTTTATG ATTTGACAAT 6420
 AGTTCTAACA AGCTAAATTC ATTGGCGTC AAATGTACCT CCTGATTATT AATAACAACA 6480
 5 GATTGGGAGC CAAAGTCGAT GCTTAGCAAA CCGTTAGTAA ATACAATGTT AGTTTCTTGA 6540
 TGTGACTTAG CGATTCTCTC GATGACTCGT ATTCGTGCCC GAAGCTCATC AACATTAAAA 6600
 GGTTTAGTCA TATAGTCATT CGCACCCTTA TCTAAAGCTT GAATAATTGT TTGTTCTTCT 6660
 10 TGTCTTGACAC TTATTACAAT GATAGGAATG TCAGTATGTT GCCTGATTTC TGAAATCAAA 6720
 CATAATCCAT CTTTATCTGG TAAACCTAAA TCTAATAAAA TGACATCTGG TTTATCAATT 6780
 TGAATTTTAA AGTGTGCTTG TGTGGCATTG TCGGCTGTAG TTACATTGTA ATAATCTAAA 6840
 15 GTTAATGCAA CATCAAGTAA ATGTGTGATT GCGTGATCAT CTTCAATTAT CAATATTTTA 6900
 GATTGCATTA TACGTCTCCT TCGTTAAAGT CTGTATATAT ATTGAAATAG AATATACTGC 6960
 CGTGTGGTTG GTTCGGTTTA TATTGTAAGT TTGATTGATG TTTGTGTAGG ATAGTCTGTA 7020
 20 CTAAATATAA GCCTAGTCCC ATGCTTTCTT TTTGGTTATC TTTAAATAT TTATTTGATC 7080
 CTGTGTAAAA AGGCTCGAAT ATCTTTTGTt GTTCTTCTAA ACTAATTCCA GGTCCTTCGT 7140
 25 CTATAACGGC AAATTCGATT TGTTCATAGC TAGCATAACG AATAGATAAA TTGATTTTGG 7200
 TGTCAGTAGA AGTGTGTTTA ACTGCATTTT CAATCAAATT GAAtAAAgCT TGTAAATCA 7260
 ACTTACTGTC AATGTGTATA AAcTGTAAT TTAAGGAGGA TGATACAGTT ATACGCTTTT 7320
 30 TTAAATGGCG ACGTTCTAAA ATACATATCG ATTTCTTATA CTA 7363

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGTH: 10470 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TTAACAATCG ATAACCACAA TACTTCTATT GTAATTGTTT AACGATTTCn CGATTAAAAAT 60
 45 CATCTAAATC GTCTGGTACT CGACTTGTTA CAATATTGTT GTCTACAcTa CTGACTCATC 120
 AACTACATGT GCGCCTGCAT TTGATAAATC TTTGCGTACA TTAAATACTG CTGTTAACGT 180
 ACGACCTTTT AAATCGTCTG TATCTATTAG TATTTGTGGC CCATGACAAA TGGCAAATGT 240
 50 TGGTACATCA TTTTtagTAA AGTATTTAGC AAATGTGCCA TATCGACCTT CTGTATCTCC 300
 ACGTAAATGA TCTGGTGAAA ATCCTCCAGG AATTAATAAT GCATCATAAT CTTCTGGTTT 360

55

	ATTTCAGTA TCTCCAATCA CTACAGTATT AAAGCCTGCA TTCTCTAATG CCTCTTTAGG	480
	GCTTGAATAT TCTATATCTT CAAATTCGTT TGCTAGAATA ATTGCTACTT TTTTAGTCAT	540
5	TGAAAATCAC CTTTCTATAT ATCATTGATA TAATTACTAT AGACAAGTAA ATCAGTGATT	600
	AAACATACAA GATATAAAAA ATATTAAGCG ACTGTCGCGA TATCTAACCC TAACACATCT	660
10	TATGTGGCAT TTA CTTAGAT ACTAATTTAA CCTTTTCTTC AAGCTGATCT AACAAATCCAA	720
	TCCATTCATC TATATCTTCA ACACGTACTT CATCAGGATT TACATGATCG ATATCCTCAA	780
	TAAACTTATT TAAACGCGCT TTTATCTGTT CGATTGTTTG CTGTTCAATC ATAAAAAGTT	840
15	AACTCCTTTT ATTTTGT TTTTTCATT ATTATCCTAA CAGAAATTGC GTTAAAGCGA	900
	TATAATCTTA GCTATATTTA TGACATTCAA ATTATTTTGA CTTTAAAAA TCCCCTTTTC	960
	AATTA ACTAA AATTAAGAGA TAATTTGTTA CGAGTGATAA TACGAaGkGG TaTCATACCG	1020
20	ATATGAACCA AATAGAAAGA AGGAAGTTTA AGACGATGAA TAGCGTCAAA TTGAAGCAAC	1080
	CTGTTAGCAT TTACAATGAT CCATGGGAAG TGAAATTTAT ATACATTTAA ATTTCATGAG	1140
	ACAATAAACG TTGATTTAAT GCGTTTTTTT GCCTTTTTTA TTTTCCTTAT TTTTCTGTT	1200
25	TTACAACAAA ATGGTATCAA AAATGGTATC ATTTGTAGTT ATTTTAGCTT CACATATTAA	1260
	AACAACCACA CTCCTAAATT AATAGGTGGT GTGGTTTTGT TGGTTGTGTG GGGATAAAAA	1320
	TAACCGCATC AGTTAAGATG CGGTTATCTA GCAAGGGCCA CGTATTTATA AATACGTTTA	1380
30	GAATCTCTTC GGCAACTTTG CTATAGACAG TCTATGCTGT TACTAAATTA TACCACCACA	1440
	CAAACTACT CCCATTCAGG AACACAGAGC TTTGTCGCTC GTCAGCAACG TCATATGAAT	1500
	TCTCAGTTCA TGTGTGTGTG ACACTTTAAA CGGTCTGTGC CAGTAGCGAC CGAGTCATTT	1560
35	CAAGAATGAC CATTTACAT TTATATTATA ACACTTGTCTG TCGTAACTG TATAGTTTTT	1620
	CAGTTGTATT TAAAGTTAAG TTATCTACTT CGCGCTTTCC TTGCCTTAAT TGTGAAATTA	1680
40	CATATTGCGC TACGCCAGTT TGTTTGTA GAA TTTGGTAACC TGTTATATCA CTTTGTATCA	1740
	ATTCAATTAT TTTTAATTTA TAATCACTCA TATTATCTAC GTCCATTCTT TTTATCTAAA	1800
	CAATAAAAAT GTGTCTTTCT CCCGATAAAT AATAACAATG GTAGGCTTAA TAAAAACAAT	1860
45	ATTAAATACA TTTGTTCTGT CATAATTGAA AACCTCCAAA TAATATTATA TTATATAAGT	1920
	GTAAGGAGGA GCCATCAGGC TCCAAGCATA ATGTTAATCT TTGTTGTTTG GCTTTCGGTC	1980
	TAGGTAGCCG AGATGCCaTT CTCTAAGTTG TTTTAACACT TCTGGAATTA TCAGTACTGC	2040
50	CAATACTTGA TGTTCTAGAA GTGTTTTTAT TATGTCTAGC ATGAGGCTTT TCACCTCCTT	2100
	ACACATAATT TGTAAGTCAT CAACTAACCT ACAAATATAA TTATACTAAA CAAATGTTTA	2160

55

	GTTATCTACA	TTTAAATCTT	GAGAGAAATG	TTAAAAAGTT	CTAGTAAAAT	AATAGCACAT	2280
	TTTATCTTTA	AATGTAAATA	GAAAGCAGGT	ATGTAACGCA	CCTGCTTAAA	TAGaCATGAC	2340
5	TATGTCATTG	TAAGTGATTT	CTCCCCATAA	GTCACTAAT	ATCTGATTAG	GTGGGGCAGA	2400
	ACCATTCCAT	GTTCTAATAG	GCAAGTAATA	ACGTTGCCCC	TCCCATGTAT	ATCCTACCCA	2460
	AACATGACCA	TCTTGTAACA	TCACTTCTGT	ATAATCACAA	TACCCACCAG	GTTGGAAGTG	2520
10	ATAACCCACT	GGACAAGATA	AGAATGGCCC	CACTTTTCTT	ACTGTGATTG	GTTGATTGCC	2580
	GTTTGTGAAT	CTAGCACTTT	CTTCCATGTA	GTAAGTACCA	TATTTATTAC	GTTTCCATGC	2640
15	ACTTGCAACT	GGTTTAACTG	TATTACTTGA	AGCGCTTGAC	TCATTAGAGA	CAGTGGCAAC	2700
	CGGTATTTTA	CCATCCATGT	ACGCCCTAAT	CTGCTTGATA	AAGTAGTCTT	TAAGTTGCAA	2760
	CCGCTTGTCT	TCTGGCAATA	GACCGCGAGT	TACTGGGTCA	AAACCAAGTG	GTAACCAAGT	2820
20	ACTTCTATGA	GGGCATGATG	TTGAAGTAAA	TTCAATGTGC	AATCTGATTG	TATTTCTGTT	2880
	TGCTGGTAAT	CCCCATTTTT	TCAACAATCT	AGCGCATTCT	TGAAAGTTG	CCTGTTTCTT	2940
	TTTAAAGAAT	GTGCGTTTAT	CTGCGCCCAT	TGATTGACAT	ACTTCAATAC	CGTAATAATA	3000
25	TTTATTACCT	ATTTGATTAG	CGGTATGCCA	ACCTACTTGT	GATTCATCTA	AGGCTTGCCA	3060
	AACTGTGTTG	CCTGATACGT	AACTATGCGC	AATGCCCGCT	TCTAATCTTG	ATAAAGGTGC	3120
	ATTTACTAAT	CCGTTACGAT	ATGCTTCAGC	AGTCGCCCCT	TTGCTCCCTG	CGTCGTTGTG	3180
30	TATAACTATA	CCTTTAGGGT	TACTACCACG	CTTAGGTAGG	TCATAACCTT	TAACCACATC	3240
	TTTGATGATT	TTAAGTTCTA	CTGCTTTAGG	TTGTGGCTTA	GCTGTTTCTT	TTTLAGGTGC	3300
	TTGTGTAGGA	GATTGAACTG	ATCGTGGCGC	TGTCTCACTT	TTAAATTCG	GACGGATAAA	3360
35	CCACATAGGG	AAATCATAAG	CATGTTGTCT	TCTTGTAAGT	TTTTCCCAAC	CCCAGCCGGG	3420
	TTGTTCGATT	CCGTCAGTCC	AGCCACCGCC	TAGCCAATTC	TGCTCATATA	CAATGATGTA	3480
40	ATCTAAAGTT	GCTTCAATTA	CCCATGCAAC	GTGACCATAT	CCAGCACCGT	AGTTGCTACC	3540
	GAATACCACC	ATGTCGCCAG	GTTGTGCTAA	GAAGTCCGGT	GTATTTTGGT	ATACAGTAGC	3600
	TAATCCGTCG	AAGTTGTTAG	CGAACGGAAT	ATCTTTTGCA	CCTAAACCTT	TTAGAAGTAA	3660
45	TCCAAACAAA	ACTTTCCAAC	CAGCATTTGGC	ATAATCAAAG	CATTGAAATC	CATACCATAA	3720
	GTCCACATTG	AATTGTTTTT	CCTCAGAAGT	TTTCAACCAC	TCTATAAACT	CATTTTTAGT	3780
	TAATTTTGCT	TGCATTGTCT	CCACCTCCAT	GATGATACTC	ATTCACATCA	AAGCCAACAT	3840
50	CGTTAGAGGC	GTCTGTGAAA	GGTTGTGATG	TATCATATTC	TTTTGGTGCT	TTCGCGCTTA	3900
	ATTCCGGCGT	TAAACTACTG	TCTTGTGATG	ATTCCACGT	AACCTGTTGT	TCTTCTTTTT	3960

55

	TTGGGTCAGT AATAACGCCA ATACCTGTAA GTAACGTGAG GATAGCGCCT ATAATTGCGC	4080
	TAGCTTGATT TAATTGAGTA GATAAATCTA ATCCGAATAA ATCCGTGACT TGCTTGATAA	4140
5	ATAGCAACAA TGCTCCAACCT AAACCAGTTA GTACTGCTTT GTTTTTGAAT CTCAATTTCC	4200
	AGTTAATATC CATTTGTTTG CTCCTTTTAT CCAAAATAAA AAAACGACTA AAAATTAGTC	4260
10	GTTTAAAAATT ATTCAATGGT CAATGTCGGA GATCCTGAAT AAACATCACT TATAGTGACG	4320
	TACAACATCC CTGAAGGATT ACTAAAGTTG ATATTTTTAC TTGCAACTCC GCTATTGACT	4380
	CCTGATATTC CTAAATCACT TGACCCTAAA TTAGTTTGCG AAATCCTCAT TATACCGCTA	4440
15	CGTACATTTT CTATTGTCAC CTGATAACTT TTATTGGGTT CAACTCCATT TATTGTCCAT	4500
	TTTGCTGTTG ATTCTTCTAT GCTATCCGGA TATTTATTTT TAGGTAAGGG TTTTATTACA	4560
	AAAGATGAAG GCTTTTTCCTA TACTTGGATA TTCCAGCAT ATACTTTTGT ATATTCTTCA	4620
20	CCTTCGTAAA TAAACTTCTT TACATTTTTA AAATTACCTT CCATAAAAAT CACCCTTTAA	4680
	TTAAATATAA CGTATTCGGG TCTTTTTGAT ATATATAGTT ATATTCATTT TCTGTTCCCTG	4740
	TCCAAATTTT AACCGTCGGT TGAGATGCGC TTTTGTAGTT ATATAAATTA TCCGCTTGTT	4800
25	GTTTAGTAAA AGCTTGAGAT GACAAAACAT ACCGCTCGTC ATGATTATGA TTTTTTGAG	4860
	CATATAAATC ATTTAGTGTT TGTTTGAATT CCTCAAAATC TTCTGTATTA ACTTTTGAGC	4920
	CAATCTGTTG CAATACACTT TCTGAAATAG AGTTGTTTTG TATTGCTTCT GCTAATTCTC	4980
30	TTAATGTGTT CATAGATTCA GGCGCGCTAT CAACTAGTTC AGCAATTTTT GTATCCGTAT	5040
	ACGTTTTAGA GTCGTTGAGA GTTGATCTTT TGATTTTTTC AACTTCTTGC AATTTATTTT	5100
	CTAACCCCTC AACATTTGCG ATATTGATTT TGTCCAATAA CTCAGGTTCT GCTTTGATAT	5160
35	CTGTATCTTT ACCATCAATT TGCCACATTT TAGTGTGAGG ATTGATTGAT ACTACAGTAC	5220
	CGTTTATACC GGGTGCGCCT TGTCTCCTT TTTTACCTGC TTCACCTTTT GCTCCAGGTT	5280
40	GTCCCGGTTT ACCTTTATCA CCTTTGCGAC CTTTAAATCT ACTTTCATTC TTTTCGATGT	5340
	AAGAAATGAC ATCTTTATCT ATTTTCTCTT TAAAGTCTTT GCTCAATAAA TCTGTGCGGT	5400
	TATCTTTTAA AATCTCGTA ATAGCATCAT CTACCAATTT AACATCGATT TCTTTTGCTA	5460
45	CAGCAGATTC AATACCACTA TCAACGATAT TGAAAGAAAA GTTTGCGACA TGTATTTTTT	5520
	CTTCTTCTTT CTCTAAAAAC AGCTTACAGC GAACATAACC AGCGTGTGTTG ATAACCTTTT	5580
	TAGGTATCTT GTAGGTAAGG AAACCTTTTA CAACATCGTC GATAATAAGG GGCTCATTTT	5640
50	TGAATATAGA GCCATCTTCC ATAAACAAAT GTAATCTAGG TGTTAAGCCA TGTGCTTTTA	5700
	GATCGATACG ACCTTGTTTG TCATTGATAC CTATTCTTAT AGATGCTGTA TTTTCATCTT	5760

	CAACATCTTT TATTTTGTAC ATTTACACAC CTCTTTATTT ATATTTATCC CTTGTGAAGT	5880
	AGATACCTTT TAAGCCGATT TGTTTATATA ACTTAGCGAT TGTACTTGCT TGATGTTGGC	5940
5	ACCACTCTAT AGCAGTAGCG TATTGGTGGG TAGCTGGATT CTTAGGATTC CATCTAATTC	6000
	GGTACAATGT GTTTTGACCT TTATTGATGT AATCCTTTCT TACGAAGCTA GCACCGCCCA	6060
10	TGATTGCTTT TGCTGGAGAT GTCCAACCTT TATTCCTTGC AAACGTCATT GCGTAGTTAG	6120
	GATTGTTGTC GTAAGCGCCA ATGCCGAAGT AGTTGTATAC TCCATCTTTT CCGTTAGCGA	6180
	AGTTACTTGT TCCATATCCA CTTTCTAAGA AAGCATGCGC GATTAAATAA ATTTCAATTA	6240
15	TGTTGTGCTT TTTACAAGCT TCTGCGAAGC CTTTACCTTG ATTATTCAAT GTTCCCTTAC	6300
	CTTTAAGTAT CTTATTAAGT GCGCTAACTG AAACACCTTG ATACTGCCT AAATTAAGCA	6360
	TTTGGTAGCA TTGTGTGTTA CTTTCCCAT A TACGCTTTAC ATTCATTGCT GAACTCGTTT	6420
20	GTGCTCGTGT AGCGTTAseC AACCCCAAGC ATTAGATTTT TTCGGGTAC CTCTTGCCAT	6480
	TTGTTTATCC AGTGCTTGTT TGAATGTATA AGGACTCGTT TCTGTTATGA TCTGCGGTTG	6540
	TTTAGATGCC GAACCATTGT TGGCTGTTGG TGACGAGTCT CTTACATTAG CTATATCAGC	6600
25	GTTTTTATTA TCTACCATAA CTTTATTCT AGATTTTGTT ACTGTTGGCT TAGTTATAGA	6660
	ATTTAATAAT TTTTCTCTGT TTTTAAATAT ATTAAGTAAT GCCTTTTCTA ATGCTTCGTA	6720
	TTTATCTTTA GGAGGAACAC CGTTGTCAAT CATATTCCAA TTAACATGTT CCAACATTGA	6780
30	ACGCCAAATG CTGTCGTCTA CTTTTAAATT TTCAACTCTT AGAGGTATCT CATATTTGGC	6840
	CATCATATCT ACAGCTACAA CCATTGCGTG AATCTCATTA AAAATAAATT CATTTTTACT	6900
	CGCACTATAA TCTTCACATA CGTCTATAAC TATATAATCA GGTTCAATTAG GAACTTCAAA	6960
35	TACAGCTCTT CTAGGTGCCC AAATATTATG TCTATCAACA TAAAAGTGGG GATATTCTAC	7020
	ATCCTGTTTG TATTTCTTCC TACTGTTATA TAACTTTCT ACCGAGCTCA TCGTTTGTGC	7080
40	GTTTCTAATC ATTATTCCTT TAGGTTTTTC GAGTCGTCGA TTACCTTCTA CTATAAAGTG	7140
	ATAAATATAT TCTGGATAAT TAACCTCTTG GCTAGAAATA GTGTACTTTA TAGTTGTTAC	7200
	ATCTTTCCAA ATTGGAACCT TTTTATTATT TTTTTCGTTA TCATCACTAT CATCTTCTGG	7260
45	TTTAGGTGCC GGTGTAGTTT TGTCTGGATG ATATGGTGGT CTAACAAAAT ATTTAACCCC	7320
	TCCACCTGGT CCATCATGAT AAGAGTGTTC AATTTTATAA GGTGGACTTC CTGTTGCGTT	7380
	ATTTGTATAC CAGTTTTGAT CTACGCCATA CCAATAGTCT TTTGTGCATG GTCCCACTAC	7440
50	AATGTTTACA TGTCTGCCC AACCACCAGT CCAAACACCC CAGTCGCCTG GTTGTGGTAC	7500
	AAAATCTTTT GTATTTCTAA TTATCTTGAA ATCTCTACCT CTATAATTGG ATTTTTGAGC	7560

55

	TAAATCCCAG	CATTGTGCTC	CCATTCCAGA	ACCAGGTACA	TCAATAGCTA	TTTTGTTTTT	7680
	AGCGATATAT	AACGCCCATT	CAACCACTTC	ACTAGCTGTG	GGCTTTCTAT	TTTTCGGATT	7740
5	AGGTAATCCC	ATGTATGCAC	CTCATTTCAA	TCAAAATAAA	AAGCCAGTGC	CGAAGCACTG	7800
	ACTCTTAACT	GTTATTTACA	TTTACCAAAC	CAGAAGCAGC	CCCAGAAGCT	ATATCCTAAA	7860
	ATCCCTTTAA	GCATGGTAAT	CACCTCCTTT	AAATACCAAA	AACAGTTCTT	AGTAAAGCTA	7920
10	TGACAATCGT	ACTGAAGATA	GTCCCTATCA	AACCTAGAAT	CCACATTTTT	ATGTCTCTAA	7980
	TATTCTTGGC	ATTCTTTTCT	TTATTCTTTT	CATCTTCTAC	CTTGTGCGGC	TTTAATTCTT	8040
	CAAAATTTCT	ATCTAATTTG	TCATAAATCT	TTTCTTGCGC	TCTAAGACTA	TCTTCTATTG	8100
15	TGTCGAATTT	TTCAAACATA	GTCTTATCAT	TTTCTTCTAA	TCGCGTTAAA	CGCCAATCTT	8160
	GTTTCATGTC	TTTGGTAAAT	CCAAACATTA	TGCCACCCAC	TTTATTCAAA	TTAAAAAGCC	8220
20	ACAAGCATT	CACCTGTGAC	TTTTCATCTT	TTGTTTCTGG	ATATTTTTCT	CCAGTGATTA	8280
	AAGCGTATTC	TTCTTTATCG	ATTAAACCCT	TGTCTACGTA	CCACTTAATT	TGCTCGTTTT	8340
	TATAGTAACC	CCAAACATAA	AAAGTTTTAA	TGTCTTTAAA	AGTTGGATAA	ATCATCTTCA	8400
25	TTATTTAAAC	GTCCCCCTCA	GTACTTGTTC	TGTTAGTTTT	CAGTTCAGTC	AACTGTTGTG	8460
	TTAACATAGC	GTTTTGTTGA	GCTAATCCCA	TTGTTAATAC	GTTTACTTGT	GCCACCTGCA	8520
	TTTGCACTACT	CGCAACCATT	CCGCGAAGTT	CCTCATCACT	TAAATCTGAC	GCACTTTGTT	8580
30	GGTTTGATGC	ATTCGGTACG	TCTTCTTTTT	CGAAATTGCT	ATTGTATTTA	ATTTGCGCGT	8640
	TAGTGAAAAC	AACTTTTCTA	GGTTCGAACT	CTTCTTTAAA	TTAATAGGC	ACATTGTTAT	8700
	CATCTACATC	TAACTATTG	CGTAAACCGC	CAGTATTAAC	GAATCCGATA	ACTTCGTTTT	8760
35	TATCGTTTAC	TGTGATTTTC	ATTATTTCCA	CCCCATAATT	TTAGTTATAG	TAACTTTGTT	8820
	GGCAJTCGCT	CCAGAACCTG	ATGTTTTACC	TAAATCAAAG	TACACATCGT	TATCTATTCT	8880
	TAAAGTAGTG	CTACTTGTTT	TGGATAGTAA	GCACTCATAA	ATACCGCCAC	CGTTGCCGTC	8940
40	TGAGTCAACT	ACATTCGCTT	TACTCAATTG	AATCGCGTTA	GGTAATGCGG	TTAGTCCGAA	9000
	TCCCTCAATA	ACGCCACCTG	GATAAGTCC	ACTTACCAAC	AAAATAGAAT	AGTTTGTGTA	9060
	CGGTTTCAGTT	AGATTGATTG	TTGTACCTAC	ACCATTGCG	CCACCGTCGA	ACAATACCGT	9120
45	TGATTTATGT	TCATTAGGAA	CTGTCCACTG	TTGCTCAAGT	CTGCCGTTTG	TGATTGATCG	9180
	TGTGTAAATC	TTTTTAGAGT	TATAAGGTGT	GAAGTTAAAT	AGCTTGTTTG	TATCATCTTT	9240
50	AACGAATACC	GATAAATAAC	CCTCATAACT	TTCAACGCTA	CCTGGTAAAT	CCGGCACTCT	9300
	TGTTGCATAG	TAATTACCAG	CAGTTAAATA	TCCCAAATCG	CCTTGCGCAT	TATTTAAGTT	9360

GAATTTATCA TCTACATACT GCTTAGCTTG ATTTAAAGCG TTGTTAGACG TTTCTTCAAC 9480
 AAATGCTTA GTTAAGTTTC CATCATTCTT TTTATAAAAC GGGTACCATG TGCCGTAGAT 9540
 5 TTTGTATTTT GTGTACTCAT CGTTTGAATC GTCTGGGTAC CATGTTGCAC GAGCAGTATT 9600
 ATTATCAACA ACATAAACAA CTAACACACC AGATTTGCTT GATGTATAAG TTGATTCATC 9660
 GAACGAAGAA CCGTCATCAA CACCATCTTG TCCAGGCTTC TCTAACGTGC CTATATCCGT 9720
 10 CTTTCTGGC GCATCTGTTG CATTAGTAAT ATGAATAATC CTAGATGTGT TAACTGCGCT 9780
 TAAAACGCTA TCTATGGACT GCTCATAAGA TTCAATTGCT TTACCGTAAT CATCTGTAAG 9840
 TTTAGACTTT TGCCAATTTCG TTGTTGAATT ACCTTTAACA AGGTCAGCGC CATTGATTTG 9900
 15 TTGTTCAACT TCGTTAACAC GTTCAAAAAT CGCTTGCTCT TTTTCAACTA TTTTATCGAA 9960
 TTCAGCTGTA ACAGCTTGTG TTGCACTAGT TTGCGTCGCA GTAATAGCTT GTATAGCTTC 10020
 20 GTTTTGCTTG ATTTGATTT GTTGAATGCC TTTTGTGCGA CTATCATTCA CTTTGTCTAT 10080
 TAACGTTTGT GTATCAGCCA TATTTTGCTT TAATTGGTTA AAATCTTTAC CGACAGCTTC 10140
 GATAGTATCT TGAATAGATT TGATATAAAC AAGCTTTGTT ATACCATCAA ACCCACTAAC 10200
 25 TAAATCATT TCAATATTGA AGCTAAATTG ACGTTCAACA ACAACATTAT TACTCCCGTT 10260
 TTGTGTAAAG AATGCCTGAG CATGCACCTT GCCTGAATGT TTTAAAAATT CATTGCGTAT 10320
 CACATACTGC AAACGCCCAT TAATTGCGTC TACTATCGTT AATTCGTCTG AAATATAAGC 10380
 30 GCCTCTATCT ACGTTATAAT CATCGGTTTT TAANAAGATA GATGTTTTAA CATGTTTCAA 10440
 ACTTATAGAT AAGGGTCTGT TATnCTTAGT 10470

(2) INFORMATION FOR SEQ ID NO: 21:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3647 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

45 ATCAGATCTT GAGAATCGAG TTATTAAGTC TATCGAAGAC TTAATAAAA TCCAACCATT 60
 CATGCCTACA CAAGATTTTG ATTTTAAAAC TAAAGAAATT CAATCAAACA TTTCTGAAGA 120
 AAGATTTATC GAAATGATTC AGTATTTCAA AGAGAAAATA ACAGAAGGGG ATATGTTCCA 180
 50 AGTTGTGCCA TCAAGAATTT ACAAATATGC GcATCATGCT AGTCAGCATT TAAATCAACT 240
 TTCGTTTCAA CTGTATCAAA ATTTAAAACG ACAAACCCA AGTCCATATA TGTATTATCT 300

55

	TCAAATTGTA	ACAACTAATC	CTATTGCAGG	TACGATTCAA	CGTGGTGAGA	CGACACAAAT	420
	AGATAATGAG	AATATGAAAC	AACTACTTAA	TGATCCAAAA	GAATGCAGCG	AACATCGTAT	480
5	GCTAGTTGAT	TTAGGACGTA	ATGATATTCA	TAGAGTAACT	AAAATCGGTA	CCTCAAAAAT	540
	TACTAAATTA	ATGGTTATTG	AAAAATATGA	ACATGTTATG	CATATCGTAA	GTGAAGTCAC	600
	AGGTAAAATA	AATCAAAATT	TATCGCCAAT	GACAGTTATT	GCGAATTTAT	TACCAACAGG	660
10	TACCGTTTCA	GGTGACCAA	AATTACGTGC	AATTGAAAGA	ATATATGAAC	AATATCCACA	720
	TAAACGGGGC	GTTTATAGTG	GTGGTGTGG	ATACATAAAT	TGTAATCATA	ACTTAGATTT	780
	TGCATTAGCA	ATTCGAACGA	TGATGATAGA	TGAGCAGTAT	ATCAACGTAG	AAGCTGGTTG	840
15	TGGCGTTGTA	TATGATTCTA	TTCCTGAAAA	AGAACTGAAT	GAAACGAAAT	TGAAAGCTAA	900
	AAGCTTATTG	GAGGTGAGCC	CATGATCTTA	GTTGTAGATA	ATTATGATTG	CTTTACATAT	960
20	AACCTAGTGG	ATATTGTTGC	TCAACATACT	GACGTCATTG	TTCAATACCC	TGATGATGAT	1020
	AATGTGCTGA	ATCAATCGGT	GGACGCTGTT	ATTATATCTC	CTGGTCCAGG	GCATCCATTA	1080
	GACGATCAAC	AGTTAATGAA	AATCATATCA	ACCTATCAAC	ACAAACCCAT	TTTAGGTATT	1140
25	TGTTTAGGGG	CTCAGGCACT	GACTTGTTAC	TACGGTGGAG	AAGTCATTAA	AGGCGACAAG	1200
	GTTATGCACG	GCAAAGTTGA	TACACTAAAG	GTTATATCGC	ATCATCAACA	TCTGTTATAT	1260
	CAAGATATAC	CAGAACAGTT	TTCAATTATG	AGATATCATT	CATTAAATAAG	TAACCCTGAC	1320
30	AATTTTCCAG	AAGAATTGAA	AATTACTGGA	CGTACCAAAG	ATTGTATACA	GTCATTGAG	1380
	CATAAAGAAA	GACCGCATT	TGGTATTCAG	TACCATCCTG	AATCATTTGC	TACAGACTAT	1440
	GGTGTCAAAA	TAATTACAAA	TTTCATTAAT	CTAGTGAAGG	AAGGATGAAA	ACCATGACAT	1500
35	TACTAACAAG	AATAAAAACT	GAACTATAT	TACTTGAAAG	CGACATTAAA	GAGCTAATCG	1560
	ATATACTTAT	TTCTCCTAGT	ATTGGAACCTG	ATATTAAATA	TGAATTACTT	AGTTCCTATT	1620
	CGGAGCGAGA	AATCCAACAA	CAAGAATTAA	CATATATTGT	ACGTAGCTTA	ATTAATACAA	1680
40	TGTATCCACA	TCAACCATGT	TATGAAGGGG	CTATGTGTGT	GTGCGGCACA	GGTGGTGACA	1740
	AGTCAAATAG	TTTCAACATT	TCAACGACTG	TTGCTTTTGT	TGTAGCAAGT	GCTGGCGTAA	1800
	AAGTTATAAA	ACATGGTAAT	AAAAGTATTA	CCTCaAATTC	aGGTAGTACG	GATTTGcTAA	1860
45	ATCAAATGAA	CATACAAaCA	ACAACTGTTG	ATGATACACC	TAACCAATTA	AATGAnAAAAG	1920
	ACCTTGTATT	CATTGGTGCA	aCTGAATCAT	ATCCAATCAT	GAAGTATATG	CAACCAGTTA	1980
50	GAAAAATGAT	TGGAAAGCCT	ACAATATTAA	ACCTTGTGGG	TCCATTAAAT	AATCCATATC	2040
	ACTTAACGTA	TCAAATGGTA	GGCGTCTTTG	ATCCTACAAA	GTTAAAGTTA	GTTGCTAAAA	2100

55

	AAGCAACACT ATCTGGTGAT AATTTGATAT ATGAATTGAC TGAAGATGGA GAAATCAAAA	2220
	ATTACACATT AAATGCGACT GATTATGGTT TGAAACATGC GCCGAATAGT GATTTTAAAG	2280
5	GCGGTTCCACC TGAAGAAAAT TTAGCAATCT CCCTTAATAT CTTGAATGGT AAAGATCAGT	2340
	CAAGTCGACG TGATGTTGTC TTTACTAAATG CGGGTTTAAG CCTTTATGTT GCAGAGAAAT	2400
	TGGATACCAT CGCAGAAGGC ATAGAAGCTT CAACTACATT GATTGATAAT GGTGAAGCAT	2460
10	TGGAAAAATA CCATCAAATG AGAGGTGAAT AATATGACGA TTTTATCAGA AATTGTTAAA	2520
	TATAACAGT CACTTTTACA AAATGGCTAT TATCAAGACA AACTTAATAC CTTGAAAAGT	2580
	GTGAAGATTC AGAATAAAAA ATCTTTTATA AACGCAATTG AGAAAGAACC AAAGCTAGCA	2640
15	ATTATTGCAG AAATTAAATC GAAGAGTCCT ACAGTTAATG ACTTACCTGA ACGAGATTTA	2700
	TCGCAACAAA TCTCAGATTA TGACCAATAT GGTGCAATG CCGTGTCCAT TTAACTGAT	2760
	GAAAAGTACT TTGGTGGTAG TTTTGAAAGA TTACAAGCAT TGACGACAAA AACAACATTA	2820
20	CCCGTATTAT GCAAAGACTT TATTATAGAC CCGCTTCAA TTGATGTTGC TAAACAAGCT	2880
	GGTGCACTA TGATTTTATT GATCGTTAAC ATCTTATCTG ATAAACAATT GAAAGATTTA	2940
25	TATACTACG CTATATCGCA AAATCTAGAA GTGTTAGTTG AAGTACATGA TCGCCATGAA	3000
	TTAGAACGTG CCTATAAGGT TAATGCTAAA TTGATTGGTG TAAATAACAG GGAAGTAAAA	3060
	CGATTGTGTA CAAATGTGGA ACATACAAAT ACTATTTTAG AAAATAAAAA AACAAATCAT	3120
30	TATTATATTT CTGAAAGTGG TATTCACGAT GCATCTGATG TAAGAAAAAT CTTGCATAGT	3180
	GGTATCGATG GCTTACTAAT AGGTGAGGCG CTTATGCGTT GTGACAATCT ATCTGAATTT	3240
	TTACCACAAC TGAAGAAATGCA AAAGGTGAAG TCATGATGAA ATTGAAATTT TGTGGCTTTA	3300
35	CATCAATAAA GGATGTTACA GCGGCCAGTC AATTACCTAT TGATGCGATA GGTTCATCC	3360
	ATTATGAAAA AAGTAAAAGG CATCAAACAA TTACCCAAAT AAAAAAGTTA GCGTCTGCTG	3420
	TTCCAAATCA TATCGATAAA GTATGTGTCA TGGTAAATCC TGATTTAACA ACAATTGAAC	3480
40	ACGTATTAAG CAATACGTCA ATTAACACAA TACAGTTACA CgGCACAGAA TCTATTGATT	3540
	TTATACAGGA AATTAAAAAG AAATATTCAA GCATTAAAAT CACTAAAGCT TTAGCTGCaG	3600
45	ATGgAAAAACm TwATCCCCAA caTtAAtnAA tnTTAgGGGG TCCGTGG	3647

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5966 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

	CcAcCTTGAC CACCTTTACG TGAATCTTT TcMcCTkGAG CAACaTCGaT AATaTATATT	60
5	GAAAgTCAAC AAGTTCTGGA CTAAATGTTG CTGCTAAGTT ATCGCCACCA GATTCTATGA	120
	AAATTAGTTC TATATCGTCA TGACGTTCTA ATAATTCGTC TATTGCTGCA AAGTTCATAG	180
10	ATGCATCTTC ACGAATCGCA GTATGAGGAC ATCCACCAGT TTCAACACCA ATGATACGAC	240
	TTTCAGGTAG AACTCCTGAA TTTACTAATA TCCTTTTCGTC TTCTTTTGTA TATATATCAT	300
	TTGTAATAAC GCCGATACTC ATTTCTTTTG AAAGACGTTT TACAACTTTT TCAATTAATT	360
15	GTGTTTTACC TGCACCTACA GGACCACCAA TACCAATTTT AATCGGATTT GCCACAATTA	420
	TAACCTCCTA TGATATGAAA tCTAACATT GaCGTCTCA TGCGCCATTT GATTTAGTTC	480
	TAAACCAGGC GCTGTCATGC CAAAATCTGC TTCTTTTAAAT TCGAAAATCT GCTTTCTTGT	540
20	TCCTTCTATA TAAGGAATCA TGTGAGTAAC TATCTTTTGA CCAGCAGTTT GTCCAAGTGG	600
	AATAGCACGA ACAGCATTTT GAGTTAAACT TGAAACATTT TGATATAAAT AGTAATCAAT	660
25	AATCGTTTCA ATATCTACAC CTAAATGATG GCCTAGCATA GTAAAACAAA TAGCTGGATT	720
	TnACTTTGCT TTCTTATCTT GCATTTGTTG ATGATACCAA GCAATCCATG GGCTATtATA	780
	AAGTTCTAAA GCCAATTTAA CCATGCGAGT CCCCATTTGT KTTGCACCAA CACGTGTTTC	840
30	TTTAGGTAAG TTTTGtACAr ACATCAGTTT ATCTATGTGT AATACTTTTT GTGTATCATC	900
	ATTTTCCAAT GCATCATAAA CTAAACGCAT GGCTAAACCA TCAGAATAGG TAAGTTGCTC	960
	TTGTAAAAAC ATTTTTAACC AAGCAATAAA AGTATGATCG TCATGAATTA TATTTGTTG	1020
35	AATATATGTT TCAAGACCAA ATGAATGACT GAAAGCACCT GTTGGAAGT GTGAATCACA	1080
	GAACTGAAAT AATCTTAAGT GTGTATGATC AATCATGAGA ATGCCCTATA TGTCTGAAAG	1140
	CCTtATTAAC TTTACGGTCT TCTCGAACAT ATGGGATGCC TAACTTTTT AATAAATCTT	1200
40	CAACTAAATA ATCATATTGT ACTAGCATTT CAGTCTCTGT AAATTGTGCT GGCAAATGAC	1260
	GATTTCTTAA TTGATGGGCT ATATCTCCCA TTTCTTGCAA TGTCTTGGT TGAATCACTA	1320
	AAAGATCTTC TGAATTAACA TCCACAATAA TCATATTATG GTCATCTGCG TATAAAATAT	1380
45	CTCCATATTG TAAGTCAATA GGTGTTTTTA AACGAATGCC TATTTCAAGT CCATGGTCTG	1440
	TAACGACTCT TTGAATACGT TTAACAAGAT CTGAATTTTC AAGGTATACT TTTTCGACGT	1500
50	GCTTTTGTTT TTCTGAATTT GACAAATTGG CAATATTGCC TTGGATTCT TCAACAATCA	1560
	TTCTATGTTT CTCCTAGAAT AAGAAGTATC TTTGAGTTAA TGGTAACTCA GTTGCTGCAT	1620
	TACTTGTAAT TTTTTCTCCA TCTACATATA CTTCATATGT TTGTGGATCA ACGTCTAATT	1680

55

	GACGCACCAT	GCGTTTTTAAA	TTTAATGCAC	GATTGATACC	ATTTTCATAA	GCAGTTTTAG	1800
	ACACGAATGT	CATTGACGTA	CTTGTAAGGT	TTCCGCCGTA	TTGACCATAC	ATTTTACGGT	1860
5	ACTTCATCGG	TTCAGATGTA	GGTATAGAAC	CATTTGCATC	GCCATTTACG	GCAGAGTTAA	1920
	TTAATCCGCC	CTTTACAAC	AATTCAGGTT	TAACCCCAA	GAAAATTGGG	TCCCATAAGA	1980
10	CAATGTCAGC	TAGTTTGCCC	GGCTCGATAG	ATCCTACATA	TTCAGAAATA	CCATGTGTAA	2040
	TTGCTGGGTT	AATTGTATAT	TTAGCGATAT	AACGTTTGAT	GCGATTATTA	TCATTATGTT	2100
	CAAAATCACC	ATCTAAAGGA	CCACGTTGTT	CTTTCATGCG	ATGTGCTACT	TGCCATGTTT	2160
15	GTGTAATTAC	TTCACCTACA	CGGCCCATG	CTTGTGAATC	GGAACATAATC	ATACTGAATA	2220
	CACCCATATC	TTGCAGAACA	TCTTCTGCTG	CAATCGTTTC	TTTACGAATA	CGTGAATCTG	2280
	CGAATGCGAT	ATCTTCAGGA	ATAGCCGCAT	TTAAATGGTG	AGTAATCATT	ACCATATCTA	2340
20	AATGTTTCATC	TACAGTATTA	TGTGTATAAG	GCAAAGTTGG	ATTTGTAGAT	GAAGGTAAAA	2400
	TATTTGAAAA	TGCAGCGGAT	TTAATTAAAT	CAGGCGCATG	ACCGCCACCA	GCACCTTCAG	2460
25	TATGGTACAT	ATGAAGTACA	CGGTCTTTAA	CAGCAGCCAT	TGTGTCTTCC	ATAAATCCTG	2520
	CTTCATTTAA	AGTATCTGCA	TGTAATGCAA	TTTGAACATC	AAATTCATCA	GCAACATCTA	2580
	ATGCATGACT	CAAAGCAGAT	GGTGTGTCAC	CCCAGTCTTC	ATGTACTTTT	AATCCAATTG	2640
30	CTCCGGCATT	GATTTGTTCA	ATGAGTGCAG	TTGGATTTGT	TGCTTGTCCT	TTACCTGTAA	2700
	AACCGACATT	AATCGGTAAA	CcTTCCGGCAG	CTTCTAACAT	TCTATGAATA	TGCCATGGAC	2760
	CTGGAGTTAC	AGTTGTTGCT	TTAGAACCTT	CTGAAGCACC	AGTACCACCA	CCAATATGAG	2820
35	TCGTAATACC	ACTTTCTAAT	GCGACCTCTG	CTTGTTTCAGG	ATTAATAAAA	TGAACATGAG	2880
	TATCAATACC	ACCAGCAGTG	ACGATTTTAC	CTTCAGCGGC	AATGATATCT	GTTGTTGAAC	2940
	CTATAATAAT	GTCGACATTA	TCCATTATAT	CTGGGTTGCC	GGCATTACCT	ATGGCGAAAA	3000
40	TATAACCATT	TTTAATGCCT	ATATCAGCTT	TAACCACTTT	ATCGTAATCG	ATAATAACGG	3060
	CATTAGAAAT	GACAAGGTCT	GCAACGTTCA	CGTCATCACG	TGTTACACGA	GGATTTTGCG	3120
45	CCATACCGTC	TCTAATAGAT	TTACCACCAC	CAAAAGTAGC	TTCTTCACCA	TAAACCGCAT	3180
	AGTCTTTTTT	TATTTGAGCA	AATAGATTCTG	TATCACCTAA	ACGAATGGAA	TCTCCAACAG	3240
	TTGGACCGTA	TAAGCTCGTA	TATTGATTTT	GCGTCATTTT	AAAGCTCATG	ATCTTTTTTCC	3300
50	TCCTTTTTTA	TTACAGTTTT	CAGCACCGTT	ATCTCCGAAT	ACACCTGCAT	ATTCATCATT	3360
	TTCATCAGTT	GGGCGATAGA	CACGTGACTC	ATCGATAGGA	CCATTGACCA	TACCACGAAA	3420
55	ACCAAAAATT	TTACGTTTGC	CAGCATATTC	AACTAATTGA	ACTTCTTTTT	TATCCCCAGG	3480

	TTCGAAATCT AATGCTGCAT TTGCTTCATA AAAATGAAAA TGTGAGCCCA CTTGAATTGG	3600
	TCGATCTCCT GTATTTTCAA CTTGATAAC TGTTTCAGGA TGATGGTTAT TAATTTCAAC	3660
5	CTCTGTACTT TTTGTAATAA TTTCTCCTGG TATCATTTGA CTGCCTCCTT TAAACAATAG	3720
	GGTGATGTAC TGTGATTAAAC TTAGTACCAT CGGGGAACGT AGCCTCGATT TCGATATCTG	3780
10	TAATCATGTG TTCGACACCA TCCATGACAT CTTCTTTGTT TAGAATTTGT CTACCATAAC	3840
	TCATTAACTC TGCAACGGTC TTACCATCGC GTGCACCTTC TAATAATTCA TCGCTGATTA	3900
	AAGCTAATGC CTCAGGATGA TTTAGTTTCA AACCACGTGC TTTACGACGA CGTGCAACTT	3960
15	CCGCCGCCAC TACAATCATT AATTTGTCTT GCTCTCGTTG TGTAAAATGC AAATTAAAC	4020
	CCCCAATTTT ATATTAGATA CaATTTACAA AATTTATATT AATCCTAATT GTTGTGATAA	4080
	ACAAGTAATA TACAAAGTTC AATGTGTAAT TAGAAAATTA TATTTTTAGC ATATCCGATA	4140
20	TTGAAGCAAA CAATCTAATC GAAAACAAAT AGTGGAATAT ATTTATGTAA AAACCAAAAT	4200
	AGTTTTTAAT ATAACTTTTT ATAGAATAGT AGTATATTAA TGAGTAATGA TTCAAAGGAA	4260
	AGGTGAAAGA TTTGAAGATA ATAGATGTGC TTTTGAAAAA TATATCTCAG GTTGTGTTAA	4320
25	TTAGTAATAA ATGGACAGGA TTATTTATCT TAATAGGATT ATTTGTAGCC GATTGGACAA	4380
	TTGGATTAGC GGCTATTGTA GGTAGCATCA TCGCCTATAC TTTTGCGCGT TTTATAAATT	4440
30	ATAGTGAGGC AGAGATTAAT GATGGGTTAG CTGGATTAA TCCAGTGCTA ACTGCCATTG	4500
	CGTTAACAAT CTTTTTAGAT AAGTCAGGAT TAGATATTGT TATAACAATG ATAGCAACTT	4560
	TATTAACGTT ACCAGTTGCT GCTGCAGTGA GAGAAGTTTT AAGACCATAT AAAGTTCCGA	4620
35	TGCTGACGAT GCCTTTTGTC ATTGTGACTT GGTTTACAAT TTTACTTTCA GGACAGGTTA	4680
	AATTTGTAGA TACATCGTTA AAGTTAATGC CTCAAAACAT TGAAACGGTT AATTTTAGCA	4740
	ACAATGATAG AATaCATTTC ATTCAATCAT TATTTGAAGG ATTCAGTCAA GTATTTATCG	4800
40	AAGCGAGTGT AATTGGTGGC GTATGTATTT TAATCGGCAT ATTGATAGCA TCAAGAAAAG	4860
	CAACACTCTT AGCTGTTATA GCTAGTTTGT TAAGCTTTAT CATTGTAGCT CTATTAGGTG	4920
45	GTAATTATGA TGATATTAAT CAGGGATTAT TCGGTTATAA CTTTGTATTA ATGGCAATCG	4980
	CACTAGGATA TACATTTAAA ACAGCGATTA ACCCTTATAT TTCGACTTTT TTAGGTGTGT	5040
	TATTAACAGT AGTGGTGCAA CTAGGTACAA CAACATTGCT TGAACCGTTT GGCTTACCTG	5100
50	CATTAACATT GCCATTTATT ATCGTGACAT GGATTTTATT ATTTGCTGGT ATTAAACATG	5160
	ACAAAGTAGA TGCTTGATAG TTAAATCAAA CCTAATATTG TTTGAATATC ACCTTAACT	5220
55	ATACAGCGAA TTGTATAGTT TAAGGTGTAT TTTTATGGAT AAAATTAAGT GCATACTTAA	5280

GTGTTAAACT AGGAATAAAT AATTTATATT GTGTGTTGTG TGGGGTGA CT AATATGAATG 5400
 ATATGGATAA TTCCTTTTATA ATAACAACGG AAATTCAAAG AAAATGGATT GAAAAATTCA 5460
 5 AAGTAATTAG AGATACATTT AAGGCTAAAG CTGAATATAA TGATCAACAT AGCCAATTTT 5520
 CATATAAAAA TATTGAATGG TTAATTAAAG AAGGTTATGG AAAATTAACG TTACCAAAAAG 5580
 CATATGGTGG TGAAGGTGCG ACCATAGAAG ACATGGTTAT TTTGCAATCA TTTT TAGGCG 5640
 10 AACTTGATGG TGCCACAGCA TTATCTATTG GTTGGCATGT GAGTGTCTGTA GGACAAATTT 5700
 ATGAACAGAA ATTATGGTCT CAAGATATGT TGGAGCAATT TGCTGTTGAA ATTAATAATG 5760
 15 GTGCATTAGT TAATAGAGCA GTTAGTGAAG CTGAAATGGG TAGTCCAACA AGAGGGGGAA 5820
 GACCAAGTAC ACATGCTGTT AAAGCTGATG ATGGGTATAT TTTAAATGGT GTGAAGACAT 5880
 ATACATCAAT GAGTAAAGCA CTAACACATA TTATTGTTGC TGCTTATATA GAAGAATTAG 5940
 20 AAAGTGTGG TTTTTTCTTA GTAGAC 5966

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 17310 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CTGTGTCATC GCGAAATAGT TAGGGTCATT CATTAACTCT TTTGAACGTA TTTTCATCAA 60
 35 ATATAACAAT TTCATTAGTA AAGGGGACTT GTTCAAACCA GCTATAATAC AAAATAGACC 120
 TATAGTCACA CTGCTTATAA TATAAGAGGT AACGATCACT TTTTGTCTAT TACCTAACTT 180
 AAAGTGATC ATCCCTAAAT AGAAATAAAT GACTACAAAT GCATATTTAA CTGTAGATGC 240
 40 AAGAACTTCC TTAACCGTAA TAAATATCAA ATCATCAAAA AATaGCaAAC AArGCGTAAT 300
 AATCATACGA TATGTATACA AAATAATGAm,AAACTGTmAA AAATGATTTG CCTTTAATAA 360
 ATGGTTAGCG AAAACAGTA AATAAACTAA TATTAGTAAT GTGATAAAGT CAGCTATAGA 420
 45 AACATTCA CA CCGCAATAA CCGAAGATTG CTGAATAAAA ACCGCTAAAC CGATAAGTAA 480
 CAATGTTAGT AATTTACTAT TGTGTTGATT TTCCATTATA AACGTCTTCC ACTTCTTTAA 540
 50 TCATTTTCTC CTCAGTAAAA CATTCTAAAT AACGTTTCT AGATTGATTA CTCATTTTGA 600
 TGTAATCACT GTCTATTAAA TATTTTTC CA GGACTTTAGC AATAGTTTCG GGTGGTTGT 660
 TCATCATACA TATACCATTA TCAGCTACTA ATTCTGAAAT ACCGCCAACA TGACTGGCTA 720

55

	TTATTAAAAT AAACGTATCG TATTGTGATA ATAAATGACT CGCATTAAATG ACATTGCCCA	840
	AAAATGTGAC ATCATTITTTCT AACCCAGCTT GTACAACTTG TTGCTGACAA TCATTTAATG	900
5	TAGGTCCATC GCCTATAAAT GTAAAATGCG CATGATTACT GTTATGTAAT TTCAATATCT	960
	CTATTGCCGC GATTAGATT TGTGGCAATT TTGGATAAGC AAATCTTGCA ATCATAACAA	1020
10	ATTGATGCTT TGTGCGGGCA TTAATCTGTA AATCTTGTTT ATTAGGCAAC ATTCCAACTA	1080
	CTTCGCCAAT ATTGTTATGT GATTGGCTTT TTAGCGTTTG CTTAACAGCG GGAACATCTG	1140
	CAATACCATT ATGTATTGTG GTTAATTTCA ATCGATTAAA TCGATATTTT AACGCTAACT	1200
15	GTTTATCGAA ATCTGAAACA CAAATAATGC TATCTGTAAT AAGTGACATT AATTTTTCGA	1260
	TAACATAATA TAGAAATTTT TTAGCTGGTT TAACACCCTC TGTAAAAGCC CATCCATGTG	1320
	CAGTAAAAAC TATACGTGTG TCTTTTCGATT TCGAAATGAA CTtCGCAATT CGTCCGACCG	1380
20	TtCCAGCTTT GGAAGAATGT AAATGGATAA CATCAGGTTT AATTTTCGAG AATAACTGTG	1440
	CTAACACTTT GACAGCTAAA ATATCTTGTT TAAAGTCAAT TGGACCTACT AAATGTTTGA	1500
25	TAATAATTAC ATTAACCTCT GCATCTAGTT GTTCAATCAT TGGTCCATGA TTGCCTACAA	1560
	TGACATAAAC ATCATTGTGT ACGCAAAAAT GGTGGCGAG TTGAATGAGA TGTGTTTGTG	1620
	CACCACCATT GTCTGCTTTA GTAATACAAT ATATAATTTT CAACTGTTAC AAACCCCTTT	1680
30	AATGCTATAC TTTCAATTTT TTAACATGGC TATCTCATCA GATGAATAGT ATTTATAGCC	1740
	ATGCAAATCA ATGATGGCAC ATATTTCTTA ATGCCATTTG ATACTGTCTC AAGGGATTCC	1800
	TCGTTATACT GTAACAATTG GTCACAATCT TTAATAATA ACTTTTATTT GAACTTATTA	1860
35	AGTAAATTAA GACTACCTTG AGCCTTCCCC TGTAATAACA ACCATCAATG TTCTAATTGA	1920
	TATATATAGT TCCATCATT AACTACCTTT ATGTATATAT TTCATGTCAT ATTTAGTTT	1980
	TTGTTGCGGT GTTAAGTCAT ATCCACCTTG AATTTGCGCA AGTCCTGTTA ACCCTGGTGT	2040
40	AACAAGACAT CTTTGCTCGA AACCTATCAC TTCTGAACTA AATAATTCTA CAAATTCCGG	2100
	ACGTTCCGGG CGTGGTCCAA TAAAACTCAT TTCCCTTTA ACAACATTAA TTAGTTGTGG	2160
45	TAATTCATCA ATGCGTGTTT TACGAATAAA CTTCCCGACA TTTGTTATAC GATCATCATC	2220
	TTTATCAGCC CATTGCGCAC CGTTTTTCTC TCGTTTTTTG CACATCGAAC GTAATTTGTA	2280
	TATTTTAATT AATTTACCCA TCTTCCCAAC TCTAACCTGA CTATAAATAG GGTTCCTGG	2340
50	CGAATCTATG ACGATAGCAA TGGCGAATAT AACCATAATC GGTAAAGTTA AAAATAATAA	2400
	AACAATGCTT AAAATTAAGT CAATCGCAGC TTTAATTGGG TAATAGCTTT TTCTCACTTC	2460
55	TTCTAGTTTG TCTAATTTTC TTTGATAGGC ATAACCCTTA TTATTATGGA CAGCTTCAAT	2520

	AATTAAAGTA ATCCTTTAAA CCTGTTTCTA CTGTATATTT AGGAACAAAT CCTAATGCCT	2640
	TTAAGTTAGA AATATCTGCA TAAGAATGCT TAATATCTCC TTTTCGTGCT TCTTTAAATT	2700
5	CATGCTCGAC TGATTTTCCA TATAATTCAC CAATAATACG ATAAACCTCT AATAAATTAG	2760
	TAAAAGTGCC TGTACCAATG TTATAACCGT GTCCAATTGC ATCTTTGTGT TCCATAATTA	2820
10	AGCGTACAGA TTGAACAACA TCATATACAT ATACAAAATC TCTAGTTTGC AGTCCGTCAC	2880
	CAAAAAATGT AAATGGCTTG TTATGCTCAA ATGAATCGAA CATCTTTGAA ATCACACCTG	2940
	AATATTGTGA CTTAGGATCC TGTCTTGGCC CAAATACATT AAAAAATTTA ACAACCGCTG	3000
15	TTGGTATGTT ATATAACGAA CAATAATTTA ATGTCGTCCG TTCGCCGTAA TATTTATCTA	3060
	TTGCATATGG TGATAATGGT AAGATTAATG ATTGATCACT TTTAGGCAAA TCAGGAAGAT	3120
	CACCATAAAC AGCTGCTGAC GAAGCAAAGA TAAAACGTTT TATATGATTA TTATATTTTT	3180
20	TAATGATTC TAACAATCTT AATGTTGCTA CGACGTTTAT TTCTTGAGAT AAGATAGGTT	3240
	TCTCAACCGA CTCAGCAACA CTAACATAATG CTGCTAAATG AATAACATAA TCAAATTGAT	3300
25	ATGTCTTCAT GATTTGTTCA ACTGCATCAT ATTCACGAAT ATCTAATTCA AACACATGAT	3360
	CGTCAGCCAA ACTTTTAATA TTTTCTCGTT TACCTGTTCT ATAGTTATCT AGAACATAAA	3420
	CATCATAATC TTGTTGTAAA TCATCTACTA AATGCGACCC AATAAAACCA GCCCCACCAG	3480
30	TTATCAAAAC TCTTTCCAAA TCTTCCACCT CATTTATACA TTAAAAATAT ATCATAAAAA	3540
	CATAAGTAT TGTAAGCTTT TTATCGATAT TTTTATTTA TAAAAATAAA ATGAGATAAC	3600
	TTTGTGAATT TTTATTGAGA TAAATTAGAT AGTGGTGTTT TTGTGATGTT TTATAATATC	3660
35	TTGGGTGTGT TAATACTAAT AATGCTTTCA ACTGATGCAT TAGACTGTGA CATCATAACT	3720
	CACTTAAGAA CTTGCTTAT TAATTTTCTA CCAATACACT CCCTTCTAAG TGCACTAAAA	3780
	AATCCTTACT GCTAAGTGAT TAACTTAAC AATAAGGATT TATTTATCAT TAGTGGATGA	3840
40	TTATTAACGG AATCTCATA CACCATCTAC AATAATTGTT TGTCCAGTAA TGTAATCAGA	3900
	GTCTTTACCA GCTAAGAAGC TCACTACATT TGAAACATCT TCTGGTTGAG AACTCTGCC	3960
45	CAAAGCAATC TGAATTGTAA ATTGTTCCCA ACCCCATGCT TCAGGTTTAC CTGCTTCTTC	4020
	GGCTGTTGCC ACTGCGATAC TTTCCATCAT TGGTGTGTTGA ACGATACCAG GTGCGAATGC	4080
	ATTCACAGTA ATACCTTCAG ACGCTAAATC TTGTGCGGCT ACTTGTGTTA AACCTCGCAC	4140
50	TGCGAATTTT GTAATGCAAT ATAAAGACAA GCCTGGGTTA CCCTCAACGC CTGCTTGAGA	4200
	TGTTGCATTG ATAATTTTAC CGCCATGATT GAATTTTTTA AATTGTTTAT GTGCGGCTTG	4260
55	AATACCCCAT AGCACACCTG CAACGTTTAC GCCATATACT GTTTTAAACT GTTCTTCAGT	4320

5 GCGCAATTGC GCGGCAGTTT GTCTTAcTGC GTTAAATACA TCATCACGGT TTGATACATC 4440
 TGCTTTGATA GCAATAGCTT TTGTACCATC ACTTGATAAT TTAAGTGCAG CTGCTTTTGC 4500
 CCCTTCTTCA TTGAAATCAA CAACTGCTAC TTTGAAACCA TCTTCCACTA AACGTTCTGC 4560
 AATTTTAAAA CCAATCCCTT GTGcTCCGCC AGTTACTAAT GCTACTTTGT TGTTTGTCTAT 4620
 10 AAAGATCACT CCTCAAATTT CTTTCCTTTA ATTACATTTT ACTCCTCTTC ATTTGAATAG 4680
 TACAACAAAG GTAGCTCCAT TTAACAAAAT ATTCAGATAT TTAAGGTATA GTTAAACGCA 4740
 CTACCATTAG TGATTGGCAA TCGGTTTAAA TGTCGTTTAA AAAGTTCTTA TGTTGAATAT 4800
 15 TATTTTTTTA AGTCTCTCGA TTAGTTTGTC ATCAATCTTT TTTGAGACA TGGTCTTTTG 4860
 ATTCAATAGG CGGTTCCGTG TTATCACTGA CAACTTTAGT TGTAGCTTCA TCTTTATGTA 4920
 TTTCTTCGTT AAATCCTTCA AGGTTTTTAG TCGTGGGATT TTTAACCTCA GGATGTTCCA 4980
 20 TCATGTCTTG ACTATCAAGT TCCTTTTTAC ACGTGTCTTT ATGTGATGCT TGATTGCGT 5040
 TCCCTTTACT TTTTGAATA GTGGTAGTAT CTGCTGCAGC TACTAATTTT TTTCTACCTA 5100
 AAATAGATAT GGCTGAAACA AACCAGAGTA TTGCAGATAC AAAGTTGCAT AATACTAAAG 5160
 25 CGATAATAGC CAATACAATT AATATGACAC CTTTGAAT CTTTCTTTA AATAAGTCAG 5220
 ATGCCAATAC GATGACAGGT ACGATTGAAA GTATAATTAC AAATATAGAA ATTATTGCCG 5280
 30 ATATAACTAT TGTTACTATT AAATAATCAG CTCTGCTACC TGATAATAAA TAGAAAAGGC 5340
 CGAAATTAG TCCATAGCAA ATTACAAACC CACATAAAGT TATAGCCATG AGTACTATAT 5400
 AAGCTATTG AAAATATAAA CCTATCTTTA TGAATGATTT TTCTACATTT TTTTCCATGT 5460
 35 CTATTCCCCA TTTATTTAAA ATTTATACTT TACCTTAAAT ATTCTCTTTA TTCTTTAGTG 5520
 ATTTTATCTT TAGATTCAAA TTGATTCTCT GTACTTTCAA TATCAACTTT TTCATTTTCG 5580
 TCTGTCGATT CATCTTTTGA GTATTTATTC CAAATCAGCA AAATACCACC AATCAGCCAT 5640
 40 AAAATTGACG AAAGGAAATT ATATAAACAC AGTGCAATAA TAGCATAAAC AATAAAAAGT 5700
 GCACCTCCGA TTACAGAGTA ACTTTCCATA TAAATCGCAG TAAAGATGGT TGGTAAAACA 5760
 GTGAAAAGAG CCAATATTAA TCCTAATAAA AAAATTGTTT CGTAATCAGA TCCTCCAGCA 5820
 45 ATATTAATAG ATATCATCCT AACAAAAACG AACTAAAAT ATATTTGAGC TACGATGCCT 5880
 ATCCAAATTG CTATTTTTCC TATAATTGAG CTCATACTCA TTCCCATT TTAAAAATT 5940
 50 TATACTTTAC CTTAATATAC CTTATTTTAT TTAATTTTAA TATGCAAAAT AAAAAATGG 6000
 AGAACTTCAA TATTTATAAA ATATCAAAAG TTCTCCACAC TATATTGTTT TATTATATTT 6060
 TCGCTATCAA TACGCTAAAT CATCATATTT CCTCAACAT CACAGTAAAA CTATTGCTCC 6120

55

	TTCCAATTGC	GCA GTTGTTC	AACATCATCA	TCTTGT TTTAA	GTAATGCCAG	TGGTACTTGA	6240
	AGATTAAGAC	ATCGTCCTGA	AATATTAAAG	CGTGTCCACAC	CTGCTGGCAC	AGTTTCCCCT	6300
5	TTATGAACAA	CCGCTTCAAT	TTCCTTATAA	CTCAATGGCT	GATACTTCAT	GAGTACATCT	6360
	TGTTGAGAAA	GACAAGGATA	TGTACCTTGT	GCAATTCTCT	CTACAGAACA	ACAACCACTA	6420
	TAAC TTGCGA	CAACCTTTTC	CCATACTTGA	AAATGTGCTT	CGCCTAAATC	TTTTGTATAC	6480
10	AAATATTGTT	CTGTATCACC	ATGACACATT	GTAATAAATG	GCGCTTCTTG	TCTTGTCTCA	6540
	GTAGTCCATG	GCAAGCGATG	TTCTTGT TGT	AACGTTTCCC	ACCACACACC	AAATGGA ACT	6600
15	TTATGTTGCC	ATGTACTAAT	TGAATATTGT	GTTTCATGGA	TTTCTTG CAC	TGGA ACTTTT	6660
	TTACATCCTA	ACGCTTTCAA	ACTTGTATAC	CGATGCACAC	CATCTATAAC	CATATATCTA	6720
	CCATGTTGCA	TCGCTGTCAC	TAAATAGGA	TGACGTATAA	AATCATCTGC	TTCAATACTA	6780
20	CTTTTCGTTT	TTTCCAATCT	TAAAGGTTTG	AATGTTTCGT	GAAGATCAAT	CTTATCTACT	6840
	GGTACCAATT	TTAAATGTTT	ATGAATATGA	TTCAATAGTT	ATTCATCCTC	CTTTGTTTGT	6900
	GTTAAATAAA	TAAATTCAGG	ATGTGGATGG	CTTAAGAAAT	CGTGATGTGA	AATAGACCAT	6960
25	CCGTATGCAC	CTGCATATTT	GAAAACAATA	ACGTCGCCCTG	TACTGATTGC	GTCTATCTGT	7020
	ACTTCTCTAG	CAAAGACATC	TTTCGGTGTA	CATAATTGAC	CGACTAACGT	TGTGTCCTGT	7080
	CTCGAAATTG	AACTTTTTTC	AAATGAATAT	GGATTGTCTT	TATAGCGATA	AATGTCAAAA	7140
30	GGATGGTTAT	GTTGCCAAGA	TACCGGCAGT	CTAAATTGTT	GCGTACCTCC	TCTTAATATG	7200
	GCATACCAAG	CACCATGTAC	TTCTTAATG	TCTAGCACTT	CTGTCACATA	GTAACCAATA	7260
35	TGTGCCACAA	TAAAGCGCCC	ACATTCAAAG	TTCAATGTCA	CATCTTCCAT	TTCTTGCTCA	7320
	ACGATAAGTG	TTTTAAACG	TTCTACAAAA	TTATCCCAT	CAAATTGGTT	AGTTAAATCT	7380
	GCAATAGTTAA	CGCCTATGCC	ACCACCAAGA	TTGATATGTT	TGAGTGGA AA	TCGATGTTTT	7440
40	TCAGACCATG	CCTTTGCTTT	TTTAAAATAA	AGTTTCACTA	CATCGACATG	TAAATTCGAG	7500
	TCTAAATTGT	TAGAAATAGA	ATGAAAATGA	AATCCATCTA	GATGAATCTT	TGGCATTGCG	7560
	AGCGCAGcTT	CAATGACATC	ATCAACTTCG	TCTTCAGAAA	TACCAAATTG	TGTTGGGCGT	7620
45	CCTGCCATAT	GCAACGTTGC	ATTGGGAAAT	GGTCCTGCTA	AATTAACACG	CAATAAAATG	7680
	TGTTGTGTCT	TATCTTCATC	TTCTAAGATG	GCATTTAGCC	GTTGTAATTC	ATGCATACTT	7740
50	TCAACATGAA	TACGCTGAAC	ACCTTCACTT	ACTGCATATC	TTAGTTCCTC	GTCTGTCTTA	7800
	CCAGGGCCAC	CAAAAATAAT	ATGATTTGCT	GGTTTAAAAG	CAAGACCTTT	TGCTATTTCA	7860
	CCTTGAGATG	CAACTTCGAA	TCCTTCAACA	TACTGACTAA	TTGTATCTAG	GATTTTTCGT	7920
55							

	TGTTGCAAAT GATGTTCCAG TCCGACTAAA TCATAGATAT AATGACAAAC TGGATGAGAT	8040
	TGTGCTTTTA ATTGTTCAAT AACAGGTTGA ACTATACGCA TTAGCCTTCA TCCCCTTTCT	8100
5	GTTTAGACGT CGCTAGAGAT GCACTTAAAT GGCGATATAT TTTTCCGCGA TCATCACCTA	8160
	AAATAAATGT TTGTACACCT TGTGCCTGCC ATTTTGCAAT ATCTTCATCT TCACGTGGTA	8220
10	ATGCACAAAA ATGTTTACCA TGTGCATTCA CAACTTCAAA AATATGTTGA ACATGTGATG	8280
	TTACTTGATC ATCACGCGTT TGCCATGGTA TGCCAAGTGA CTGCGATAAA TCTGCGGCAC	8340
	CTTCGACTAT CATGTCTAAA CCTTCGACTT GTGCTATATC GTCAATGGCC ATAACCCCTT	8400
15	CAACATCTTC TATCATGGCA ATCACCATAA TATGCTCATT AGCCATCTCC ATTGCATCAA	8460
	GTAATGGTGT ACGTCCAAAT CTTGCCATGC GACCACCATT CAAACTTCTT AATCCTTGCG	8520
	GGTAATAACG ACTTAATTTT ACAATATGCT CAACTGTCTC ACGATCTTTA ACGTGTGGCA	8580
20	CAATAATACC TCTCGCACCC ATATCCAACA CTTTAATGAT ATCTCTATCT ATCACTGCAG	8640
	TGACACGTAC AATTGGTATA ATATGCGCTG CTTCACTGTC ACGAATTAAA TGCGCTAGTG	8700
	TCTCATCATT AATCGCCACG TGTCTGTAT CAATCACAAC AAAGTCATAC CCGCTTGCTG	8760
25	CGATAACCTC GATCATCAAT GGGTCCGGTA TAGAATTAAA AATGCCATAA ACTGAATCAC	8820
	CATTGTTTAA TCTATGTTTC AGAGATAGTT GTTGTCATCAT TGATACCTCC TACACCTAAT	8880
30	GGATTTGTAA CATGATGAAT TCTTAACTCG GAGTCACTTA ATAATCGACG TGTCGTTAAC	8940
	TTTTCAACTT GAATCGTAGG TTCAAACAAA TCGAAATGTT GATAGTTATT CAACTCTGGA	9000
	AATGCTTCTT GATACGCCTC GATGATGCCT TTAACCCATT GCCATTGCAG CTCCTCATCG	9060
35	ATACCATATT GCTTTTCAAT AAATAAGATG ATTTCCGGCGA TATTAATAAA GAAAAATGCA	9120
	TCATGTAAAA AGTCGCGTAC TAAACGTTTCG TCATCTGTTT CAATAAATGA ATTACTATTC	9180
	ACTTTTATAT GTGCTTCTGG CATTGCTTTT AATGTCAGGT GTGAAGCAGC TTCACCTAAA	9240
40	TGctCACGCT TAAAACGAAC ACCATCATGG AAATCTTTTA AGGCAATACG TGTAGGCCAA	9300
	CCATTTTCAT GAATGAGCAT CATATTTTGT GCATGCGATT CAAAGGCAAT ACCGTGATAA	9360
45	TAAAGCATAT GAATCATTGG ACGAATCGCT ACAGCTAAAA ATTGCTTTGT CCAAGCTTCA	9420
	GAACCATATT GTTTAATCCA ATTTTCAATG AATGGTACAC CATCCTTATC ACTTGCATAA	9480
	AGTGCAATTA ATGGTATCGC ATCCTCTTCA TCGATTAAAC TATGATATAT ATTTTCACGC	9540
50	CATATAACAC CTAACGCACC ATAAACTTGA GTTTGTTTAT AAGGCGAAAG TTGTGTATTT	9600
	AAATAAGACT GTCCTAAGAC TTCCCCTAGA AAAACTGTCT TTAATTCATC TTTTAAATAC	9660
55	ATATCTTGTT GCTGTATCTG CTTTAACCAA TCCGTAATTT GCGCTGCATT TTCAATTGTA	9720

	TATTTTGTGCG TGTCTATTGG CGACATCGTA CGAATCGATT GTTGAGGGTG ATATAGCTCA	9840
	TCACCTTTCCC CTAACCATAG TACTGTGCCA TTAAGCCTTT CTTCAGCCAA ATCAACTTGG	9900
5	ATGACATGTT CAAACTGCCA TGGGTGTACA GGTATCATCT CAACATCATT TACATGTTTG	9960
	CCAGATGCTT CAATTGCTG TACAAAATGT TCATAAGTCT TATCGCCAAC TTGTTGACGT	10020
	AACATTTGCT TAACTACAAC ATTTCTTGAT ACCGTCGTTT CTACTTTATC TTTGTGCGATA	10080
10	GCTAACCCTT GCAGTTTAAAC GTTTGGTACA AAATCAGGAC CAAATTTCAA ATTATCACTC	10140
	AACGTAAATC CTAAACGTGA TTTGTAACCT GGATGATACT GATGCCCTTC CATCGCATAA	10200
15	AATTCATAGT CGTTAAATGT CTCAGGTGTT GCTGGTGGGT TTGATTCTCG ATACTGCATA	10260
	CTTTGCGTAT CTTTTAATTC TGTCTGTAAT AACTCGACAA TAAATTGTTC TAGCTTTTCA	10320
	TCATTTTITAG GAAATGTAAA TACAACCTCT CTCAATAATT GTGTATAGTC TGTGTTGTA	10380
20	TCTGCCTCAT CTCCTACGAC ACGCTCAATT GGTGATGTGA TACGTATACG ATCAAAGCTA	10440
	TGTGTCTTTT CAGCAGTAAA ACGATACTCT GAATCATGTC CTTCTATTGT AAAATGACCG	10500
	ACACCGTCTT GATATGACGC TTTATACACA ACAATATTCT CATAAATAAG TGATGATACC	10560
25	AGTTGGTGCA TCACTCTAGT CTTTACACGA TTAAGAAATG TTTGATTAC AATACGATAC	10620
	CTCCTTGTTA TGACAAATTG GATTTGGTAT ATGTGTATAA ATAGGGTTTG CACCACAATC	10680
	ATTCAATTTA CTCATCAAAT TCGCTTTAGC CGCAATGGTC GCGCTTTGAT ATAAATCTTC	10740
30	TACACAGTCA ACAAATACTG CGTTATTGCG GTATTCTTTT TTCCAAGTCA TAAGACGATG	10800
	CGCTACAAGT TGCCATAACA CAACTTCATT TCTAGTCGCT TTACCAATAG TTGATACTAA	10860
35	ATGTCCTAAG TGATTACTA CAACGTAATA TTTAAGACGA TGCCATGCTT CATCATGTGC	10920
	ATATACAACA GGGCTTGATG CTGCCACAAC ATTTGGCACA AGCTGTTTTT CAGTAGCAAT	10980
	CGTTCTAGAT AGACAAATGC CTTCAAGATC TCTGACAAAG CATACGTCGG GTATGCCATC	11040
40	TTTTAATTCA ATTAATGTAT TTTGTACATG TGCTTCTAGA CTAATGCCTG TGTTACTAAA	11100
	CAGCTTTAAT ATCGGCAATA ATGTACGATT CAAATAACAT TCAAGCCATG CTTCTGGTGC	11160
	TAAACCACTT TGCTCAATCA CTTGTGATAA CTTAGACATC GGTGAATCAG GCATCGTTTC	11220
45	AAATAATGAC GCCAATACAT GAATATCTTT ATCAGCATGG TAATTCGGTA TCCCTTCACG	11280
	AACAATCATG GCACTATTTG TTAATAAATC CATTTTCAGGT TCAACTGTTT GCCCTAATGG	11340
50	ATTCGGTAAC AATGCACGAT ATCCTTCTTC AAACATCAAT TTAAATGGG GTGTTTCAAC	11400
	CTCATCTTTG ACTGATGCGA TAACTGCGC GGCATCAATT GTCCGTTCAA TCTGTTCAAG	11460
	GTCATTGCGTA CGTATAAAAT TAGTGATTTT AACGTGTATC GGTAATTTTA AATAAATGTT	11520

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

GCCAAGGTCT TTTATTAAAC CTTGTTCACT ATATTGCATA TACTGTGGAT GCTGTGCGAA 11640
 CACATTGATT TGATAAGGAT GTGTTGGTAA TAAAATAAAA TCTTTGGGTA TCTCTGATAT 11700
 ATCTATGTCT GCTAATTGAT ACAACACTTT CTCAACCTGA TCTTCTTTAC CTTCTACATA 11760
 GCGCGTGAGC AGAACATCTT GATGCACAGC TAAATAATGC AATTGGAATG ATGTATGACA 11820
 TTCGGGTGCA TATTTCTCTA AATCTGCTTC TGAAAACCCA CTTGCACTCT TAGGAGTCGG 11880
 ATGAAATGGA TGACCTAAGT ATAAAGATTG TTCTGAAACG ATATAACGAT CCTCTACGTA 11940
 GTCTATTGTG TTACTTTGCA AATAACGTGC CGTGCGATGA ATGCTATTAT CGATGTCAGA 12000
 CATAATTTGC GCCATATGTT GTTGCACTGC CGTTTGATTA TCTGCACTTT GAGCCATATG 12060
 TTGCAAAATA CGCGCAATTG CTTCTTTTATA AGTTGTTATT TTTTACTTTT TTCCATCGAT 12120
 AAGCCATACC TCTGGATGAT ACATATGATG CCCCATCGCA GACCAATAGC GAAATTCACC 12180
 CGTTAAAGTT TCGAGCTCTG ATAATTGTAT AGACCATTGA TGATTTTGAG GTGGTACTTG 12240
 ATATAAATTT TCTTCTCTAA AATATTCATT TAAAATGCGT TCGATAGCCG CATACGCTGC 12300
 ATGTTGTATT AATTCTTTAT TTTGCACTTT TTTGTTTCAA CTCCCATAAT TTCATTAATG 12360
 TGTGATCGTT GATTTGATTA GTGATGGTTG AACAAATTAA AAATAAACTA CTTACTGCAA 12420
 ATACTACGCC CATAACGATA AACGTAGTAG CTGGTGTAGT ATAACTTGTA ATGGCAGCGC 12480
 CACTaAGACT GCCAATAATT TGACCAACAA CTAACATACT GTTCGTCGTT CCAACAAATG 12540
 TGCCTTTAAG TTGTTGATGA CACGCATTCA CGACAACAAA CATGACACTT TGAATCAATG 12600
 CACTATATGT TAATCCTTGA AGTATTCTTG CAGCCATTAA AAACCTCTATA TTCGTCGCTA 12660
 AACCTTGCAG TATCGCACTA CAACCACATG CAATCGTGGC AAATATATAT ACTGATTTAA 12720
 CATATGATTT ATCATTAAAG CGTCCCCATA AAGGCGCGCT TAATATCGAA GCCGTCCAAA 12780
 ATGCGGACTG TAAAAATCCA ATCACACTAC GGTCACTCTAT CGCTGTATGA TTCACTGATG 12840
 AAGCAAGTGG TGATAATGCA GTTAGCATGC CATACATAGC AAAGTTTGCT AAAACGCCAA 12900
 CGATAATAAA TCGACATGTT TGTGTGTGTC ATAATAGACA TTGAAATGAA CGGCGAATAC 12960
 CTTTATTAAT ATTTGGTGTG TGTGATTTTG GCATATGTGT CGTTTCAATC AATTTTAATG 13020
 CACCGAAAAT ACAGACAATA AAAGTAATAA CGGCAATACT CATCAGTAAC GCACTAAAAC 13080
 CTAATATCGA AGCTGTAACA CCGCCAATTA ATGGCCCCAC AAGAGACCCT GCGCTGACTG 13140
 AACTTTGCAG TCTTCCTAAT ACCTTTCCAC GATCTTCAGC TGGCGCCTCT GCACTCGCAA 13200
 ACGCACTTGA TGCATCAACA ACACCACCAA ATAGTCCCTG CAATAACCTC ACAAGTACAA 13260
 ACTGTAATGG TGTCGTACAC AATGCCATTA AAAATAAGCA TACCGCCAAA CCAAGTAACG 13320

	ctATCATCGT	CGTTACAGCT	GGAGCAGCAA	TCGCTATACC	ACTCCACAAC	TGTATTTCTA	13440	
5	CGACTGATAG	ATTTTGTAGT	GATGCCATAT	AAATTGGCAA	TAATGGCACA	AGTACTGTCA	13500	
	GTCCAGCAAT	CGCTATAAAC	TGACTGAGCC	ATAAAATGCG	AAAGTTACTG	CGCCATATAG	13560	
	ACTGATTAAT	CATATGTCAC	CATTGGATTT	GGTACGGTAG	TTAAACCTGA	AGGCATACTA	13620	
10	CCTCCACCAC	TATCACGTTG	ATATAGCAAT	GGTAATAAAA	TTTGTTTGAA	TGGCCACGTC	13680	
	TGTTTATCAA	ATAAAATGTG	TCTGACAGCT	AGCTGATCAG	TTGTAACCCA	GGAAATAGTT	13740	
	GCCACTTCAT	TTTTTAAAAT	TTGTTTTAAC	AACGACATAA	GTTTCATGCTC	ACTTACACCA	13800	
15	AATAAATCTT	GAATTGCATC	AATAATGGCA	TATAGATTTA	CCGATACAGC	TAATGTTTGA	13860	
	AAATAAGCAA	AGAATGTTTC	CAAATCCTCA	TTAATTAGCG	TATTAGGTGT	ATCTTCTCTG	13920	
	ACGACATACT	TCGGCAATGA	AAGCTGATGT	GCTGTTAGCC	ATGGTTTATA	AATTCTGACA	13980	
20	GTATCATGAT	CACGTAACAC	GCATTTTTGT	ACACGTCCAT	CTTCAAATGA	CAACAATATA	14040	
	TTTTGACCAT	GCAACTCTGG	TAATGCGCCG	TATTGCATAA	ATGATAGTGT	TACCTTTAAA	14100	
25	AAGACTTGCG	CGATATCTTC	AAATAACGTC	ATGACATCAT	TTTTAGAAAT	ATTATCTTTT	14160	
	CCACAAATCA	TTTGATATAA	AGTGCGATCA	TTTGCCGCGA	GTGCTGCCAT	TGACACTAGC	14220	
	TGTTGCGTAT	CATTTTTGGC	TAGCACTTCG	GGATACTTTC	TTAGCTGAAC	AGTTAGATGA	14280	
30	CCTAATTGAT	CTTTGAAAAT	ATCATTATCT	TGACCCATAT	ATGACCACCA	AGCTGTTTCA	14340	
	TCACAAACCA	TGACATACTT	AGCTAGTGCT	TCATCTTTTT	CTATAAGCTG	ACGTAATAAT	14400	
	TGTTCTGCTT	GTTCTCCGTT	TTTCATGTAA	CGCGTAGGCG	TTAGCCTTAA	TGCGCCTAAT	14460	
35	GACTGCATTG	CAAATGGTAC	TTTGACATGG	TTATACGGTG	CGCCAATATC	AATTAATGAA	14520	
	CGCATACTTG	AAGACGACAG	ATAATCTCCA	AATTTTAAACG	GTAATAGTAC	AACCAACTTT	14580	
	TCACTAATCT	CTTTGCAAAA	GACGTTGCGC	AGAATATGCT	GATATTGCCA	AGGATGTACC	14640	
40	GGAAATAGTA	CATAGTCATC	TATTGATAAC	CCTTGATCAT	TTAACATGTC	TGTCGCTTGT	14700	
	TCTTTTATAG	GTA	CTACTGTCAA	ATTTTCTAAT	TCATCGATAT	TTGCAGTATC	GCCATGAATC	14760
45	ATATGTGTCT	TTTTAACTGC	TGCAACCATT	AAAGGAAATG	ATTGATTTAA	TTCAGCTTGA	14820	
	TACACTTGAT	AATCCGCTTC	TCTTAATCCT	CTTTTTTCTT	TAGCTAATGG	ATGAAATGGA	14880	
	CGATCTTTTA	AACCTGCAAA	CTGCTCTGAC	ATCACAAAAG	GATGTGACGC	TAAATCTAAT	14940	
50	TCTGATAATT	GTTTAGCAAG	CTGTGTGGCA	GCAGTAGTCA	GTCCTTCTTC	AACGCGAGCC	15000	
	ACTTCCCAT	CATGACTTAG	ATCACAATTC	ATATTAGCAA	TTGTTTGCCA	AAATTCAGCT	15060	
55	GCCGTTAAAG	GTTGCTTAGA	CACCCCTCCC	TCTATCGTAA	TTGGTTGTGA	ACTTTCGTAA	15120	

TATATCAAAA GCGTTTGTCC GTTTTCTTTA GTAATCTCAC TATTCGATAC AATTCCGGCT 15240
 ATATCTTCAA ATAATAATGC ATCAACTAAA TCTCTTAATA TTATCGCTTG TGCTGTATTG 15300
 5 ACTGCTGTAT GATTCTGCAA TGTTCAGACA CCTCGCATTG TTAATATAGG TTCAATGTTG 15360
 TCCCAATATT TTGTTGTTGT GCCTGTTGAT AAATAAAATA AGCACTTGAA ATATCTTCGA 15420
 10 TAGCCATACC CATCGGATTA AGTAATATGA TCTCATCATC GTCTTCACGT CCTGGTATGT 15480
 CACCTGTCAC AAGTTGTCCT AGTTCAGCAT GAAGAGCTTC TTTGCTGAAT TTACCTTCTA 15540
 ACACCAATTG GTTAATAGTT TTCTTTTCTC GATTACATTG TGACCAGTCA TCTACTACGA 15600
 15 CTTTGTGAGC TTTAATAAAG ACTTCTTTAT GCACATCCAT GATAGAAATG TTGCTAATAA 15660
 ATGCACCCTT TTGTAACCAA TCATATTCAA TGTATGGTTG ATCCGTTACG GTACATGTAA 15720
 TGACTACTTC ACCATTTGAT ACTGCTTCTT TAGCATTTTC TGTCGCAATA AAATTAATTT 15780
 20 CCGGACGCTG TTGTTGCCAT CTATCAACAA AGCGTGCACA TGCTTCAGAG AATTGATCGT 15840
 AAACAAACAC GCGTTCAATA TGATCGAATT GCTCTAACAT ACTTTGTAAT TGCTTGTCTC 15900
 CGATTAGCCC GCATCCAATG ATTGTTAAGT CTTTAAATCC TTTTTTAGCC AAATGCTTTG 15960
 25 CTGCAATCAC TGAAACTGCT GCAGTACGCA TACTACTAAT TAAACTTGCT TCCATAACTG 16020
 CAATTGGATA ATTCGTTTCT GGATCATTCA AAATAATGAC GCCACTTGCA CGCTCCATAT 16080
 30 TACGTTTCGA TGGATTGTCG TGCTTACTAC CTATCCACTT AATACCTGAA ATTGCGTGTT 16140
 CACCACCGAT ATGACTTGGC ATTGCAATAA TTCGATCTGC GATGTGTCCA TTTTCAGGAT 16200
 CCTGTCTTAA ATACGGCTTA AGCGGTTGTA CAAAATCATT GTGCGCATGG GCTGTTAATG 16260
 35 CTTCTGTTAA TGCGTCCACA TAAACTTGTG AATGATTACC TCCCGCTTGT TCAATATCTG 16320
 ATCTATTAA ATACAACATC TCTCTatTca TTCTGaTTTA ACTCCTTGTC TTGATTTCAT 16380
 TTTTCTAAC CATGTATCTG AATAAACTAA ATCTAAGTAA CGATCGCCTC GATCTGGTAA 16440
 40 AATCGTGACA ATTGTTGCAC CTTCTTCAAT TGACGTTATC AACTGCTCAA TCGCTGCAAT 16500
 AATCGAACCT GTTGAACTC CGGCAAATAT GCCTTCATAA TCAATCAGTT TTCGACAGCC 16560
 45 CAAAGCAGAT TGATAATCAT CTACATGGAT CACTTGATTA ATTTCTGATC TATTCAATAT 16620
 TTCGGGTACA CGACTAGCAC CGATACCAGG TAATTCTCTA TTAATAGGTT TGTCACCAAA 16680
 AATGACTGAC CCTTTCGCAT CAACAGCAAC AATTTGTGCG TTTGGATGCA CTTCTTTTAT 16740
 50 TTTTCTACTC ATACCCATAA TGCTACCTGT CGTGCTGACT GGCGCGACAA AATAATCTAT 16800
 AGGTTGCTTA ATTGTTTCAA CAATCTCTGT GCCTGCACCA TGATAATGGG ATTGCCAATT 16860
 TAACTCATTG GCATATTGAT TAATCCAATA TGCATCGTCA ATAGTGGCTA ACAGTTCTTG 16920

55

TACATTGGCA CCATAACTTT TAATAATTTT CAAATTTGTT GGTGATATTT TAGGATCAAC 17040
 AACACACGTG AGTTTTAATC CCTTGATTTT AGCTATCATT GCCAACGCAA TGCCTAAATT 17100
 5 ACCAGAAGTA CTTTCAATTA AATGTGTATT CTCAGTGATT AAACCATGTT TAATACCATG 17160
 TTCAATGATG TACTTGGCAG GTCGATCTTT CATGCTGCCT CCAGGATTCA TATACTCTAA 17220
 10 CTTTGCAAAC ACTTCATGTT TCGGAAATAG TTGATGAAGT TGAACCATAG GTGTTTGCCC 17280
 TACAGAATCT AACAATGAAT CGTGACATG 17310

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5423 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

ATACTAGTAA GCGCATCGGT TATTGACATC GAATTCAACT TTAACAGTTT TCATGTTTCGG 60
 25 TGATGTTTCa ATAGAATGTG TGTGTTGTAC TTGCGCATTT ATATTTCCAC CTAAATTACT 120
 TAAGTTTCCT GTAATACTAG AAATGTCAGG TCGTTTTAAT GTAGTTGAA ATGCATCAAC 180
 30 TACTTTATCT GCAACATTAG AACATTACG GATAACTTTA CTTGAATGAT TATCTATACC 240
 TTTAACGAAA CCTAACATTG AATACATACC AACATCCATG AATTCACGTG AAGGTGAGTG 300
 AATACCTAGC GCTCTTTTGG CTGCATTAA AGCACCTTTT GCTACACTAG CTGCTTTTTTC 360
 35 AGCTAAGTCT CTAGCCATAT TACCAATACC TCTCATCAA CCACGGATCA TATCAGCACC 420
 TGCTGATACA AAGTCATCCA CAAAGCTTTT AACTTTATTT ACTGCATTG TCATACCTTG 480
 ACTAACTTTG TTTACAACAT TAACGAATCC TTGAACAACT CTATTAACAA rGTTAATTAG 540
 40 CGTACTTGt ATAGTAGATA CCCaTnGCAT ACCTTTAGTG ACmATGAAGT TCCAAGCTTG 600
 AGACATTTTG TCTGATATAG TTGAAACAAC TTGTGTGAAT ATGCTTACAA CTTTATTCCA 660
 45 AATTGTCGTT AATATACCAG ATAAGAACT CCAAATCGTA TTCCATATAT TAGAAATAAA 720
 ACTCCATGCC GCTTGTAACG CAGTAGATAT AGCTGTAGTG ATAGCGTTCC AAACCTTAGT 780
 TGCCACAGTA ACTATAGTGT TCCACAACGT TTGTAAGAAC GTCCAAATAG CGTTCCAAAT 840
 50 TGTTATTGCG ATAGTCATAA TTGTGGTAAA CACTGTAGTT ATTACAGTGA CTAACAAATT 900
 CCAAATCGTA GTAGCGATTG TAATTATCGT ATTCCAGATT GTACTTAAGA ACGTCCAAAT 960
 AGCTGTCCAT ATCGTCATAA CTATTGTCAT TATCGTCGTG AAAACAGTTG TAATGATTGT 1020

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATAAGCGACT ATTTGATTCC AAACAATCAT TATAAAATTG TAAACATTTCG ATACTGCTGT 1140
 AGTGATAGCT GTTAAAATAG CATTCCATAC AACCGAAGCT ACAGCTTTTA ATACATTCCA 1200
 AACATTAACC ATAAACGTTT TTATCGCATT CCAAGCATTT ATAATAAAGT TTCTGAATCC 1260
 TTCATTTTTA TTCCACAATA AAACGAATAT AGCTATTAAT GCAGCAATTA CACCAATTAC 1320
 TATTGTTATT GGACCGCCTA AAATACCAA CACAGTTACT AGTCCTGTGA TAGCATTCTT 1380
 AATTAATCCA ATCTTACCGA ATAACAATTG GAATATAACT GATATAATTT TTAATGGTCC 1440
 TTTTAATAAC ATGAACGCAC CTTTTAAAAT TGTTAATCCC GCTCTTAATA AACCGAACTT 1500
 ACTTACTAAT GCAATGTTTC TACCTATTAA TCCGCCACCC ATAAAGTTAG ATACAGCAAG 1560
 AATAATCGGT ATTAAAAATC TAAATGCACC AACTAAAGTT ATAATGACAC CAACTAATTG 1620
 TGCTGTAGCT GGATGCGCCT CAAACAAGTT AGCTATCCAA CCAGTTATTG CAACTGCAAC 1680
 GCGTAATACT GCACTAGCTA TAGGAGCCAT TGCTGTTGCG AATGCAATTA ATCCTCTTGC 1740
 GATGTTTCCA ATCAATTGCA TTATTAGTGG TCCATTTGTT TGTATATAAC TGACAAAGTC 1800
 TTTAAACCCT TGAGATTGTC CTAATTGTTT AGACCATTCC CTAAACTTAG CTGTCATTTG 1860
 TTCAAGAGAT TGGAAATATG CAGTTGATGA TCCGCTGAAT GCATTCAATCA AATTGTTAAT 1920
 TCCAACGAAA ACATTTTGA AAATATTACC AATGATAGGT AAGTTTGTTT TTGTGTATTG 1980
 AATAAACGA GTTATCGAAT TTTCTCCAGC TGCATTATTA GCCCAGTTAG AGAAAGATTG 2040
 ACCTAATCTA TCCAACCAAT CAGCCGACCA TTGAAACAGT GGTGCTAATT GCGTGAATAC 2100
 ATTGACTAAT CCGTCACCAA AACCACCTGC AGCACTTAAT AGCTTGTTAA ATACCGAAAC 2160
 ACCCGTTGTA TTCATCATAT TAAAGAATCT TGAAGCTACA CTGCTATTTT CAGCCCATT 2220
 AAGCACGCTT TGAGACGCTT CTTCCATTCC TCTTGAAATA CCACTAAAAA ACGGTTGTAA 2280
 GCTCTGCATT GCAGTTTTAA CAGTATTTAA ACCATTTGCA AGAGTTGTGA AGATAGCGGA 2340
 TTGATTTTGC TTTATAATAT CAGTCCATGC TGACTTTACG CCATCTAACG CTTTTTTGTA 2400
 TTCGTTTGTT GCTGAGCTAG CTTGTAAAGT GCCATCATTA AGCATCTTTA TAGCGCTGAT 2460
 AGCCATTGCG CCAAACGCTA CAAATCCTGC TCCGCTATT GCTACGGCAC CACCTAAAGC 2520
 AAGTACACCA CCAGTTAACA CTTTGATAGC GTTTAATAGC GCAAATACTA CAGGTACTAC 2580
 GCTCGCTATT ACAGGTATTA AGATACTAAA AGATGATGTA AGTAATCCAC CAACCATATT 2640
 AGAACCTACA GTACCGAACA CACGGAACAT ATTAGCTAAA TTCCCCATCT GTCTTTGAAA 2700
 ATTGTCATTT GCTTTTATTA TGTAGGCATA AGCTTTCTTT AAACCATTAG TATCGACATC 2760
 TACCTTTGTT GTTTTTTGT TCGGCAATGC GTCTAATGAT TTTTAAACG CATAAATAGT 2820

	AAGTTCTTCT	TTAGTACGTT	TGATTTTAGA	GTTAGCAACA	CCATTGTCCA	CGTCTATAAT	2940
	AGCTTTGGCT	TTAGACCTAT	TTAATGCTTC	GAGACTAGCT	TTAGATACTT	TTAACACTCG	3000
5	ATTGAATTTA	CTGTTATCTG	CATTGACGTC	AATATTGACA	CGTTTCTTTT	CTAATTCTGA	3060
	TAATTTAGCT	TCTGTTTCAG	CGATATCTTT	AATCAACTTT	TGTTTTTGCA	ACTTAACTTC	3120
10	TGGTGTAAC	TCTTTAGAGT	TTAGTTTGTC	TAGTTCAAAA	TTCGATTCTA	GTACCTTTTG	3180
	TTGTAAATCT	TGTATACTAG	CATCTAATTT	AGCTTTTACA	TTTTTGTTAC	TAAAGGCATC	3240
	TAAAGACTTT	TTAGCAACTT	TGATAGTTTT	TTGTAAATTT	TTATCGTTAG	CGTTTAATTC	3300
15	AACATCTTTA	GTTTGATCTG	CTACTCGTTT	AAATCTTTGC	ACAGACTTAA	CCGCACTATC	3360
	AATTTGCCTT	TTGAATTTGG	CTACACTAGC	TTCAATAGTC	GCTTTAATTT	TATATTCCGT	3420
	CACATTAACA	CCTCTCTTTC	TATTGCTTAT	TAAATTCTGC	TATAACTTTA	AAGAATTCAT	3480
20	TATTTGTGG	TTCGTATTCA	TCACGTTTCG	TACTAAATCT	TATATCTTTA	CCTTCGTTAA	3540
	GCCGTGGAT	ATTTTCTTCA	TAAGGCAATA	CGTCGTTTGC	ATTGTTAAAA	ACATATTCCT	3600
25	CTTTAGGTTT	ATTTTCTGTC	CCAACATTTT	TAGTAGCTGC	AGCATCACGA	ATAGCAAACG	3660
	CAAGTTTGTA	ACGTTCGAAT	TCTTGGGTTA	GCATTTTATA	CTCTTTCGCA	TACATTCGAT	3720
	AGTTATATTC	TGTTAATGTC	ATTTGCTCAA	TAACGTTCAA	ATCTGTAATA	CCAAGTGTG	3780
30	ACATACAAGT	TATAACGATT	CTGTCGTAAG	TTATTAGGCT	TCCGCTGGTT	TTTCTTCCGT	3840
	TTCCACTACT	TCGACTAGGT	TTCCGGTCAT	AGGTCGCTTT	CCCAACTCCG	TTAAATATC	3900
	CGAACCGAAT	TCTTCTAGTC	CGATATTTTC	TGCGATTTC	TCTAATGCTT	CATCAATGTT	3960
35	ATTAATAGTA	ATTGCTTGTT	TTTTTAAGTG	AGATGTAGCT	GCGATTAAAA	CTTCGCCAAT	4020
	CACAACCGGA	TTTCCACTTT	CTAAACCTAC	AGGCAACATT	GATACACCTT	GACCGATAGA	4080
	AGCTTGTTC	ACTTTTAAAC	CTAATCGGTT	ATCGATTCTT	CTTAAAAATT	TAAAACCAAA	4140
40	ACTTAATTCT	AATGACTTTC	CGTTAATTTT	TACATTCATA	ACTTAAAATC	TCCATTCATA	4200
	ATTAATTTAA	ACAAAATAAA	mAAGCTTAAC	GCCCTATTTT	TATACCTCTC	TTGGTGCAAC	4260
45	CGGTGGTGAA	TCTACTTTAG	GTTGTGGAAT	TGCTGTTAAA	TCTTCGCCAG	TTAATGCATC	4320
	TGCTTTTGTA	GTGTCGTGGA	ATCTGTATcC	AGTCGCCTTA	AGTTTCTTTG	TTACAGCCTC	4380
	AGGTAGTGTT	GCAAATCCAC	GTTGGAAACG	ACCATTCACT	CCATATTCAT	ATTCATATTC	4440
50	ATCAATACCG	TTAGCTTCTG	CTTTTAATTC	AAATTTATTG	TGGAAACCTT	GGAAATATTT	4500
	CGCTTTAAAT	TTAGCGGAAT	CCCCATTTT	GCCTGGTATT	CTACTTTCAA	CTTCCCAAGC	4560
55	TTCATACAAT	ACGCGATCTA	CAACTGCATC	TTCAATTTCA	TCTGCAAAAT	CGTCACCATA	4620

GTCCATTGTA TCCTCTGTAT CTGTATCAGC TTCATGTGAT AAGCCGTATT CAGTTAAAAA 4740
 AAGCATTTTA GTAGCATCTA CTTTTTCGCC AGCTTTTCTA AATAAAATAA TACGATCATT 4800
 5 ACTATTTTTC ATATTTGCCA TTCAATATTC CTCCGTTTTT TAAAATGTTT TGTAAGATAT 4860
 CGTTACTGAT GTGTGTAGCA ATTCTTGATT GGTAAGTATCA TCAACTAACT GTGTGATGTT 4920
 10 AGTATCTTCT TCTTCAAAGT CATAATCGTT TGTTTTAACG CTAGGTGTTA AATCATCAAT 4980
 ACATCTTTTA ACAAGTCCGT CATGATGTCC TAAATCATCG CTTACACTCC AAATATCAAT 5040
 AACTAAATTC GATCGCCAG AATAACTATC AAACGTGTAC TTACTTCTAT TTGACTCCGG 5100
 15 CATTTTTATT ACAAAAAAAG GATACGGAAT CTCTTGTTGC ATCTCTTTAC GAGAAATAAC 5160
 AGGGAATCCA TATCCTTGTA GCGTTTCATA CGCTTTATTA TAAAGTTGTA AGTTCGGTGT 5220
 CATGCTTTTA TCTCCTATTC AAACAACGCT TTCAATTCTT CTACAGTTGA TTTCTAATC 5280
 20 ACTTCGTATA CCGGCCACAT AAAAGGTTCA GCCTCCATGT ATCGAGTACC AAATTCTAAG 5340
 AAACCACTAT AAGCTGCGTG CGATGTGATA GTGTATTGCA AATCGCCAGT TTTTTATAT 5400
 CTGATATTGC GTGATAAATT ACC 5423

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6251 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AAACGCAGAT GTTCAATTAG AACCAGTCTA TCGTATTAAG GAAGGTATTA AACAAAAGCA 60
 AATAAGAGAC CAAATTAGAC AAGCGTTAAA TGATGTGACA ATTCATGAAT GGTTAACTGA 120
 40 TGAACCTAAGA GAAAAATATA AATTAGAGAC CTTGGACTTT ACTTTGAACA CATTACATCA 180
 TCCTAAAAGT AAAGAGGATT TATTACGTGC TCGTAGAACC TATGCATTTA CTGAACTGTT 240
 TTTATTTCGAA TTACGTATGC AATGGCTAAA TAGATTAGAA AAGTCATCTG ACGAAGCAAT 300
 45 TGAAATTGAT TATGACATAG ACCAAGTTAA ATCATTATTAT GATCGTTTAC CTTTGAAGT 360
 AACTGAAGCA CAGAAATCCA GTGTTAATGA AATTTTTAGA GATTTAAAAG CACCAATACG 420
 50 TATGCATCGA TTACTTCAAG GTGATGTAGG TTCAGGAAAA ACAGTAGTTG CTGCAATTG 480
 TATGTATGCG TTAAAACTG CTGGTTATCA ATCAGCATTG ATGGTACCAA CTGAAATTTT 540
 AGCAGAGCAA CATGCTGAAA GTTTAATGGC TTTATTTGGA GATTCTATGA ACGTTGCATT 600

	TACGATTGAT TGTTTAATTG GAACCCATGC TTTGATTCAA GATGATGTGA TTTTCCATAA	720
	TGTTGGTTTA GTAATTACAG ATGAACAACA TCGATTGGT GTGAATCAAC GCCAGCTTTT	780
5	AAGAGAAAAA GGTGCAATGA CGAATGTGTT ATTTATGACA GCAACGCCGA TACCAAGAAC	840
	ACTAGCAATA TCAGTTTTTG GTGAGATGGA TGTGTCTTCA ATTAAACAAT TACCAAAAGG	900
10	TCGTAAACCT ATCATTACTA CTTGGGCAAA GCATGAGCAA TACGATAAAG TTTTGATGCA	960
	AATGACCTCA GAGTTGAAAA AAGGTCGTCA AGCATATGTC ATTTGCCCGC TAATAGAAAG	1020
	TTCTGAGCAT CTCGAAGATG TTCAAAATGT TGTCGCATTG TACGAGTCTT TACAACAGTA	1080
15	TTATGGTGTT TCCCGTGTAG GGTATTGCA TGGTAAGTTA TCTGCCGATG AAAAAGATGA	1140
	GGTCATGCAA AAGTTTAGTA ATCATGAGAT AAATGTTTTA GTTCTACTA CTGTTGTTGA	1200
	AGTAGGTGTT AATGTACCGA ATGCAACTTT TATGATGATT TATGATGCCG ATCGCTTTGG	1260
20	ATTATCAACT TTACATCAGT TACGCGGTG TGTAGGTAGA AGTGACCAGC AAAGTTACTG	1320
	TGTTTTAATT GCATCCCCTA AAACAGAAAC AGGAATTGAA AGAATGACAA TTATGACACA	1380
	AACAACGGAT GGATTTGAAT TGAGTGAACG AGACTTAGAA ATGCGTGGTC CTGGAGATTT	1440
25	CTTTGGTGTT AAACAAAGTG GaTTGCCAGA TTTCTTAGTT GCCAATTTAG TTGAAGATTA	1500
	TCGTATGTTA GAAGTTGCTC GTGATGAAGC AGCTGAACTT ATTCAATCTG GCGTATTCTT	1560
30	TGAAAATACG TATCAACATT TACGTCATTT TGTGAAGAA AATTTATTAC ATCGTAGTTT	1620
	TGACTAATTG CCATGCTGAT TTGTCAATTT GAGTGCAACa CTTGTTAAT TGAGTGATAT	1680
	GACACTTGAA CTATTTAAAT GTAAAGTGGT ATTTTAACAA TTTATAAATT TTCGACTAAA	1740
35	TAATAGCTAA ATATTACAGT TATTTGTTGA GTCGGTTAAA TAGAAAGTGT TATGATATGT	1800
	GAGGAATGTT TAAGACTAGG TACTAAAAAA TGAGGGGTGA GACGTTGAAA CTAAAGAAAG	1860
	ATAAACGTA AGAAGCAATC AGACAACAAA TTGATAGCAA TCCCTTCATC ACAGACCATG	1920
40	AACTAAGCGA CTTATTTCAA GTGAGTATAC AAACAATTCG TTtAGaTCGC ACTTATTTAA	1980
	ACATACCAGA ATTAAGGAAG CGTATTAAAT TAGTTGCTGA AAAGAATTAT GACCAAATAA	2040
45	GTTCTATTGA AGAACAAGAA TTTATTGGTG ATTTGATTCA AGTCAATCCa AATGTTAAAG	2100
	CGCAATCAAT TTTAGATATT ACATCGGATT CTGTTTTTCA TAAACTGGA ATTGCGCGTG	2160
	GTCATGTGCT GTTTGCTCAG GCAAATTCGT TATGTGTTGC GCTAATTAAG CAACCAACAG	2220
50	TTTTAACTCA TGAGAGTAGC ATTCAATTTA TTGAAAAAGT AAAATTAAAT GATACGGTAA	2280
	GAGCAGAAGC ACGAGTTGTA AATCAAACCTG CAAAACATTA TTACGTCGAA GTAAAGTCAT	2340
	ATGTTAAACA TACATTAGTT TTCAAAGGAA ATTTTAAAT GTTTTATGAT AAGCGAGGAT	2400
55		

	TTAGAAGCCG TACAAAAGGC TGTGAAGAC TTAAAGATC TAGAAATTAT ACTTTTCGGT	2520
	GACGAAAAA AGTATAATCT GAACCATGAA CGAATCGAAT TTAGACATTG TTCTGAAAAG	2580
5	ATTGAAATGG AAGATGAGCC TGTTAGAGCG ATTAACGTA AAAAAGATAG CTCAATGGTA	2640
	AAAATGGCTG AAGCTGTGAA ATCTGGTGAA GCAGATGGAT GTGTGTCAGC AGGTAATACT	2700
10	GGTGCTTTAA TGTCAGCTGG TTTATTCATT GTTGGACGTA TTAAAGGTGT AGCTAGACCG	2760
	GCTTTAGTAG TAACATTGCC AACGATTGAT GGAAAAGGTT TTGCTTTTTT AGACGTTGGT	2820
	GCAAATGCTG ATGCTAAACC TGAACACTTA TTACAGTATG CGCAACTAGG GGATATTTAT	2880
15	GCTCAAAAAA TTAGAGGTAT TGATAATCCG AAAATCTCAT TATTAAATAT AGGAACCGAG	2940
	CCAGCTAAAG GTAATAGTTT AACGAAAAA TCATATGAGT TATTAAATCA TGATCATTCA	3000
	TTGAATTTTG TTGGGAATAT TGAAGCGAAG ACATTAATGG ATGGCGATAC AGATGTTGTA	3060
20	GTTACCGATG GCTATACTGG GAACATGGTC CTTAAAAATT TAGAAGGTAC TGCAAAATCA	3120
	ATCGGTAAAA TGTTAAAAGA TACGATTATG AGTAGTACTA AAAATAAATT AGCAGGTGCA	3180
	ATATTGAAGA AAGATTTAGC TGAATTCGCT AAAAAGATGG ATTACTCAGA ATACGGTGGT	3240
25	TCCGTATTAT TAGGATTGGA AGGTACTGTA GTTAAAGCAC ACGGTAGTTC AAATGCTAAA	3300
	GCTTTTTTATT CTGCAATTAG ACAAGCGAAA ATCGCAGGAG AACAAAATAT TGTACAAACA	3360
30	ATGAAAGAGA CTGTAGGTGA AtCAAATGaG TaAAACAGCA ATTATTTTTC CGGGACAAGG	3420
	TGCCCCAAAA GTTGGTATGG CGCAAGATTT GTTAAACAAC AATGATCAAG CAACTGAAAT	3480
	TTTAACCTCA GCAGCGAACA CATTAGACTT TGATATTTTA GAGACAATGT TTAGTGATGA	3540
35	AGAAGGTAAA TTGGGTGAAA CTGAAACAC ACAACCAGCT TTaTTGaCGC aTAGTTCGGC	3600
	ATTATTAGCA GCGCTAAAAA ATTTGAATCC TGATTTTACT ATGGGGCATA GTTTAGGTGA	3660
	ATATTCAGT TTAGTTGCAG CTGACGTATT ATCATTTGAA GATGCAGTTA AAATTGTTAG	3720
40	AAAACGTGGT CAATTAATGG CGCAAGCATT TCCTACTGGT GTAGGAAGCA TGGCTGCAGT	3780
	ATTGGGATTA GATTTTGATA AAGTCGATGA AATTTGTAAG TCATTATCAT CTGATGACAA	3840
45	AATAATTGAA CCAGCAAACA TTAATTGCCC AGGTCAAATT GTTGTTTCAG GTCACAAAGC	3900
	TTTAATTGAT GAGCTAGTAG AAAAAGGTAA ATCATTAGGT GCAAAACGTG TCATGCCTTT	3960
	AGCAGTATCT GGACCATTCC ATTCATCGCT AATGAAAGTG ATTGAAGAAG ATTTTTCAG	4020
50	TTACATTAAT CAATTTGAAT GCGGTGATGC TAAGTTTCCT GTAGTTCAAA ATGTAAATGC	4080
	GCAAGGTGAA ACTGACAAAG AAGTAATTAA ATCTAATATG GTCAAGCAAT TATATTCACC	4140
55	AGTACAATTC ATTAACCTCA CAGAATGGCT AATAGACCAA GGTGTTGATC ATTTTATGTA	4200

	AACATCAATT CAAACTTTAG AAGATGTGAA AGGATGGAAT GAAAATGACT AAGAGTGCTT	4320
	TAGTAACAGG TGCATCAAGA GGAATTGGAC GTAGTATTGC GTTACAATTA GCAGAAGAAG	4380
5	GATATAATGT AGCAGTAAAC TATGCAGGCA GCAAAGAGAA AGCTGAAGcA GTAGTCGAAG	4440
	AAATCAAAGC TAAAGGTGTT GACAGTTTTG CGATTCAAGC AAATGTTGCC GATGCTGATG	4500
	AAGTTAAAGC AATGATTAAA GAAGTAGTTA GCCAATTTGG TTCTTTAGAT GTTTTAGTAA	4560
10	ATAATGCAGG TATTACTCGC GATAATTTAT TAATGCGTAT GAAAGAACAA GAGTGGGATG	4620
	ATGTTATTGA CACAACTTA AAAGGTGTAT TTAAGTGTAT CCAAAAAGCA ACACCACAAA	4680
15	TGTTAAGACA ACGTAGTGGT GCTATCATCA ATTTATCAAG TGTTGTTGGA GCAGTAGGTA	4740
	ATCCGGGACA AGCAAATCTAT GTTGCAACAA AAGCAGGTGT TATTGGTTTA ACTAAATCTG	4800
	CGGCGCGTGA ATTAGCATCT CGTGGTATCA CTGTAAATGC AGTTGCACCT GGTTTTATTG	4860
20	TTTCTGATAT GACAGATGCT TTAAGTGTAG AGCTTAAAGA ACAAATGTTG ACTCAAATTC	4920
	CGTTAGCACG TTTTGGTCAA GACACAGATA TTGCTAATAC AGTAGCGTTC TTAGCATCAG	4980
	ACAAAGCAAA ATATATTACA GGTCAAACAA TCCATGTAAA TGGTGAATG TACATGTAAT	5040
25	ATATTTGAGC TAAAGCTCAT TGACGCAGTG GTTGACTGGT CATCCAATGG AGAATTGTCT	5100
	GACCTAGTCA ACTTTGCGGG GGAAATTCTA AGCAACCTAG ATAAGGTTCC AGAATTTCTC	5160
30	CCTAAGAAAC ACTAATCAAT aAATTGwTAA GTGTTTCTAA AATTTCTACT TGTTTTTTAG	5220
	AATTTAAAT GGGAAAATAT AGTAGTCTAT GTATAGGCAT TTTTAAAGGA GGTGAATCGA	5280
	CGTGGAAAAT TTCGATAAAG TAAAAGATAT CATCGTTGAC CgTTTAGGTG TAGACGCTGA	5340
35	TAAAGTAACT GAAGATGCAT CTTTCAAAGA TGATTTAGGC GCTGACTCAC TTGATATCGC	5400
	TGAATTAGTA ATGGAATTAG AAGACGAGTT TGGTACTGAA ATTCCTGATG AAGAnGCTGA	5460
	AAAAATCAAC ACTGTTGGTG ATGCTGTAA ATTTATTAAC AGTCTTGAAA AATAATAAAT	5520
40	CTTACATCTG GGTGTCAGT ATTGTCGACT CAGTTTTTTT CTTTAATTAT CAATAGTTTT	5580
	AACGTAAAT TAAAGATGAT TCAAGAGCAA CACATAAAGG AGATAAAATA ATGTCTAAAC	5640
	AAAAGAAAAG TGAGATAGTT AATCGTTTTA GAAAGCGCTT TGATACTAAA ATGACAGAGT	5700
45	TAGGCTTTAC TTATCAAAAT ATTGATTAT ACCAACAAGC ATTTTCGCAT TCGAGTTTTA	5760
	TTAATGATTT TAATATGAAT CGTTTAGACC ATAATGAGCG TTTAGAGTTT TTGGGTGATG	5820
50	CGGTATTAGA ATTGACGGTT TCACGATATT TATTTGATAa ACATCCCAAC TTGCCAGAAG	5880
	GGAATTTAAC AAAATGCGT GCCaCTATTG TATGTGAGCC CtCACTkGTA ATATTTGCGA	5940
	ATAAAATTGG ATTGAACGAA ATGATTTTAC TTGGTAAAGG TGAAGAGAAA ACAGGGGGAC	6000
55		

ATCAAGGACT AGATATAGTT TGGAAATTTG CTGAGAAAGT CATTTTCCCA CATGTAGAAC 6120
 AAAATGAGTT ATTAGGCGTG GTAGATTTTA AAACACAATT CCAAGAATAT GTGCACCAGC 6180
 5 AAAATAAAGG TGATGTAACC TATAATTTAA TAAAAGAAGA GGGACCGGCA CATCATCGTC 6240
 TATTCACTTC A 6251

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4920 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ACCTACTGAA GTTGCTAATT TTTTGGAGCA ACTAAGCACT GAAATTGAAC GTCTTAAAGA 60
 AGATAAAAAA CAACTTGAAA AAGTAATCGA AGAGAGAGAT ACTAATATTA AGTCTTATCA 120
 AGACGTGgCA TCAATCTGTA AGTgATGCTT TGATACAAGC TCAAAAAGCT GGTGAAGAAA 180
 25 CTAAGCAAGC TGCAGAGAAA CAAGCTGAAG CGATTATAGC TAAGGCAGAA GCGCAAgcTA 240
 ATcAAATGGT TGGTGACGCG GTAGAAAAAG CACGCCGTTT AGCATTCCAG ACTGAAGATA 300
 TGAAACGTCA ATCAAAAGTA TTTAGATCGC GTTTCCTAT GTTAGTTGAA GCGCAATTAG 360
 ACTTATTAAA AAACGAAGAT TGGGATTACT TGTTGAATTA TGATTTAGAC GCTGAACAAG 420
 TGACGCTTGA AAATATTCAT CATTTCATG AAAATGATTT AAAGCCAGAT GAAGTTGCAG 480
 35 CAAATGCACA AAATAATGCA TCAAATACAC CAGACAATAA TCAACAATCC AATGATTCAG 540
 AAACAACATA GAAGTAAGAA TTAAATAAAG ACAGACGCGT AATATACATT TAACTTTTCA 600
 CAGCGAATTA GGTAATGGTG AGAGCCTAGT AAAAGCATGT ATGTTATATC ACTGGCTTTT 660
 40 TAATATTTAA ATAATGTAAT GAGAGAACTC TAAGTTGAGT TAATAAGGGT GGTACCGCGA 720
 GCAATCGTCC CTTTAAATTT AACTTAGAGT TTTTAAATTT TTTAAGGAGT GAAAAAATG 780
 GATTACAAAG AAACGTTATT AATGCCTAAA ACAGATTTC CAATGCGAGG TGGTTTACCA 840
 45 AACAAGGAAC CGCAAATTCA AGAAAAATGG GATGCAGAAG ATCAATACCA TAAAGCGTTA 900
 GAAAAAATA AAGGTAACGA AACATTCATT TTACATGATG GCCCACCATA CGCGAATGGT 960
 50 AACTTACATA TGGGACATGC CTTGAACAAA ATTTTAAAG ACTTTATTGT ACGTTATAAA 1020
 ACTATGCAAG GGTTCATATG ACCATACGTA CCAGGTTGGG ATACACATGG TTTACCAATT 1080
 GAACAAGCAT TAACGAAAAA AGGTGTTGAC CGAAAGAAAA TGTCAACAGC TGAATTCCGT 1140

	TTAGGTGTTT	GTGGTGACTT	TAATGATCCA	TATATTACAT	TAAAACCTGA	ATACGAAGCT	1260
	GCACAAATTC	GTATTTTTTG	AGAAATGGCA	GATAAAGGTT	TAATTTATAA	AGGTAAAAAG	1320
5	CCAGTTTATT	GGTCTCCTTC	AAGTGAGTCT	TCATTAGCAG	AAGCAGAAAT	TGAATATCAC	1380
	GATAAACGTT	CAGCATCAAT	TTACGTTGCA	TTTGACGTTA	AAGATGACAA	AGGTGTCGTT	1440
10	GATGCAGATG	CTAAATTTAT	TATCTGGACA	ACAACGCCAT	GGACAATTCC	ATCAAATGTT	1500
	GCGATTACCG	TTCATCCTGA	ATTAAAATAT	GGTCAATACA	ATGTAAATGG	cGAAAAATAT	1560
	ATTATTGCAG	AAGCCTTGTC	TGACGCTGTA	GCAGAAGCAC	TGGaTTGGGA	TAAAGCATCA	1620
15	ATCAAATTAG	AAAAAGAATA	CACAGGTAAA	GAATTAGAGT	ATGTTGTAGC	ACAACATCCA	1680
	TTCTTAGACA	GAGAATCGTT	AGTGATTAAT	GGTGATCATG	TTACTACAGA	TGCTGGTACA	1740
	GgTTGTGTAC	ATACAGCACC	AGGTCACGGG	GAAGATGACT	ATATTGTTGG	TCAAAAATAT	1800
20	GAATTGCCAG	TAATTAGTCC	AATCGATGAT	AAAGGTGTAT	TTACTGAAGA	AGGCGGCCAA	1860
	TTTGAAGGGA	TGTTCTATGA	TAAAGCTAAT	AAAGCCGTTA	CTGATTTATT	AACAGAAAAA	1920
	GGTGCACTAT	TAAAATTAGA	CTTTATTACA	CATAGCTATC	CACACGACTG	GAGAACAAAA	1980
25	AAACCTGTAA	TCTTCCGTGC	TACACCACAA	TGGTTTGCCT	CAATCAGTAA	AGTAAGACAA	2040
	GATATTTTAG	ATGCAATCGA	AAATACAAAC	TTCAAAGTAA	ATTGGGGTAA	AACACGTATT	2100
30	TACAATATGG	TTCGTGACCG	TGGCGAATGG	GTTATTTCTC	GTCAACGTGT	GTGGGGTGTA	2160
	CCGTTACCAG	TATTTTATGC	TGAAAATGGC	GAAATTATCA	TGACGAAAGA	AACAGTGAAT	2220
	CATGTTGCTG	ATTTATTTGC	AGAACACGGT	TCAAATATTT	GGTTTGAAAG	AGAAGCGAAA	2280
35	GACTTACTAC	CAGAAGGATT	TACACATCCA	GGCAGCCCTA	ACGGTACATT	TACTAAAGAA	2340
	ACAGACATTA	TGGACGTTTG	GTTTGATTCT	GGTTCATCAC	ACCGTGGCGT	GTTGGAAACA	2400
	AGAGCGGAAT	TAAGTTTCCC	AGCGGATATG	TATTTAGAAG	GTAGTGACCA	ATATCGTGGT	2460
40	TGGTTCAACT	CTTCTATCAC	AACTTCAGTT	GCTACAAGAG	GAGTATCACC	TTATAAATTC	2520
	TTACTTTCTC	ATGGTTTTGT	TATGGACGGT	GAAGGTAAGA	AAATGAGTAA	ATCTTTAGGT	2580
45	AATGTGATTG	TACCTGACCA	AGTGGTTAAA	CAAAAAGGTG	CTGATATTGC	GAGACTTTGG	2640
	GTAAGTAGTA	CGGACTATTT	AGCTGATGTT	AGAATTTCTG	ATGAAATTTT	AAAACAAACA	2700
	TCTGATGTTT	ATCGTAAAAT	CAGAAATACA	TTAAGATTTA	TGTTAGGTAA	CATTAACGAT	2760
50	TTCAATCCTG	ACACAGATAG	CATTCTCTGAA	TCAGAGTTAT	TAGAAGTGGA	TCGTTACTTG	2820
	CTAAATCGTT	TACGTGAATT	TACTGCAAGT	ACGATTAACA	ACTATGAAAA	CTTGACTAC	2880
	TTAAATATTT	ATCAAGAAGT	TCAAACTTTT	ATCAATGTTG	AGTTAAGTAA	TTTCTATTTG	2940
55							

	CAAACAGTGT TATATCAAAT TTTAGTTGAT ATGACGAAGT TGTTAGCACC AATCTTAGTG	3060
5	CATACAGCTG AAGAAGTTTG GTCTCATACA CCACATGTTA AAGAAGAAAG TGTTCACTTA	3120
	GCAGACATGC CTAAAGTTGT AGAAGTAGAT CAAGCTTTAT TGGATAAATG GCGTACATT	3180
	ATGAATTTAC GTGATGATGT GAACCGTGCA TTAGAACTG CTCGTAATGA AAAAGTTATT	3240
10	GGTAAATCAT TAGAAGCTAA AGTTACGATT GCTAGTAACG ATAAATTTAA TGCATCTGAA	3300
	TTCTTAACTT CATTTGATGC ATTACATCAA TTATTTATCG TGTCAACAAGT TAAAGTTGTA	3360
	GATAAGTTAG ACGATCAGGC AACAGCTTAT GAACATGGTG ATATTGTCAT CGAACATGCA	3420
15	GATGGTGAAA AATGTGAAAG ATGTTGGAAC TATTCAGAGG ATCTTGGTGC TGTTGATGAA	3480
	TTGACGCATC TATGTCCACG ATGCCAACAA GTTGTAATAA CACTTGTATA ATTGAAATTG	3540
	TATAAGTAC TCATACAGAT GATATAAATT AAAGCTCTCT TCATAATCAT GTTGTAGTTT	3600
20	TTGTTGACAT GATGAAGAGA GTTTTTTTGT GAATAAAAAA ATGACCAAGT TACCGGTCAT	3660
	ATATGTAAAA AATGTGCGAT TTAATAAAT AAAAATTATT CAGGAATGGT ACAAATTCTC	3720
25	TGAGGCATAT AAATGCGTTA TAGTTGCTAT TCTCAATTAT GTTCGCGATA ATTTTAAGTA	3780
	AAAGTAAGCA CAGATATTGA ATTTGATAGG AGTTAATTGA ATGTATCATA ACAGTAACGC	3840
	AAACTTTGTC AATGGTATCA CTTTAAATGT GAGAGATAAG AATGAATTAA AGCCATTTTA	3900
30	TGAGGACATA TTAGGATTAA ATATTATAAA TGAGACATTA ACATCGATAC AATATGAAGT	3960
	AGGTCAAAAT AATCATGTCA TTACACTTGT TGAATTACAA AATGGACGTG AACCTTTAAT	4020
	GTCCGAAGCG GGAAGTTTC ATATCGCAAT TAACTACCT CAAATTAGTG ATTTAGCTAA	4080
35	TTTACTAATT CATTTAAGCG AATATGATAT TCCAGTTAAC GGAGGTATAC AGCCTGCTTC	4140
	GTTATCATT TTTTTTGAAG ACCCGGAAGG AAACGGTTTT AAATTTTATG TTGATAAAGA	4200
40	CGAAGCGCAA TGGACGAGGC AAAATAATTT AGTAAAAATT GATATTAGAC CATTAAATGT	4260
	ACCGAGATTA GTGAGTCATG CAACAAAATT GTTATGGTTA GGTATTCCAG ATGACGCTAT	4320
	TATAGGTGCA TTGCATATTA AGACAATTCA TTTATCAGAG GTAAAAGAGT ACTACCTCGA	4380
45	TTATTTTGGA TTAGAGCAAT CGGCATATAT GGATGATTAT TCAATATTTT TAGCATCGAA	4440
	TGGCTATTAT CAACATTTGG CCATGAATGA TTGGGTATCA GCAACGAAAC GTGTAGAAAA	4500
	TTTTGATACG TATGGATTAG CAATTGTTGA CTTTCATTAT CCTGAAACAA CACATTTAAA	4560
50	TTTACAAGGT CCGGATGGTA TCTATTATCG CTTTAATCAT ATCGAAGTTG AAGATTAGTA	4620
	TATACTTTGA ATGGACGAAC CATATAATGA ATCGTTTTTA ATGATCTTTT TATACAAGTT	4680
55	ATGAAGGAGG CTGGGACATT AAGTTCTTAG GCAATGTAAA AAGCTGATTT CTATTAATTA	4740

TTTTCCTTAT ATTAATTGCC ATTAATACAA AACCTAGCTC TCGTTTAACT TTATTTATTC 4860
 CTCGAACTGA CATTGnGTG AACTCAAAAT ngCCTACTTh CTAAATTAC CAATATCTAT 4920

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 626 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

TGGATTGCCA TTACATGGAC AAGATTTAAC TGAATCAATT ACACCATATG AAGGTGGTAT 60
 CGCTTTTGCA AGTAAACCAT TAATTGATGC TGATTTTATT GGTAAATCTG TATTAAAAGA 120
 TCAAAAAGAA AATGGTGACAC CAAGAAGAAC AGTGGGATTA GAATTACTTG AAAAAGGAAT 180
 TGCAAGAACT GGTATGAAG TTATGGATTT AGATGGAAAT ATTATTGGAG AAGTAACTTC 240
 AGGAACACAG TCTCCATCAT CAGGAAAATC AATTGCACTT GCAATGATAA AAAGAGATGA 300
 GTTTGAAATG GGTAGAGAGT TGCTTGTTCA AGTTCGTAAG CGTCAATTAA AAGCGAAAAT 360
 TGTTAAGAAA AATCAAATTG ATAAATAATT AAAAAGGGGT GTGCATTGTG AGTCATCGTT 420
 ATATACCTTT AACTGAAAAA GACAAGCAAG AAATGTTACA AACAATTGGT GCAAAATCTA 480
 TAGGAGAATT ATTCGGTGAT GTACCAAGTG ACATTTTATT AAATAGAGAT TTAAATATTG 540
 CTGAAGGCGA ACGGAGAACA ACGTTACTTA GAAGATTnAA TCGCATTGCA AGCAAGAGTA 600
 TCACTAGAGG AACGCGTACA TCGTTT 626

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1126 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ngGAAGTGGT GTATATATTT GTAATGAGTG TATTGAATTA TGCTCAGAAA TCGTCGAAGA 60
 AGAATTAGCT CAAAACACTT CTGAAGCGAT GACAGAATTA CCTACTCCTA AAGAAATTAT 120
 GGATCATTTA AACGAATATG TTATTGGTCA AGAAAAAGCT AAAAAATCTT TAGCTGTAGC 180
 TGTTTATAAC CACTATAAGC GTATTCAACA ATTAGGACCA AAAGAAGATG ATGTTGAATT 240

AACCTTAGCC AAGACGTTGA ATGTACCATT TGCAATTGCA GATGCGACAA GTTTAACTGA 360
 AGCTGGTTAT GTAGGCGATG ATGTTGAAAA TATCTTGTTG AGATTAAATC AAGCAGCTGA 420
 5 CTTTGACATT GATAAAGCCG AAAAAGGTAT TATTTATGTA GATGAAATTG ATAAAATTGC 480
 ACGTAAATCT GAAAACACAT CTATAACACG TGACGTTTCA GGTGAAGGTG TTCAACAAGC 540
 10 ATTGCTTAAA ATCTTAGAAG GTACGACTGC AAGTGTTCCG CCACAAGGTG GACGCAAACA 600
 TCCAAACCAA GAAATGATTC AAATTGATAC AACAAATATC TTATTTATTC TTGGTGGTGC 660
 CTTTGATGGT ATTGAAGAAG TGATTAAGCG CCGTCTTGGT GAAAAAGTTA TTGGTTTCTC 720
 15 AAGCAATGAA GCTGATAAAT ATGACGAACA AGCATTATTA GCACAAATTC GCCCAGAAGA 780
 TTTGCAAGCC TATGGTTTGA TTCCTGAATT TATCGGACGT GTGCCAATTG TAGCTAATTT 840
 AGAAACATTA GATGTAAGTG CGTTGAAAAA CATCTTAACG CAACCTAAAA ATGCACTTGT 900
 20 GAAACAATAT ACTAAAATGC TGGAATTAGA TGATGTGGAT TTAGAGTTCA CTGAAGAAGC 960
 TTTATCAGCA ATTAGTGAAA AAGCAATTGA AAGAAAAACA GGTGCGCGTG GTTTACGTTT 1020
 AATCATAGAA GAATCGTTAA TCGATATTAT GTTTGATGTG CCTTCTAACG AAAATGTAAC 1080
 25 GAAGGTAGTT ATTACAGCAC AAACmATTAA TGrAGaACTG AACCAG 1126

(2) INFORMATION FOR SEQ ID NO: 29:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4392 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

40 ATTGACTTCT TAGCAAThAA TaTGAGTGAA GAACGTACTG TTGAAGTACC AGTTCAATTA 60
 GTTGGTGAAG CAGTAGGCGC TAAAGAAGGC GCGTAGTTG AACCAACCATT ATTCAACTTA 120
 GAAGTAACTG CTA CTCCAGA CAATATTCCA GAAGCAATCG AAGTAGACAT TACTGAATTA 180
 45 AACATTAACG ACAGCTTAAC TGTTGCTGAT GTTAAAGTAA CTGGCGACTT CAAAATCGAA 240
 AACGATTCAG CTGAATCAGT AGTAACAGTA GTTGCTCCAA CTGAAGAACC AACTGAAGAA 300
 GAAATCGAAG CTATGGAAGG CGAACAACAA ACTGAAGAAC CAGAAAGTTGT TGGCGAAAGC 360
 50 AAAGAAGACG AAGAAAAAAC TGAAGAGTAA TTTTAATCTG TTACATTAAA GTTTTTATAC 420
 TTTGTTTAAC AAGCACTGTG CTTATTTTAA TATAAGCATG GTGCTTTTTG TGTTATTATA 480
 AAGCTTAATT AAAC TTTATT ACTTTGTACT AAAGTTTAAT TAATTTTAGT GAGTAAAAGA 540

55

	CTTACTAAGC TAAAGAATAA TGATAATTGA TGGCAATGGC GGAAAATGGA TGTTGTCATT	660
	ATAATAATAA ATGAAACAAT TATGTTGGAG GTAAACACGC ATGAAATGTA TTGTAGGTCT	720
5	AGGTAATATA GGTAAACGTT TTGAACCTAC AAGACATAAT ATCGGCTTTG AAGTCGTTGA	780
	TTATATTTTA GAGAAAAATA ATTTTTCATT AGATAAACAA AAGTTTAAAG GTGCATATAC	840
10	AATTGAACGA ATGAACGGCG ATAAAGTGTT ATTTATCGAA CCAATGACAA TGATGAATTT	900
	GTCAGGTGAA GCA GTTGAC CGATTATGGA TTATTACAAT GTTAATCCAG AAGATTTAAT	960
	TGTCTTATAT GATGATTTAG ATTTAGAACA AGGACAAGTT CGCTTAAGAC AAAAAGGAAG	1020
15	TGCGGGCGGT CACAATGGTA TGAAATCAAT TATTAAATG CTTGGTACAG ACCAATTTAA	1080
	ACGTATTCGT ATTGGTGTGG GAAGACCAAC GAATGGTATG ACGGTACCTG ATTATGTTTT	1140
	ACAACGCTTT TCAAATGATG AAATGGTAAC GATGGAAAAA GTTATCGAAC ACGCAGCAGC	1200
20	CGCAATTGAA AAGTTTGTTG AAACATCAGC ATTTGACCAT GTTATGAATG AATTTAATGG	1260
	TGAAGTGAAG TAATGACAAT ATTGACAACG CTTATAAAG AAGATAATCA TTTTCAAGAC	1320
25	CTTAATCAGG TATTTGGACA AGCAAACACA CTAGTAACTG GTCTTTCCCC GTCAGCTAAA	1380
	GTGACGATGA TTGCTGAAAA ATATGCACAA AGTAATCAAC AGTTATTATT AATTACCAAT	1440
	AATTTATACC AAGCAGATAA ATTAGAAACA GATTTACTTC AATTTATAGA TGCTGAAGAA	1500
30	TTGTATAAGT ATCCTGTGCA AGATATTATG ACCGAAGAGT TTTCAACACA AAGCCCTCAA	1560
	CTGATGAGTG AACGTATTAG AACTTTAACT GCGTTAGCTC AAGGTAAGAA AGGGTTATTT	1620
	ATCGTTCCTT TAAATGGTTT GAAAAAGTGG TTAACCTCTG TTGAAATGTG GCAAAATCAC	1680
35	CAAATGACAT TCGGTGTTGG TGAGGATATC GATGTGGACC AATTTCTTAA CAAATTAGTT	1740
	AATATGGGGT ACAAACGGGA ATCCGTGGTA TCGCATATTG GTGAATTCTC ATTGCGAGGA	1800
	GGTATTATCG ATATCTTTCC GCTAATTGGG GAACCAATCA GAATTGAGCT ATTTGATACC	1860
40	GAAATTGATT CTATTCGGGA TTTTGATGTT GAAACGCAGC GTTCCAAAGA TAATGTTGAA	1920
	GAAGTCGATA TCACAACCTGC AAGTGATTAT ATCATTACTG AAGAAGTGAT CAGCCATCTT	1980
45	AAAGAAGAGT TAAAACTGC ATATGAAAAT ACAAGACCCA AAATAGATAA ATCAGTGCGC	2040
	AATGATTTGA AAGAAACGTA TGAAAGCTTT AAATTATTCTG AAAGTACATA CTTTGATCAT	2100
	CAATACTAC GTCGCTTAGT AGCGTTTATG TATGAAACAC CTTGACAAAT TATTGAGTAT	2160
50	TTCCAAAAAG ATGCAATCAT TGCAGTTGAT GAATTTAATC GTATTAAAGA AACTGAAGAA	2220
	AGTTTAACAG TAGAGTCTGA TTCGTTTATT AGCAATATTA TTGAAAGTGG TAATGGATTT	2280
55	ATAGACAAA GTTTTATAAA ATATGATGAT TTTGAAACAT TGATTGAAGG CTATCCTGTC	2340

	TCATGTAAAC	CTGTCCAACA	ATTTTATGGG	CAATATGACA	TTATGCGTTC	TGAATTTCAA	2460
	CGATATGTTA	ATCAAACTA	TCATATCGTG	GTTTTGGTCG	AAACCGAAAC	TAAAGTTGAA	2520
5	CGTATGCAAG	CGATGTTAAG	TGAAAtGCAT	ATTCCATCAA	TAACAAAATT	GCATCGCTCA	2580
	ATGTCATCGG	GGCAAGCAGT	GATTATTGAA	GGCAGTTTAT	CTGAAGGATT	TGAACTACCT	2640
10	GATATGGGAT	TAGTTGTCAT	TACTGAGCGT	GAgcTTTTTA	AATCAAAACA	GAAAAAGCAA	2700
	CGAAAACGTA	CGAAAGCTAT	CTCAAATGCT	GAAAAAATTA	AGTCTTACCA	AGATTTAAAT	2760
	GTGGGAGATT	ATATTGTTCA	TGTGCATCAT	GGTGTGGTA	GATATTTAGG	TGTTGAGACG	2820
15	CTCGAAGTGG	GGCAAACGCA	TCGTGATTAT	ATTAAATTGC	AATATAAAGG	TACGGATCAA	2880
	CTATTTGTTT	CAGTAGATCA	AATGGATCAA	GTTCAAAAAT	ATGTAGCTTC	GGAAGATAAG	2940
	ACGCCAAAAT	TAAATAAACT	CGGTGGCAGT	GAATGGAAAA	AAACAAAAGC	TAAAGTTCAA	3000
20	CAAAGTGTTG	AAGATATTGC	TGAAGAGTTG	ATTGATTTAT	ATAAAGAAAG	AGAAATGGCA	3060
	GAAGGTTATC	AATATGGGGA	AGACACAGCT	GAGCAAACAA	CATTTGAATT	AGATTTTCCA	3120
	TATGAACTTA	CGCCTGACCA	AGCTAAATCT	ATCGATGAAA	TTAAAGATGA	CATGCAAAAA	3180
25	TCGCGTCCAA	TGGATCGCTT	GCTATGTGGT	GATGTTGGTT	ATGGTAAAC	TGAAGTTGCA	3240
	GTGAGAGCAG	CATTCAAAGC	TGTAATGGAA	GGAAAGCAGG	TTGCATTTTT	AGTTCCTACA	3300
30	ACTATTTTAG	CTCAGCAACA	TTATGAGACG	TTAATTGAGC	GTATGCAAGA	TTTTCTGT	3360
	GAAATTCAAT	TAATGAGTCG	TTTLAGAACG	CCTAAAGAGA	TAAAACAAAC	TAAGGAAGGA	3420
	CTTAAAACTG	GATTTGTTGA	CATAGTTGTT	GGTACACACA	AATTACTTAG	TAAAGATATA	3480
35	CAGTATAAAG	ATTTAGGGCT	GTTGATTGTA	GATGAAGAAC	AACGATTGGG	TGTACGCCAT	3540
	AAAGAGCGTA	TTAAACATT	AAACATAAT	GATAGATGTAC	TAACATTGAC	TGCAACCCCA	3600
	ATAGCTAGAA	CATTGCATAT	GAGTATGCTA	GGTGTGCGGG	ATTTGTCAGT	GATTGAAACG	3660
40	CCGCCAGAAA	ATCGTTTCCC	AGTTCAAACA	TATGTATTAG	AACAGAACAT	GAGTTTTATC	3720
	AAAGAAGCTT	TAGAAAGAGA	ACTATCCCGT	GATGGCCAAG	TGTTTTATCT	TTATAATAAA	3780
45	GTGCAATCCA	TTTATGAAAA	ACGAGAACAA	CTCCAGATGT	TAATGCCAGA	TGCTAACATT	3840
	GCAGTTGCTC	ATGGACAAAT	GACAGAGCGC	GATTTAGAAG	AAACGATGTT	AAGTTTTATC	3900
	AATAATgAAT	ATGATATTTT	AGTAACGACG	ACGATTATTG	AAACAGGTGT	CGATGTCCCA	3960
50	AATGCAAATA	CTTTGATCAT	TGAAGATGCA	GATCGCTTTG	GATTGAGTCA	GTTGTATCAA	4020
	TTAAGAGGTC	GTGTTGGTCG	TTCAAGTCGT	ATTGGTTATG	CATACTTCTT	ACATCCAGCA	4080
55	AATAAGGTAC	TAACGTGAGAC	TGCAGAAGAT	CGATTACAAG	CGATTAAAGA	ATTTACGGAG	4140

TTAGGTAAAC AACAGCACGG CTTTATTGAT ACAGTTGGAT TTGATTTGTA CAGTCAAATG 4260
 TTAGAAGAAG CTGTAAATGA AAAACGTGGT ATTAAGGAAC CAGAATCTGA GGTGCCAGAA 4320
 5 GTCGAAGTTG ATTTAACTT GGATGCATAT TTGCCAACAG AATATATTGC AAATGAACAA 4380
 GCTAAAATTG AA 4392

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 729 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

20 TTTCTTTTGA ATCTATATCG AGGTGGTTGG TAGGTTTCATC TAAAATAAGT ACATTGTCAC 60
 GTTGCAACAT AAGTAGTGCT AGTTGTAAAC GTGCTTTTTT ACCACCAGAT AAATCATTAA 120
 25 TTATCTTTTT AACATCGTCT TGTACAAATA AGAAACGTCC AAGAACTGCT CGAATATCTT 180
 TTTCAATTCAT TAACGGATAT TGATCCCACA CATAATCTAA AATCGTTTTA CTAGATTTAA 240
 ATTCTGCTTG CTTTTGATCA TAATAACCAA TTTGTAAATT TCGCGCGAAA GTAATATCGC 300
 30 CATTAGCGC TTTTTGTTGA TTAGCAATAG TTTTAATTAA GGTGATTTT CCAATACCAT 360
 TTGGCCCAAT GATTGCTATA TGATCGCCTT TAGAGACCTC TATACTCATA GGTTTGGTAA 420
 TTGCAGTTTG ATAACCGATT TCTAAATTTT TTACATGCAT GACGTCATTA CCTGTATTCC 480
 35 GGTCAAAGCC AAATTGAATA TTTGCACTTT TGGCATCTAA CATTGGTTTA TCAATGCGTT 540
 CCATTTTTTC TAAAATCTTA CGTCTACTTT TTGCCATTCC ACTTGTGAA GCACGGGTAA 600
 TATTTTTCTC AACAAAAGTT TCTAATCGTT TTATTTCTGC TTGTTGACTT TCATATTCTT 660
 40 GCATTCGTTT TTGATAATAT AAATCCCGTT GCTGTATAAA TTCCTCGTAA TTACCAACAT 720
 AGCGTTTGA 729

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13856 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

	TGATGTTTCG ATACATTTGT TGCACCTTGT GGATATACTT TAAAGGTTGT GTCGTATGTT	120
	TCCTTACTAT CTTTAGCTTC AGATTCCTGT GATTCAACCG TTTTATATTT TTCAAGTGCA	180
5	TGTCCTTCAA TATCAACTCG TGAATAATG CGATTCAACC ATGCTGGTAA ATACCACGAA	240
	CCTTTtCCAA ACAATTTcGt TAATGCAGGA ATTAACATCA TtCTGACTAC GAAGGCATCA	300
10	AAGAGTACAC CAAACGCTAA TGCCATACCC ATTGATTAA TCATGACATC TTCTTGGAAT	360
	ACAAACGCAA AGAAGACACT AAACATAATT AATGCAGCTG CTACAATAAC AGGACCGCTT	420
	TCTTTCAATC CTACTTTGAT AGAATAATCA TTATCCCCTG TTTTACTATm yyCTTCATGr	480
15	ATTGCGGACA TAAGGAAGAC TTCATAATCC ATCGCTAATC CAAATAAGAT ACCTATAGTA	540
	ATAACCGGTA AAAATGCTAG CATTGGTCCT GTCGTTTCAA TACCAAACAG ACCTTTCATA	600
	AAACCATCTT GCATTACTAA TGTGTAAAT CCTAATGTTG CCATTAATGA CAAGACGAAT	660
20	CCTAAAACTG CTTTAAATGG TATTAGAATT GAACGGAAGA CAATCATTAA TAAGAAAAAT	720
	GCTAATACAA CAATGACTGA GGCAAATAAA GGTATCGCCT CATTTAACTT TTTAGACATA	780
25	TCAATATTAA TGACACTTTG TCCCGAAATC TCCGTTTTGA ACCCATATTT ATCTTGTGCA	840
	TCTTTATGAT AATCTCGTAA ATCATGCACT AAATCATTTG TACTCTCTGC ATTAGGCCCT	900
	TGCTTAGGTA TCACGACCAT CAAAGCGTAA TCATTATCTT TACTCATTG TGGTGGCGTA	960
30	ACGATATCTA CATTTTTCTT ATCTTTAATA TCTTTATATA CAGACTGTAA ATCTTGTTGT	1020
	AATCCTTGTG GATCATCCTT TTTATCTTTC ACATTTATCA ACATCGGTAT TTGGCCATTA	1080
	AATCCTTCAC CAAATTTATC CGAGATAATA TCGTAAGCTT TTTTCTGTGT AGAATCTGCT	1140
35	GGTTTAACAC CGTCATCTGG AATACCAAGT CGCATATGAC TAACTGGTAT TGCAGCTGCT	1200
	ACTAATATGA TTAAACCTAG TAATACTGCC GCAAGTGCAT TTCCTGTAAT AAATTTAGAC	1260
	CATGGCGTAT CAATATCTTT TTTGAATTTA GACTGTAATT TATTCACCTT AATGCGTTtA	1320
40	TGGAAAATGC TTATTAATGC AGGTAATAAA GTTAAAGCGC TAAGTACTGC AAAACAACA	1380
	CTAATTGCCG AAGCAAATCC CATTACCGCT AAGAAGTCAA TGCCTACTAA TGATAAACCA	1440
45	CATACTGCAA TTACAACTGT TACACCAGCA AAAACAACCTG CACTACCTGC TGTTCCCTATT	1500
	GCAAGACCAA TGCCTTTAAT GTAATCTGTT TCAGTTTTCA TAACTTGTCG ATATCTGAAT	1560
	AAAATAAATA ATGCATAATC GATACCAACT GCTAGTCCAA TCATTACGGC TAATGTCAGT	1620
50	GTGACATTTG GTATATCGAA TGCATAAGTT AACAACTGA TAATACCTAC ACCAGAGGCT	1680
	AGACCAATCA ATGCACTTAT AATTGGTAAT CCTGCAGCAA TGA CTGAACC GAATGTGATT	1740
55	AACAGTACAA CAAATGCAAC AATAATACCA ACTAGTTCAG AATTACCGCC TACTTCTGTA	1800

	AAATGACTTT TAACATTATC TCTAGAGCCA TCTTTTAAAG ATGTTTGAAT AACGTCATAT	1920
	GTGATATCTG CAAATGCAGT TGTTTTATCT TFACTAATTT GCTTATTTTC ATAAGGATCT	1980
5	GATATTTTAT CAATGTGCTT GTCATCTTTT TTAATATCAT CTAACGTTTT CTTAATATCT	2040
	TTAGTAATGT TCGGTTGCAC AATACCATCA TCTTTAGTCG TCTTAAAGAC AACACGTATT	2100
10	TGTGCCTTTT CACTATCTTG ATTAAAATGT TTTTCAATCT TTTTATTCGT ATCTAACGAC	2160
	TCTAATCCTG TCATTTTAAAT ATCATTGTCA AATTTGCGTG CATTTGTAGC AAGTGGTATC	2220
	AATATTGCAG CTACAATCAC TATCCATGCA ATGACCGCGG ACCATTTATG TTTTGCGATG	2280
15	AATGTCCCCA TCTTATATAA AAATTTTGCC AAAGTATATT GCCTCCTTTT AAAATCAACG	2340
	TTATAGTTTA AATATACAGT GTAGATTATT GTTCGATTAT AGTATCTATC CCCGACCTCT	2400
	TAAAGAATCA ATTGGAAAAT TTTGTATATT AAACATACACA CAAAGGAGAA ATGTAGATGA	2460
20	AAGAGACTGA TTTACGAGTT ATAAAGACAA AAAAAGCATT GTCGAGTAGC TTGCTACAAT	2520
	TGTTAGAACA GCAATTATTC CAAACGATTA CTGTCAATCA AATTGCGAC AACGCACTCG	2580
25	TACACCGTAC AACATTTTAT AAACATTTT ATGATAAATA TGATCTTCTA GAGTACTTGT	2640
	TCAATCAATT GACTAAAGAC TACTTTGCTA GAGATATCAG TGACCGTCTT AATCATCCAT	2700
	TCCAAACGAT GAGTGATACG ATTAATAATA AAGAGGATTT GAGAGAAATC GCAGAATTCC	2760
30	AAGAAGAAGA CGCTGAATTT AATAAAGTAT TAAAAAATGT CTGCATTAAA ATTATGCATA	2820
	ACGATATCAA AAATAATAGA GACCGTATCG ATATTGACAG CGACATCCCA GATAATCTCA	2880
	TATTTTATAT TTATGACTCG TTGATTGAAG GTTTTATACA TTGGATAAAA GATGAAAAAA	2940
35	TTGATTGGCC TGGCGAAGAT ATTGATAACA TTTTCCATAG ATTAATCAAT ATTAAGATTA	3000
	AATAGTAGAT GAGAAACTCA TGAGCGTTAC CAACATTCAT AATAAAAACG ATAGTGKACA	3060
	CGTTAATGAA TTCGTGTACT ACTATCGTTT TTTATTTTTA TCGTGCTTAT CGCTATTAAA	3120
40	ACAACTGATA CACAACACAT AAACATGAA GAAAAAATA AATCCGCTAT CTAAATGACT	3180
	TTGACTCAGT TGTTTAAATG ACCAAATTGC TAATACAATT CCCATTATTA TTGAAATAAC	3240
45	GTATCTCACA TTCTTATACC TATAATCCTT TTCTAAAAAT ATGGTTGCTA TTAATTAAAT	3300
	TTTAAAGTTA TAAATAAAAA GAGCCAACCG CAATGGATGG CCCTTGTTCA TTATGAAGCA	3360
	TTAGAACATT TCTGAAACAA CCTTTTGTTT TAAGAAGTGT AATAAGTAGT CTGGACTACC	3420
50	TGTTTTAGCG TCCGTACCTG ACATTTTGAA ACCACCAAAT GGATGGTATC CAACAACTGC	3480
	TGAAGTACAG CCTCTGTTAA GGTATAAATT GCCTACATCA AATTCGTTTA CCGCTTTAAT	3540
55	CCAATGCTCG CGATTATTTG TAATCACTGC ACCAGTTAAA CCGTAATCTG TATCATTTGC	3600

	TTCTTCTTGC ATGATTCTAT CTTTAGATTT AAGTCCTGAA ATGATTGTTG GTTCTACAAA	3720
	GTAACCTTTT GAATCATCAG TGCCGCCACC TTGTTCTAAT TTACCTTCTT CTTTACCAAT	3780
5	CTCAATATAA TTTTAAATCT TATCAAATTG TTTTTTATTA ATAAGTGGGC CCATATACGT	3840
	ATTGTCTACA GTATTGCCCA ACGTTAATTC TTTTGTTAAT TTGATTGATT TCTCTAATAC	3900
10	TTCGTCATAA ACGTCTTTAT GCACAATTGC ACGTGAACAT GCTGAACATT TTTGACCAGA	3960
	AAAACCAAAT GCTGACGTTA CAATAGCTTC TGCTGCCATA TCTGTATCAA TATTTTCATC	4020
	AACTACAATG GCATCTTTAC CACCCATTTT AGCGATAACA CGTTTCAAGA AGTTTTGACC	4080
15	TTCTTGAACA ACGGCACTAC GTTCATAAAT TCTAGTACCT GTCGCACGTG ATCCTGTAAA	4140
	TGTAACGAAA TGCGTATCTT TATGATCAAC TAAGTAATCA CCAATTTCTT TCGGATCACC	4200
	AGGAACAAAG TTAAGTACGC CTTTGGTAA TCCTGCTTCT TCTAAAATTT CCATTAATTT	4260
20	ATAAGCGATA TAAGGTGTAT CCTCAGCAGG TTTCAATAAC ACTGTATTAC CTGCCACAAC	4320
	TGGTGCTAAA GTTGTACCAG CCATAATCGC AAACGGGAAG TTCCACGGCG GAATTGTAAC	4380
	ACCTGTACCA ATTGATTTAT AGAAATATTT ATTGTGTTCA CCTTCACGAT CAAGTACTGG	4440
25	CTTACCTTGA GCCAAGTCCA TCATTGAACG TGCATAGTAT TCAATAAAAT CAATACCTTC	4500
	AGCTGCATCA CCAACTGCTT CATCCCATGG CTTACCTGCT TCATAAACCA TAATTGCTGC	4560
30	AATTTCCGCT TTTCGACGAC GAATAATTGC CGAAACACGT AACATAAGCT CTGCACGATC	4620
	ATTTGCTGAC CATGTTTTCC AAGATTTATA AGCTTCGTTT GCTGCTTTAA ACGCATCTTC	4680
	AACATCTTGT TTTGTTGCCT TTGATGCATT TGCAATCACT TGTGATGTGT CTGCAGGATT	4740
35	GATTGATTTA ATTTTGTGAT CTTTGAAAAT CTTCTCTCCA TTAATCACTA ATGGTATGTC	4800
	TTGACCTAAT TCTTTTTCCA CGTCTTTCAA TGCTTTCTTA AACATATCCA CATTTTCTTG	4860
	GACTGAAAAA TCGTAACCAG GTTCATTTTT AAATCTACT ACCATGTACA CTTACCCCTT	4920
40	ATAAATTTTG AAAGTGGTTT AACCCTTTGA TTTAATGATA TAACATCATT TAAACTCATT	4980
	TTACTATGAT TAAGGTTAGT TTTGCAATCG CTTTCATTTT TATGTTTTAT CACTTATTCT	5040
45	CAAGTATTTT GAAATTGATT GGTACTTTT TAAATTTTAT ATGGGTCGCA ACTGCTACTT	5100
	TATCGTTTCG TCATTTAATG TTTCCGATGG TAGGTCATTA TCAATTTTAC GAACGACTTT	5160
	ACAAGGGTTT CCAACCGCTA AGCTGTGTGG CGGAATATCT TTAGTGACAA CACTACCAGC	5220
50	ACCAATCACA CTGCCTTCTC CAATCGTCAC CCCTGGTAAC ACGGCTACAT GACCGCCAAA	5280
	CCAAGTATTA CTGCCAATAT GAATGGGTCC GGCTTTTTCA AAACCTTCAT TTCTATGATG	5340
	GAAATTAAGT GGATGTGTCG CTGTGTAGAA TCCACAATTA GGTCCTATAA AAACATTATC	5400
55		

	TCCTAGTTTA	ACGTTCCAAC	CATAATCTGT	ATCAAAAGGA	ATCGAAATAC	TTACATTGTC	5520
	TGTTGTTGTT	TGAAATAATT	GATCAATTAA	TTCCTTTCTT	TTATTTGTAG	CACTCGGTCT	5580
5	TGTATGATTT	AATTCAAAGC	AAATATCTTT	CGCTCGTGCA	CGTTCATTGA	TTAAGTATTG	5640
	ATCAAAGTTT	GCATCGTACC	ATTTTCTGTC	TAACATTTTT	TCTTTTTCAG	TCATTACACC	5700
10	TTTCAACTCC	TAATAACTTA	TTTACTTGTT	TAAAAGTTAA	TCAAATAAAC	CTTCGCCTAT	5760
	GCAACTAATA	CGCTATAACA	TTATGAAATC	ATGACCTTAT	CACCCTTATC	TATACAATTC	5820
	TCGCATCAAA	TACTGCTAAA	GTAGTAGATA	AATTCAATAC	TACAGACGCA	TTCATTTTTT	5880
15	AATCTATTAA	CGTACAATGT	GAGTAAGAGA	AATATAAAGG	AGTATGATAG	CGATGAGAAT	5940
	ATTAATTACA	GGCACAGTTG	CTATCTTAAT	CATTCTAGGT	TTGGTCAAAA	CGATACAAGA	6000
	TTACGAAATG	ACAAACGACA	CGAGTCGTCA	GTTGTCAGAC	AACAAAGATG	ATGATAAAGT	6060
20	CATCCATCTT	AATAATTTTA	AAAATTTACA	TGCGAAAGAA	TTTAACCCAT	CTGATTTCTT	6120
	TTAAGTCACC	TAAGAATTGC	AAATCCAGAA	GTCATTTAAG	TTTTACCTTT	CATTCATACA	6180
25	TCCTTTAATA	TTAATTACGA	CTTCTTTTAT	ATAGATGCTA	AGTAGAGAGA	TTGTTGTGCA	6240
	ATGTTTGCAC	GGCAATCTCT	CTTTTCTTTT	TTAAAATTGG	TAAAAGTAAA	ACGCAACGAT	6300
	TGACTTATAT	ACCTATAGGG	GGTACATTAG	ACGTGTAACA	ATGAATCACA	GGGAGGCAAT	6360
30	AATGTGGCTA	ATACGAAAAA	AACAACATTA	GATATCACTG	GTATGACTTG	TGCCGCATGT	6420
	TCAAATCGTA	TCGAAAAGAA	ACTGAATAAA	CTTGATGACG	TTAATGCCCA	AGTGAATTTA	6480
	ACTACAGAGA	AAGCAACTGT	TGAGTATAAC	CCTGATCAAC	ATGATGTCCA	AGAATTTATT	6540
35	AATACGATTG	AACATTTAGG	TTACGGTGTC	GCTGTAGAAA	CTGTCGAATT	AGACATTACA	6600
	GGTATGACTT	GTGCTGCATG	CTCAAGCCGT	ATTGAAAAAG	TGTTAAATAA	AATGGACGGC	6660
	GTTCAAAATG	CAACGGTCAA	TTTAACAACA	GAGCAAGCTA	AAGTTGACTA	TTATCCTGAA	6720
40	GAAACAGATG	CTGATAAACT	TGTCACCTCG	ATTCAAAAAT	TAGGTTATGA	CGCGTCTATT	6780
	AAAGATAACA	ATAAAGATCA	AACGTCACGC	AAAGCTGAAG	CGCTACAACA	TAAATTGATT	6840
45	AAGCTTATCA	TATCAGCAGT	ATTATCTTTA	CCACTATTAA	TGTTAATGTT	TGTACATCTT	6900
	TTCAATATGC	ATATACCAGC	ACTATTTACG	AATCCATGGT	TCCAATTTAT	TTTAGCTACA	6960
	CCTGTACAAT	TTATTATTGG	ATGGCAATTT	TATGTAGGTG	CTTATAAAAA	CTTAAGAAAT	7020
50	GGTGGCGCCA	ATATGGATGT	ACTGTTGCT	GTTGGTACAA	GTGCAGCATA	TTTTTACAGT	7080
	ATTTATGAAA	TGGTTCGTTG	GCTAAATGGC	TCAACAACGC	AACCGCATTT	ATACTTTGAA	7140
55	ACAAGCGCCG	TACTAATTAC	CTTAATCTTA	TTCGGTAAGT	ATTTAGAAGC	TAGAGCGAAG	7200

TTAAGATG GTAATGAAGT GATGATTCTT CTAAATGAAG TACATGTTGG AGATACACTT 7320
 ATCGTTAAAC CAGGTGAAAA GATACCTGTT GATGGCAAAA TTATTAAAGG TATGACTGCC 7380
 5 ATCGACGAAT CTATGTTAAC AGGTGAATCT ATCCCTGTTG AGAAGAATGT TGATGATACT 7440
 GTAATTGGTT CAACGATGAA CAAAAACGGT ACTATTACTA TGACAGCAAC AAAAGTTGGC 7500
 10 GGGGACACTG CGTTGGCAAA TATTATTAAA GTTGTGCAAG AAGCTCAAAG TTCTAAAGCG 7560
 CCGATTCAAC GATTGGCAGA TATTATTTCT GGTTATTTTCG TTCCTATCGT TGTTGGTATC 7620
 GCACTATTAA CATTATCGT GTGGATTACT TTAGTTACAC CAGGTACATT TGAACCTGCA 7680
 15 CTTGTTGCGA GTATTTCCGT TCTCGTCATT GCTGTCCAT GCGCATTGGG ACTTGCTACA 7740
 CCAACTTCTA TTATGGTAGG TACTGGTCGC GCTGCTGAAA ATGGTATTTT ATTTAAAGGT 7800
 GCGGAGTTTG TTGAACGCAC ACATCAAATT GATACCATCG TTTTAGATAA GACGGGTACC 7860
 20 ATTACAAATG GTCGTCCAGT CGTGACAGAT TATCATGGTG ACAATCAAAC GCTACAATA 7920
 CTTGCTACTG CTGAAAAAGA TTCTGAACAC CCATTGGCAG AAGCCATTGT CAATTATGCA 7980
 AAAGAAAAGC AATTAATATT AACTGAGACA ACAACATTTA AAGCAGTACC TGGCCATGGT 8040
 25 ATTGAAGCAA CGATTGATCA TCACCATATA TTGGTTGGTA ACCGTAAATT AATGGCTGAC 8100
 AATGATATTA GCTTGCCTAA GCATATTTCT GATGATTTAA CACATTATGA ACGAGATGGT 8160
 30 AAAACTGCTA TGCTCATTGC TGTTAATTAT TCATTAAGTG GTATCATCGC AGTGGCAGAT 8220
 ACTGTCAAAG ATCATGCCAA AGATGCTATA AAACAATTGC ATGATATGGG CATTGAAGTT 8280
 GCCATGTTAA CTGGCGATAA TAAAAACACT GCTCAAGCCA TTGCAAAACA AGTAGGCATA 8340
 35 GATACTGTTA TTGCAGATAT TTTACCAGAA GAAAAAGCTG CACAAATTGC GAAACTACAG 8400
 CAACAAGGTA AGAAGGTTGC GATGGTTGGT GACGGTGTA ATGATGCACC TGCATTAGTT 8460
 AAAGCTGATA TCGGTATCGC CATTGGTACA GGTACAGAAG TTGCCATTGA AGCAGCTGAT 8520
 40 ATTACTATTC TTGGTGGCGA CTTGATGCTT ATTCCTAAAG CCATTTATGC AAGTAAAGCA 8580
 ACCATTCGTA ATATTCGTCA AAATCTATTT TGGGCATTGC GCTATAATAT TGCCGGTATC 8640
 CCTATAGCTG CATTGGGCTT ACTTGCGCCA TGGGTTGCTG GTGCTGCAAT GGCACTAAGT 8700
 45 TCAGTAAGTG TTGTCACAAA CGCACTTAGA TTGAAAAAGA TGCGATTAGA ACCACGCCGT 8760
 AAAGATGCCT AGATTCCTTA ATAATGAAGG ATTCGTTGGT GATTCTGAGA TAGGCTAGTG 8820
 50 ATTGGCTCTA TAATGTCGCG GTTTAyaGt GGATCTTCGC TCCAAGTCA TATATAGTnA 8880
 CACTTTTCGC TTGGCGAATT AGTGTATCTT ACCTAATAGc TCCGCCTATT AGGTTCCATC 8940
 ATTATTATAA ATAATAAGTA CACTACGGtT TACAGTTGGA TCTTCGCTCC AACTGCATAA 9000

55

GAAATTTTAA ATGTTGAAGG TATGAGCTGT GGTCACTGCA AAAGTGCTGT TGAATCTGCA 9120
 TTAAATAATA TTGACGGTGT CACTTCAGCT GACGTTAACC TTGAAAATGG TCAAGTAAGT 9180
 5 GTTCAATATG ATGACAGTAA AGTTGCTGTA TCTCAAATGA AAGACGCAAT TGAAGATCAA 9240
 GGTTACGATG TCGTTTAATT AGGCAATATT CAACGTCATC AACACCAAAT TAAAAAATCG 9300
 10 AACTGATGAG AATCCCAACA ATCCAAATTA TCTCATCAGT TCGATTTTAA ATTTACTCGT 9360
 AACCTAGTAT CTCCAGTCTG CAATACATCT AATGTTGCAT CTAATGCATC GACAATTAGA 9420
 TTTTAACTG CAGCTTCAGT ATAAAACGCA ATATGTGGTG TTAATATGAC ATCTTCCCTG 9480
 15 TCAATCAACG ATTCTAACAA TGGATCGTTC AGTGTTTTGC CCCTTTGATC ACTTGGGAAA 9540
 AGTTTGCGTT CAAATTCATA CGTATCAAGT GCTGCACCTT TAATCACACC ATTGTCTAAT 9600
 GCGTCTAATA ACGCCTTAGT ATCTACTAAA GAACCTCTCG CACAATTGAC AAATACTGCG 9660
 20 CCCTTTTTAA AATGTTTAA TAATTCAGCA TTAAATAGAT AATGATTATA TTTGTTGCA 9720
 GGTACATGTA ATGTCACGAT ATCAGCACCT TCAACCGCTT CCTCAATCGT ATCTTTGTAA 9780
 TCGACATACG TTGCAATTTT AGCATTAGGA AACGGTCGTA TCGGACCACA TCACTTTGAT 9840
 25 AACCATTGGC AAATATATCG GCTACTACAC GGCCAATTG ACCTGTACCA ATAACAGCTA 9900
 CTTTTAAATC TTTAATGGAT TTCGATAAAA TAGTAGGTTT CCATCTAAAA TCATGCTCCC 9960
 30 GCACTTTCGT TTGAATTTGA TTAAAATGAC GAACCACATT AATAGCCTGG TTCACAGCAA 10020
 ACTCCGCAAT TGAATTCGGA GAGTATGACG GCACATTTGA CACAATAAAG TTATACTTGT 10080
 TTGCTAACTC CAAATCATAT GTATCAAATC CAGCACTACG TTGTGCGATT TGTTTAATAC 10140
 35 CTAGTTCATT TAATCGTTTA TAAACATGCT CTGATAATGG TATTTGTTGT GATAGCGATA 10200
 AGCCATCATA ACCAGCGACA CCTTCAACAT TGTCATCAGT TAATGCTTCT TTAGTAATAT 10260
 CTACETCAAC ATGATGTTTC TCTGCCCACG CCTTGATATA AGGCATATCT TCATCACGTA 10320
 40 CACTCATGAT TTTAATTTTT GTCATTTTAA CATCACCTT AACTTTATTA TTCATATAAA 10380
 TATGCTAGTT CTGTTAATCT TATTGCAGCT TCGTCTAATT TCTGGTCATC TAACGCCAAT 10440
 GAAATTCTCA CATAACGATT ACCATTCTCT CCAAATGGTT TCCCTGGAGC AACAAGTATT 10500
 45 GACTTCTCTT GCACTAAAAA TTGCTCAAAT TGCTCGCTGT CATAACCAGG CGGTGTTTCC 10560
 AACCATACAT ATATGCCACC TTTAGCATGA ACAAATGGCA AATCAGCTTT TGCAAGCATG 10620
 50 GCTTCGAATC GGTACGACG TGTTTTAAAT ACATTGCTTT GTTCTTCTAA AAAATCATCA 10680
 TAATGATTCA AAGCATATAT TGCGGCATCT TGTAATGCAC CAAACATCCC AGCATTTGTG 10740
 TGCGTTTGGT ACTTTTTTCAA AGCTTGAATC ATATCTTTAT TACCAACTGC AAAACCGACT 10800

55

	CCATTTTCCG AAGCAAGTAT ACTAGGATTT TTAGCGTCGA AACCGAAAGC ACCATAAGCA	10920
	AAATCATGCA CGATTTTAGT GTCTGTACCT TTAAATTTAG CTATCGCTTC ATCAAAAAC	10980
5	TCTTTCGTAG CTGTCGATCC AGTTGGATTA TTTGGATACG TTAAATAAAT GAGTTTGT	11040
	TTATCTATTA TTTGTGAATC AACTTTGGAC CAATCTGGCA AATAATGTGG CGGTTCTAAA	11100
10	TTAAGCGGGA CTGGCTTGCC ATCAGCTAAA AGTACACCTG CTAAATAATC CGGTAGCCT	11160
	GGATCAGGTA GTAATACATA GTCTCCTGGA TTGATAACAC ATGTTGGTAC TGCCACTAAT	11220
	CCATTTTTTG TACCATATAA AATGCATACT TCATCTTCTT TATCTAACGT CACATTATAT	11280
15	TGTCTTTGAT AAAAATCTAC AATAGCTTGC TTGAACGCTT CTTTACCATG AAAAGCACCA	11340
	TATTTTTGAT TTTCAGGAAT AGTTAGTGCT TTTTGAATAT GATCAATAAT ACCTTGTGGC	11400
	GTGGGCCCCAT CAGGGATTCC AACTGCCATA TTAATTAATG GCAATGGTCC ATGTTGAT	11460
20	TTACGTCCCA TCGTTTTCCC GAAATAACTA TCAGGGATAT TTGCTAATTT GTTAGAGATC	11520
	ATCAAATTC TCCTCTATCA TTAAACATAG CCTGGGCGAC TATCATAATC CTAACAAC	11580
	GTATCACTCT CATTTAGATG GTTACAATGA CATCGCCATT CACCGTTATG TTCAACAGAA	11640
25	CTTATGACAC ACGTTGTATT GAATGAATTT ATTTTCATT TAGGTAGGTA TAATATTATT	11700
	GTCAATATTA GGAATTTTCA GATTAATATG CACTCAATCG TTATGATTTA ACTGTCATGC	11760
30	ATATCCGCAT GCGCAACCAG TTAGATATGC TTATATAAAG TATAACGCCC ATCAAGGTAC	11820
	GTATTCAAAC GTGAACCTTA ACAGGCGTCA TTCATTGTTA AATAAACTT CTTAAGCACA	11880
	TACTTATTTT ACTATGCCTT TTACGTTCCC CTTATACTTT TCTCACATCT TTCTCTTAGA	11940
35	CTACTCCCTT ATACGCCCCG CTCAATATCT TTAATCATT CATCTACAGT TATTTTCGCA	12000
	CTCGTTAAGA CAATAGGAAC GCCTGCACCT GGATGCGTAC TTGCACCTGC AAAATATAAA	12060
	TCTTTATAAT CTCGCGATAC ATTTTGTGGA CGATAATAAT TACTTTGCGC TAAAGTTGGC	12120
40	ATTAAACCGA ATGCCGAACC AAATTTTCGCA TGATACGTTT GCTCAAAATC ATTTGGCGTA	12180
	AAGATTGTTT CTGAAACAAT ATGCGATTTT ATATCTTCAA ATACTTCAAT CGTTGCTAAT	12240
45	TTACGATAAA TAATTTCTTT TATTTGTTGC GTCAAAGCTT CATCTGACCA ATCGATTCCG	12300
	CTACCTGTTT TAAGTTCCGG CGTCGGCATT AGCACATAAA TACCAGTTTT GCCTTCTGGC	12360
	GCAAGTGATT TATCAGCGAC CGCTGGTACA TACACATAAA TAGAAGGATC ATATGATAAA	12420
50	CGTCCCTCAA ATATTTCTTC AATATTGCCT CTAAAGTCAT CTGAAAAAAT AACATTATGA	12480
	AGTCTCACTT GATCTGTCAC ATCAATATCT ATACCGATAT ACATTAAAAA TGCTGAACAA	12540
55	GAGTAATCTA AGTCTGCAAT TTTATGTGGT GGATACTTTT TAATAGGTGC AAAATCTGGC	12600

ATGTCACCAT TCACTTTTAT CGCATCGGCC CGTTTGAATT TAGGATCAAT AATAATTTGC 12720
 TCAATTTTCAG CATTTAGTTC AATATTAACG CCTAAGTCTT TATTTAATTG CGCTAGcCCT 12780
 5 TGAGCCATGC CATACATACC GCCTTTAATA AAATGCACAC CAAACATCAT TTCAATCATA 12840
 GGAATAATTG AATATAGTGA CGGGCCTCGT TTTGGATCAA TTCCTATGTA TAACGTTTGA 12900
 AACGCTAAAA GCTTTTGTAT CTTTTCGTTA TCAATATAAT GTTCAATTAG CTGATCTGCA 12960
 10 TGATTAAACG TTTTAACTT AGCACCTTGC ACAAGTGACG TCATATTATA AAAGTCACTC 13020
 GGTTCGGAT ACGTTCTTTC TAAGAAATAG CGACGTGCAA TTTCATATTT TTTATAAACA 13080
 15 TCCGTTAAAA AGGACATAAA ACCATGCGTT GAACCAGGTT CTATACTTTC TAGCATTTCG 13140
 TGTAATTCAG CTAAATCTGT AGGCACCGTT ATACGATCAT CGTGGTCAAA ATACACATCG 13200
 TAAATATAAC GTAATTGTCT CAATTCAATA TAATCTTCAT AATTTTTACC ACACGCTGTA 13260
 20 AAAACATCTT TATAACATC TGGCATCATG ACAATTGTGG GACCCATATC AAATGTAAAG 13320
 CCGTCTTTCT TTAATTGATT CATACGCCCG CCTACATTAT TATTTTTTTC AAATATCGTC 13380
 ACTTCATGAC CTTGAGAAGC AATACGGGCT GCCGCTGCTA ATCCTGTGAC ACCTGCACCA 13440
 25 ATTACTGCAA TCTTCATTAT TCAACCACCT ATATTCTATG ATATTTACTA TTTATTTTCAT 13500
 GAAACAACTT TGCCTTTTTC CTCTTATCCA CAAAAACAG TTCATGTAAT GTATAGTTAG 13560
 CCTGTCTCAC TTCGTCCAGT ATTTCAATAT ATATACGTGC TGCTAATTCT ATGATTGGTT 13620
 30 GTGCTTCAAT ACTAAATACT TTGATTTGAT CCATAACATC TTGAAAATCT TTTTCTGCGA 13680
 TAGCTGCATA ATATTCCCAT AAGTCAATAT AATGATTATT AACACCATT TGGTACACTT 13740
 35 CAGCAATATC AACTTCATAT TGCTTTAATC GTTGCTTACT AAAATATATC CGTTCATTGT 13800
 CAAAATCTTC ACCGACATCT CTTAATATAT TAAnGGGATC CTCTAGAGTC GACCTG 13856

(2) INFORMATION FOR SEQ ID NO: 32:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10088 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

50 ATATATAAAT ATAGATTAAG TATATAGATT AATCAACTTT TTTGGAAGAG CAAATCACGC 60
 AATCAACAAA TAATATAAGA AGTTTTTGCG ATAGTTTTAA AATAGCTGTA ATAGAATACT 120
 AAATGTGACA AACTTAGAAC TAATATCAAG TGTTGATGTT TTGAATATAA AAATGCTAAT 180

55

EP 0 786 519 A2

	ATAATTGGTT AATATATGAG TAATTAGAAA ATAGACAAAG GATGACGATT TATGTATATC	300
	AATATGAAAG ATTATGGGTT AACAGGCATA AACAAAATA AAGATACTCG AGCAATACAA	360
5	CGTGCGTTAA ATCGTGGAAG ATGTAAACCA ACGACAGTTT ATATACCGAA AGGGACGTAT	420
	GATATTTGCA AACCATTAAAC GATATATGGC AATACAACAC TTTTGTTAGA TAATGAAACT	480
10	ATTTTACGCC GATGTCATTG TGGTCCTTTA TTAATAAATG GTCGTCGCTT TGGTTTTTaT	540
	CGTGTTTATA ATGGACACAG TCATATTCAT ATTAAAGGCG GCAAGTTTGA TATGAATGGT	600
	GTATCGTATC CTTATAACAA TACAGCTATG TGCATTGGGC ATGCTGAAGA TATTCAATTA	660
15	ATAGGTGTGA CCATTAAAGAA TGTAGTGAGT GGTCAATGCA TTGATGCTTG TGGGATTAAAC	720
	GGACTCTATA TTAAGAGCTG TTCATTTGAA GGATTCATAG ACTATAGTGG CGAACcTTTT	780
	ATTCTGAAGC AATACAATTA GACATTCAAG TACCTGGTGC TTTTCCAAAA TTCGGAACga	840
20	CAGATGGTAC GATAACGAAA AATGTCATTA TCGAAGATTG TTATTTTGGG CCTTCAGAAT	900
	TGCCCGAAAT GGGAAGTTGG AATCGTGCTA TTGGCTCACA TGCAAGTAGA CATAATCGAT	960
	ACTATGAGAA TATTCATATT AGAAATAATA TATTTGAAGA TATACAAGGT TATGCATTAA	1020
25	CTCCCTTGaA GTATAAGAT GCTTTTCAATTA TTAATAATAA GTTTATTAAAC TGTGaGGGTG	1080
	GCATTAGATA TTTAGGAGTT AGAGATGGTA AAAATGCAGC AGATGTGaTG ACAGGaAAAG	1140
30	ACTTAGGTTC CCAAGCAGGC ATAAATATGA ATATAATTGG AAATGAATTT AAAGGATCAA	1200
	TGTCTAAAGA TGCGATACAT GTACGTAATT ATAATAATGT TAAACATAAA GATGTATTAA	1260
	TCGTTGGGAA TACATTCAAT AATTCGACTC AATCAATTCA TTTAGAAGAT ATTGATACAG	1320
35	TGTTTTTAAG TCCTGTTGAA GCGGGTATTC AAGTTACTAC AATCAATGTA GATGAAATAA	1380
	AAAAGTAAAA AGTTTCGCAT GACATTAGGA TTAAGAATAG TAGATAATTT TTGAAAGCGC	1440
	ATTGATAAAA CGGTATAAAT ATGCTATAAT AAACCCAATT ATCTGATAAA AGGGGTATTT	1500
40	TGACGGTAAT GATAATACAA GATAGACAAC TTTCTATACT CTAATATAGT GAGTTGAAGT	1560
	AGCTTGTCAT AATCATCATG AGGGGGAAAT TTATGGCTTA TTTCAATCAA CATCAATCAA	1620
45	TGATATCGAA AAGGTATTTA ACATTCTTTT CAAAATCAAA GAAAAAGAAA CCGTTTAGTG	1680
	CGGGACAAC TATTGGACTA ATATTAGGTC CATTACTTTT CCTATTAAAC TTATTATTCT	1740
	TTCATCCACA AGACTTACCT TGGAAAGGCG TCTATGTTTT AGCGATTACT TTATGGATTG	1800
50	CGACTTGGTG GATTACTGAA GCAATTCCTA TTGCAGCAAC GAGCTTATTA CCAATTGTGT	1860
	TATTACCATT AGGTCATATA CTTACACCAG AACAAATATC ATCCGAATAT GGCAATGATA	1920
55	TTATCTTTTT GTTTTTAGGT GGATTTATTT TGGCAATTGC AATGGAAAGA TGGAATTTAC	1980

TTGGATTCA T GGTGGCAACA GGATTCTTAT CTATGTTTGT ATCGAACACT GCAGCTGTAA 2100
 TGATTATGAT TCCGATTGGT TTAGCAATTA TTAAGGAAGC ACATGATTTA CAAGAAGCCA 2160
 5 ATACGAATCA AACAGTATT CAAAAGTTTG AAAAATCTCT AGTTTTAGCA ATTGGCTATG 2220
 CAGGTACGAT TGGTGGCTTG GGTACATTAA TCGGAACCCC GCCATTAATT ATTTTAAAAG 2280
 GACAATACAT GCAACATTTT GGACATGAAA TTAGTTTTGC TAAATGGATG ATTGTAGGGA 2340
 10 TTCCAACGGT CATTGTTTTG TTAGGTATTA CTTGGCTCTA TTTAAGATAT GTTGCGTTTA 2400
 GACATGATTT GAAATATTTa CCTGGTGGTC AGACGTTAAT TAAACAAAAG TTAGACGAGC 2460
 15 TTGGCAAAT GAAGTATGAA GAAAAGGTAG TACAACTAT CTTGTACTT GCTAGCTTAT 2520
 TATGGATTAC AAGAGAGTTT CTTCTGAAAA AATGGGAAGT TACGTCATCT GTTGCGATG 2580
 GTACGATTGC TATTTTATA TCAATATTAT TATTTATTAT TCCAGCTAAA AATACTGAAA 2640
 20 AACATCGCCG TATCATTGAC TGGGAAGTTG CAAAAGAGCT CCCTTGGGGT GTATTAATTT 2700
 TATTTGGTGG CGGTTTAGCA TTAGCGAAAG GTATTTCTGA AAGTGGTTTA GCAAAATGGT 2760
 TAGGCGAACA GTTGAAATCA TTAAATGGTG TTAGTCCGAT TCTTATTGTA ATTGTCATAA 2820
 25 CAATCTTGT CTTATTTTA ACTGAAGTGA CATCTAATAC TGCAACTGCA ACGATGATTT 2880
 TACCGATTTT AGCAACGTTG TCTGTTGCTG TTGGAGTGCA TCCATTACTA CTTATGGCAC 2940
 CTGCAGCTAT GGCGGCTAAC TGTGCATACA TGTTACCAGT AGGGACACCA CCGAATGCAA 3000
 30 TTATCTTTGG TTCTGGTAAA ATATCTATCA AACAAATGGC ATCAGTAGGA TTCTGGGTAA 3060
 ACTTAATCAG TGCAATAATT ATTATTTTAG TCGTGTATTA TGTAATGCCT ATAGTTTTAG 3120
 35 GTATTGATAT AAATCAACCA CTGCCATTGA AATAGTAATT GCAGATTAGA ACGAAAAATA 3180
 AAAGGTTACA TTAGCAATTG CTTGGACGAG TGGTAACGAA ACGTATACCG CAGCATCGTG 3240
 TAAEAACAAT ACAACAAAA GAAAGTCAAC CAAGGATGGA TTCCTATTTT AATCCTTGGT 3300
 40 TGACTCTTTA TTTTATTTAA ATTGTAGAAC CTAGAAAATA AAGTTTAATT AAAAGCACCA 3360
 ATCATTTCTA CTTTGAAATC TAAGGTTTCT AAAATAGCAA TGACTTTCTT TATATCGGTT 3420
 GTAATTGCAG AATCAGCCTG AACGAAAAAT CGATACATAC CTAATTGTGT TTTTAAAGGA 3480
 45 CGAGACTCAA TCCAGGATAA ATTAATATTA AACAAAGCAA ATGTATTAAG CACACTTGCT 3540
 AACAAACCAG GTTTATCATG CATTGGTGTA ATTAAAAACA TCAATGATGT CGCATTTTGA 3600
 50 TCAAATTGCT GCTGATTTTT TATAACTAAA AAACGTGTCA CGTTATGTGG ATAGTCTTCA 3660
 ATATGTGTAT CAATAGGTGT AAAACCATAA GctTCGCCAC TACCTAAAGG TGCAATTGCT 3720
 GCAACGCCAT TTTCAATTTT AGTCAAACTT TGAATTGTAC TGTCGACATA ATCATAGTCA 3780

55

	TTTTTAATAT CAGAAATGGA ATCTGTTCCA TTACCATATA ATGCAAAGTT AATATCTAAA	3900
	CGTATTTTAC CGTGTGCAAA GACATCTTGC TGTGCAAGTG CATCTGCCAC AATGTTGATT	3960
5	GTTCCTTCTA TAGAATTTTC AATAGGGACA ACACCAATCG ATGTGTCATC ATCTGCAACT	4020
	GCCTTGATGA CTTCAAATAA ATTTGACTTT GGTGAAAAG TTGCTTCATT TTCAGAAAAA	4080
10	TACTGACGAC AAGCCAAATA TGAAAATGTA CCTTTAGGGC CTAAATAATA TAATTGCATA	4140
	TGCTACACCT CTAATACTT AATGATGGAA AGGGCACTGG TTAGCATTTG ATTCTTTCTT	4200
	TTTATAGAAA AAGTTTGGAT CTTTTACTGT ATTGTCATAT CCGTGATGAT AATTGACGT	4260
15	CAATGTTGGA GATAATGGCG GTGCTAGCCA AGACCATTTT CCGGTAACCT GACGACCTTG	4320
	TTGTGCTTCG TTACGTTTGA ATAGTTCGAA TTGCTTTGCA GCGGTCAAAT GATCGACAAT	4380
	TGATACGCCCT TCTTTTTTAA AGGAATGATA CACAGCATAG TTCAATTCAA CAAGTGCTCG	4440
20	ATCTTTATTA AATGAATTAT TTTTAAGTGT ATCAAATTCA AACGCATCTG CAACTTTTTTC	4500
	TAGTAAATTG TAACGGTAAT CATCAATAAA GTTACGTACG CCAATTTTCAG TTACCATATA	4560
	CCAACCGTTA AAGGGTGCAG TTGGATATAC AATGCCACCG ATTTTTAAGT CCATATTGGA	4620
25	AATGATAGGG ACTGCATACC ATTTTAAGTT CAATTTTCTT AATTTTGGAT AATGATTATG	4680
	TTCAATAGGT ACTTCTTTAA TTAATGAAGT AGGATATTCG TAAAATTTAA CTGACTCATT	4740
30	AGGTAATTGG TAAATCAGTG GTAACACGTC AAAATTAGTA CCTTTTCCTT TCCAACCTAA	4800
	GTGATTTGCT AAGCGTGTA CTTCTTTTTT AGCAGGATCA CCACAATTGT CATAGCCAGC	4860
	ATAGCGAATT AATTGATTGT TGAAAATTTT AGGTCCATCC TTTGGAGCAT ATATAGTAAT	4920
35	ATACGGCTTT AATTACCTT CATTGTAGC CTGTGTAATA TGATAAGTAA TTGATGATAA	4980
	GAACGATGCT TCGTCAGTAA CATCTCTTGC ATCAATGACA TTAAACGAAT CCCAAAATAA	5040
	ACGACCAATG CAACGATTTG AATTACGCCA AGCCATTTTA GCACCATAAA TAAGTTCTTC	5100
40	TTCTGTATGT GTATATGTCC CAGTTTCTTT TATTTCTAGT TCAATGTCAT GTAAACGTTT	5160
	ATTGATAATT TGCCTTTCAT AATGACACTC TTTATACATG TTTTCTATGA AAGCTTGAGC	5220
	CTCTTTAAAT AACATTAACA ACACCTCGCT TTATATTATA GTCTACATTA TTAAAATACT	5280
45	CTTAAAAATT ATGTATATGT CATTAAATTG TTGGTTGATT TTAATTAAAA GTATGGAAAT	5340
	TAAGGGGCTC TTATGTATAT AAAAAATGA ATTATGATAA AATGTAAGAA AATATTTAGG	5400
50	TCGATTGGAG AGATACAAGT GTACCAATTA GAAGACGACA GTTTAATGTT ACATAATGAC	5460
	TTATATCAAA TAAATATGGC TGAAAGTTAT TGGAATGATA ATATTCATGA AAAAATGGCT	5520
55	GTATTTGATT TGTATTTTAG AAAAAATGCA TTTAATAGTG GCTATGCTGT TTTAATGGT	5580

	TTAAAGTCTA TTGGCTACAA GGATGATTTC TTATCATATT TAAAAGATTT AAAATTCACA	5700
	GGCAGCATCC GTTCGATGCA AGAAGGCGAA TTATGCTTTG GTAACGAACC ATTGTTACGC	5760
5	GTAGAAGCAC CATTGATTCA AGCGCAATTA ATAGAAACAA TTTTATTAAA CATTGTAAAT	5820
	TTCCATACAT TAATTACAAC AAAGGCTAGC AGAATTCGTC AAATTGCATC AAATGATAAA	5880
10	TTAATGGAGT TTGGTACACG TCGTGCACAA GAAATTGATG CAGCATTGTG GGGCGCTAGA	5940
	GCTGCTTACA TCGGGGGCTT TGATTCTACA AGTAATGTTA GGGCGGGGAA ATTATTTGGT	6000
	ATACCTGTGT CTGGTACACA TGCACATGCA TTTGTCCAAA CTTATGGAGA CGAATATGTT	6060
15	GCCTTCAAAA AATATGCTGA AAGACATAAA AATTGTGTGT TCCTAGTAGA TACATTCCAT	6120
	ACTTTAAAT CTGGCGTGCC AAATGCAATA AAAGTTGCAA AAGAATTAGG TGACAAAATT	6180
	AACTTTGTAG GTATTGATT AGATTCTGGA GATATCGCTT ATTTATCTAA AGAGGCAAGA	6240
20	CGTATGCTTG ATGAAGCAGG ATTTACTGAA ACTAAAATTA TCGCGTCTAA TGATTTGGAT	6300
	GAAGAAACGA TTACGAGTTT GAAAGCACAA GGTGCAAAAG TAGATTCTTG GGGCGTTGGT	6360
25	ACAAAGCTGA TTACAGGATA CGATCAACCA GCATTAGGTG CAGTATATAA ACTTGTAGCT	6420
	ATTGAAAATG AAGATGGTTC ATATAGTGAT CGTATTAAAT TATCAAATAA CGCTGAAAAG	6480
	GTTACGACGC CAGGTAAGAA AAATGTATAT CGCATTATAA ACAAGAAAAC AGGTAAGGCA	6540
30	GAAGGCGATT ATATTACTTT GGAAAATGAA AATCCATACG ATGAACAACC TTTAAAATTA	6600
	TTCCATCCAG TGCATACTTA TAAAAAGAAA TTTATAAAAT CTTTCGAAGC CATTGATTTG	6660
	CATCATAATA TTTATGAAAA TGGTAAATTA GTATATCAAA TGCCAACAGA AGATGAATCA	6720
35	CGTGAATATT TAGCACTAGG ATTACAATCT ATTTGGGATG AAAATAAGCG TTTCTGAAT	6780
	CCACAAGAAT ATCCAGTCGA TTTAAGCAAG GCATGTTGGG ATAATAAACA TAAACGTATT	6840
	TTTGAAGTTG CGGAACACGT TAAGGAGATG GAAGAAGATA ATGAGTAAAT TACAAGACGT	6900
40	TATTGTACAA GAAATGAAAG TGAAAAAGCG TATCGATAGT GCTGAAGAAA TTATGGAATT	6960
	AAAGCAATTT ATAAAAAATT ATGTACAATC ACATTCAATT ATAAAACTTT TAGTGTTAGG	7020
45	TATTTTCAGGA GGACAGGATT CTACATTAGT TGGAAACTA GTACAAATGT CTGTTAACGA	7080
	ATTACGTGAA GAAGGCATTG ATTGTACGTT TATTGCAGTT AAATTACCTT ATGGAGTTCA	7140
	AAAAGATGCT GATGAAGTTG AGCAAGCTTT GCGATTCAAT GAACCAGATG AAATAGTAAC	7200
50	AGTCAATATT AAGCCTGCAG TTGATCAAAG TGTGCAATCA TTAAGAAG CCGGTATTGT	7260
	TCTTACAGAT TTCCAAAAAG GAAATGAAAA AGCGCGTGAA CGTATGAAAG TACAATTTTC	7320
55	AATTGCTTCA AACCGACAAG GTATTGTAGT AGGAACAGAT CATTGAGCTG AAAATATAAC	7380

	TAAACGACAA GGTCGTCAAT TATTAGCGTA TCTTGGTGCG CCAAAGGAAT TATATGAAAA	7500
	AACGCCAACT GCTGATTTAG AAGATGATAA ACCACAGCTT CCAGATGAAG ATGCATTAGG	7560
5	TGTAACCTTAT GAGGCGATTG ATAATTATTT AGAAGGTAAG CCAGTTACGC CAGAAGAACA	7620
	AAAAGTAATT GAAAATCATT ATATACGAAA TGCACACAAA CGTGAACCTG CATATACAAG	7680
10	ATACACGTGG CCAAAATCCT AATTTAATTT TTTCTTCTAA CGTGTGACTT AAATTAAATA	7740
	TGAGTTAGAA TTAATAACAT TAAACCACAT TCAGCTAGAC TACTTCAGTG TATAAATTGA	7800
	AAGTGTATGA ACTAAAGTAA GTATGTTTAT TTGAGAATAA ATTTTTATTT ATGACAAATT	7860
15	CGCTATTTAT TTATGAGAGT TTTCGTACTA TATTATATTA ATATGCATTG ATTAAGGTTA	7920
	GGTTGAAGCA GTTTGGTATT TAAAGTGTA TTGAAGAGA GTGGGGCGCC TTATGTCATT	7980
	CGTAACAGAA AATCCATGGT TAATGGTACT AACTATATTT ATCATTAAAG TTTGTTATGT	8040
20	AACGTTTTTA ACGATGCGAA CAATTTTAAC GTTGAAAGGT TATCGTTATA TTGCTGCATC	8100
	AGTTAGTTTT TTAGAAGTAT TAGTTTATAT CGTTGGTTTA GGTTCGGTTA TGTCTAATTT	8160
	AGACCATATT CAAAATATTA TTGCCTACGC ATTTGGTTTT TCAATAGGTA TCATTGTTGG	8220
25	TATGAAAATA GAAGAAAAAC TGGCATTAGG TTATACAGTT GTAAATGTAA CTTCAGCAGA	8280
	ATATGAGTTA GATTTACCGA ATGAACCTCG AAATTTAGGA TATGGCGTTA CGCACTATGC	8340
30	TGCGTTTGGT AGAGATGGTA GTCGTATGGT GATGCAAATT TTAACACCAA GAAAATATGA	8400
	ACGTAAATTG ATGGATACGA TAAAAAATTT AGATCCGAAA GCATTTATCA TTGCGTATGA	8460
	ACCTCGAAAC ATACATGGTG GATTCTGGAC TAAAGGCATT CGTCGTAGAA AGCTTAAAGA	8520
35	TTATGAACCA GAAGAACTGG AAaGTGTAGT AGAaCATGAA aTTCmAAGTA AaTGAGAAaTG	8580
	AaMCAATtGC TGATTGTTTG TCACGAATGA AAtGCAAGGG TATATGCCGG TAAAACGTAT	8640
	TGAAaAaCCC GTGTTTCAAG AGCAaAaAGA TGGCACGGTT GAAGTATCAC ATCAAGAAAT	8700
40	CGTTTTTGTA GGTAAGAAAA TCCAATAACA TAATCCAATT TAAATAAAGA CTATTTGAAG	8760
	AGGAAAGGCT ATTCAAAGTT TGAGTAATTT TACTTTGAAT AGCCTATTTG TTTATACATG	8820
	CAAGATGCTC GATCCATATT GTATGAGAAA CCCCAGCAA GCTATATAAA GCATATGCTG	8880
45	GGGGTTCTTA ATATTTTAAA AATTATTGTT AGATTATATA TATCGTCGCT TTTTCTAAAA	8940
	CAATCTCATC GCATGAAATT TTTTCTTCCT AGAGACCTTT AATAAGATTA ATAGTTTACT	9000
50	TAATCATATC TAGATAGTCT TATGACTTAT GCTTAATGAA AGTCATTCTA GGAGAAGTTC	9060
	CCAAAGCTTC TGTGTTTATA ATTGTTAGTA GTATTTTATT ATCATTGGT ATAAATATTT	9120
	CAATAACAAT TGAGCTATTA TTTTATTAT ATAATGTGAG TTGTTTGTGT TCTGTATTTA	9180

55

CATTTAAATC TTGAGGATGC CATTCTCCCT CAATAATATT AAGATAATAC TTAGCCTCTG 9300
 AATTACATTT GAATTTATCA ATACTAAATA ATTCAATTTG TTCCATAATA TTATTTACCT 9360
 5 TTCTAAAATA CAAATTTTAA TAACCATAAA TAGATGAATA CCATCGATAA TGGTCGCCAT 9420
 TGGATACTGG AATAACATTG TTTTTCAGCAT CTTGAGTCAT AAAACCATTG TCCCATGGAT 9480
 TCCATATAAT TATAACCTCT TGTCCATTAT CTAATTTAGC GTTCCCAACA ACTGCCATGG 9540
 10 CATGCCCTGC GTGCATACCA TTTCTTGATT CTAATCTACT ACCTAAAACA GCAATTCCTT 9600
 TATTATTTTT AGTAAGATTG TCAACTTCAT TATATGTAGT CATTCTATTA AGAAGTTGTG 9660
 15 GACTTCTTCC CTGAGTTTGT CCAAAATAAA TCATCTCTCT TGGCGTTAAA CCAGTAAATT 9720
 GGAATCGTTG TCCTTGTAAG TTTGGGTGTA AAAATCTCAT CACAGCTTCT GCATGATATT 9780
 TGTAGTATT ATAAGTCGCA TTTAGTAATT CAGACATCGT ATAGCCTGCA CACCAACCAT 9840
 20 TGTTACCTTG AGTTTCTCTT ATCTTGAAAT TCTCAAGTTT ATTTATATAT TGSTCGTTGT 9900
 AAGTATAATT ATTACTTTTA AATTGACTAG TTGGCATAGT GACAGAAGCT TTTTGCTTTA 9960
 GTTGCCTTAC ATTATTGCCA GTAGGTATAC TCTCAGTCTT TnTnAACTnT nTATCTTCTA 10020
 25 GACGTGGTGT TTTTAGTACT AGTTTAGCTT TATGATTTTG AGTACCACAT AGTAACCTTT 10080
 TGAGTTGT 10088

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7563 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

40 CGGAAACGnA CCCnATGCGT ATGCTTGACG TGCCAAAATT AAATACGAAG TTCATAGCTT 60
 TGAGGTACCA GAAGAACATT TATCTGGTCA AGAAGTCGCA GnACTCATAC AAGCAAATGT 120
 TAAAACAGTA TTTAAAACGC TTGTTCTAGA AAATACAAA CATGAACATT TTGTATTTGT 180
 45 TATCCCAGTA AGTGAACTT TAGATATGAA AAAGGCAGCT GCTTTGGTTG GAGAGAAGAA 240
 ATTGCAGCTT ATGCCTTTAG ATAATTTGAA AAATGTAACG GGATACATTC GTGGTGGGTG 300
 TTCGCCTGTT GGTATGAAAA CATTGTTTCC AACAGTCGTT GACAAATCGT GTGAAAATTA 360
 50 TAGTCATATC AGTGTGAGTG GTGGGCTTCG AACAAATGCAA ATCACAATAG CTGTTGAGGA 420
 TTTGATTACA ATAATAAAG GCAAAATTGG AGCAGTTATC CATGAATGAT TAATAACAAC 480

TGCCACACTC CTTTTTGATT GAATTAGCAT TTTACGATCA TAAACAGTCA TTATAATTGA 600
 GTATTTGAAC ATAAAAATGT AATTTTATCG TAACAATTTG AGTGTTTG TG ATTGTTTTTG 660
 5 GTAATTTATG ATTGAAAAGT GAAAGCGTAC TCATTATAAT ACAAAGTGAG ATGGGGTGAT 720
 GATGATAATT ACTGaAAAAA GACACGAGTT AATATTAGAA GAACTTTCGC ACAAAGATTT 780
 TTTGACTTTA CAAGAATTAA TAGATCGAAC TGGTTGCAGT GCTTCAACAA TACGArGAGA 840
 10 TTTATCTAAA CTACAACAAT TAGGGAAATT GCAACGTGTG CATGGTGGTG CAATGTTAAA 900
 AGAAAATCGT ATGGTTGAGG CGAATTTAAC TGAAAAATTA GCAACGAATC TTGATGAAAA 960
 15 GAAAATGATT GCTAAAATAG CAGCTAATCA AATCAACGAT AATGAATGCT TATTTATCGA 1020
 TGCTGGTTCA TCTACATTGG AGCTAATTAA ATATATTCAA GCGAAAGATA TCATTGTGGT 1080
 AACCAATGGT TTAACACATG TAGAAGCTTT ACTTAAAAAA GGTATTAAAA CAATTATGCT 1140
 20 AGGTGGTCAA GTTAAAGAAA ATACACTTGC TACGATTGGT TCTAGTGCTA TGGAGATATT 1200
 AAGACGATAT TGTTTCGATA AAGCTTTTAT CGGGATGAAT GGATTAGATA TTGAACCTTG 1260
 ATTAAC TACT CCCGATGAGC AAGAGGCATT AGTTAAACAA ACAGCAATGT CATTAGCCAA 1320
 25 TCAATCATT GTACTTATAG ATCATTCTAA GTTTAATAAA GTATATTTTG CTCGTGTACC 1380
 TTTGCTAGAA AGTACGACAA TCATCACATC TGAAAAGCA TTAAATCAAG AATCGTTAAA 1440
 AGAATACCAA CAAAAGTATC ACTTTATAGG AGGGACTTTA TGATTTATAC AGTGACTTTC 1500
 30 AATCCTTCAA TTGACTATGT CATTTTACG AATGATTTTA AAATTGATGG TTTGAACAGA 1560
 GCAACAGCAA CATATAAATT CGCTGGGGGG AAAGGTATTA ATGTCTCGCG CGTCTTAAAG 1620
 35 ACATTGGATG TTGAGTCAAC TGCCTTGGGA TTTGCAGGTG GATTTCTTGG GAAATTCATT 1680
 ATAGATACAT TAAATAACAG TGCAATTCAA TCGAATTTTA TTGAAGTTGA TGAAGATACA 1740
 CGTATTAAATG TGAAATTAAA AACAGGACAA GAAACAGAAA TCAATGCACC GGGTCCTCAT 1800
 40 ATAACGTCAA CACAATTGA ACAACTGTTA CAACAAATTA AAAATACAAC AAGCGAAGAT 1860
 ATAGTTATTG TTGCTGGAAG TGTACCAAGT AGTATTCCAA GCGATGCGTA TCGCAAAT 1920
 GCACAAATTA CAGCACAGAC AGGTGCTAAA TTAGTAGTCG ACGCTGAAAA AGAATTGGCT 1980
 45 GAAAGCGTTT TACCATATCA TCCACTATTT ATTAAACCTA ATAAAGATGA ATTAGAAGTG 2040
 ATGTTTAATA CAACAGTGAA CTCAGACACA GATGTTATTA AATATGGTCG TTTGTTAGTT 2100
 GATAAAGGTG CGCAATCTGT TATTGTCTCG CTTGGCGGTG ATGGTGCTAT TTATATTGAT 2160
 50 AAAGAAATCA GTATTAAAGC AGTTAATCCA CAAGGGAAAG TGGTTAATAC AGTTGGCTCT 2220
 GGTGATAGTA CAGTTGCAGG CATGGTGGCT GGAATTGCTT CAGGTTTAAAC GATTGAAAAA 2280

55

	CGGGACGCTA TAGAAAAAAT AAAATCACAA GTTACGATTA GCGTACTTGA TGGGGAGTGA	2400
	AAATAATGAG AGTAACAGAG TTATTAACAA AAGATACAAT AGCAATGGAT TTAATGGCAA	2460
5	ATGACAAAAA TGGTGTTATT GATGAGTTAG TAAATCAATT AGACAAAGCA GGTAAATTAA	2520
	GTGATGTCGC GTCATTTAAG GAAGCGATTG ACAATCGAGA ATCACAAAGT ACAACTGGTA	2580
	TCGGCGAAGG TATTGCCATT CCACATGCCA AAGTGGCCGC AGTTAAGTCA CCAGCTATTG	2640
10	CGTTTGGTAA ATCTAAAGCA GCGGTAGATT ATCAAAGTTT GGATATGCAA CCAGCACACT	2700
	TATTCTTTAT GATTGcAGcG CCAGAAGGTG GCGCCCAAAC ACATCTAGAT GCTTTAGCTA	2760
15	AGTTGTCTGG TATTTTAAATG GATGAAAATG TACGTGAGAA ATTATTACAT GCTTCATCAC	2820
	CTGAAGAAGT ACTAGCGATC ATAGATGAGG CTGATGATGA AGTGACAAAA GAAGAAGAGG	2880
	CAGAAGCTGA AGCACAACAA GTTGCAACTG CAGAACAATC ATCTAAACAA TCTAATGAGC	2940
20	CATATGTGTT AGCAGTAACT GCTTGTCCAA CAGGTATTGC ACACACATAT ATGGCACGTG	3000
	ATGCATTGAA AAAGCAAGCG GATAAAATGG GTATTAAAAT TAAAGTAGAA ACGAATGGTT	3060
	CAAGCGGCAT TAAAAACCAT TTAAGTGAAC AAGATATTGA AAATGCAACA GGTATCATTG	3120
25	TTGCTGCTGA TGTTTCATGTT GAGACGGATC GCTTCGATGG TAAAAATGTC GTAGAAGTAC	3180
	CAGTAGCAGA TGGTATTAAA CGCCCAGAAG AATTAATTAA TAAAGCATTG GATACAAGTC	3240
	GTAAACCTTT TGTTGCCCCG GATGGTCAAA GAAAAGGTAA CTCAAATGAC AGTCAAGAAA	3300
30	AATTAAGCCC AGGTAAAGCA TTCTATAAAC ACTTAATGAA CGGTGTTTCT AACATGTTGC	3360
	CACTTGTAAT ATCTGGTGGT ATTTTAATGG CAATTGTATT TTTATTTGGA GCAAATTCAT	3420
35	TTAATCCAAA AAGCTCAGAG TACAATGCGT TTGCAGAGCA GCTTTGGAAC ATTGGTAGTA	3480
	AAAGTGCATT CGCGTTAATC ATTCCAATTT TATCTGGATT CATTGCACGT AGTATTGCGG	3540
	ATAAACCTGG TTTGCTTCA GGTCTTGTAG GTGGTATGTT AGCAATTTCA GGTGGTTCAG	3600
40	GATTTATTGG TGGTATTATT GCAGGTTTCT TAGCAGGTTA CTTAACACAA GGTGTTAAAG	3660
	CCATGACACG TAAGTTACCA CAAGCATTAG AGGGATTAAA GCCAACATTA ATTTATCCAC	3720
	TATTAACAGT GACGGCTACA GGCTTATTGA TGATTTATGC CTTTAATCCA CCAGCATCTT	3780
45	GGTTAAATCA TTTGTTATTA GATGGATTAA ACAATTTATC AGGTTCTAAT ATTGTATTAT	3840
	TAGGTTTAGT TATTGGCGCT ATGATGGCGA TTGATATGGG CGGTCCATTC AACAAAGCGG	3900
	CATATGTTTT TGCAACAGGT GCGTTGATTG AAGGTAATGC AGCACCAATT ACAGCTGCAA	3960
50	TGATTGGTGG TATGATTCCA CCGTTAGCAA TTGCGACAGC GATGTTAATT TTTAGACGTA	4020
	AATTTACAAA AGAACACGT GGTTCATTA TCCCTAACTA TGTGATGGGT ATGTCATTTA	4080

55

5 TGATTGGTTC AGGTATAGGT GCGCAATTG CTTTAGGCTT AGGTTACGA ATTACTGCGC 4200
 CACATGGTGG TATTATTGTA ATTGTTGGTA CTGATGGTGC ACACTTACTT CAAACTCTTA 4260
 10 TTGCACTTCT AGTTGGCACA TTAGTTTCAG CATTAAATTTA CGGTTTAATC AAACCAAAGT 4320
 TAACTGAAAC AGAAATCGAA GCTTCAAAAT CAATGGACGA GTAGTTTAA TGATGTAAAA 4380
 TGATTGTTAG CAAAGAGCTT CATATTAAGT TGTATGTTCA ATGAATATAT GTTAGTTTAA 4440
 15 TATATCGTGT TAACGGTAGC TTATACAAAG CTGTAAAAAC ACTTCTATT AATTCAGTTT 4500
 TTATGAATTG ATATGAAAGT GTTTTTATTT TTAGATAAAT GAATGAAGAA ATAGACACCA 4560
 CAAATGTATA GACTTTTTTA ATATTTTGCA AAAAGTTATG CCAAACGAAG CAGATATAGT 4620
 AAAATATGAG TGTCTTAAAG TGAAAATTTA TAAATAAGA AGGGTTTATA CGTGTCAGAA 4680
 TTAATTATAT ATAACGGCAA AGTTTATACT GAAGATGGCA AAATCGATAA TGGTTACATT 4740
 20 CATGTGAAAG ATGGACAGAT TGTGCAATT GGAGAAGTGG ATGATAAAGC AGCAATTGAT 4800
 AATGATACGA CAAATAAAT TCAAGTGATT GATGCTAAAG GTCATCATGT ATTACCAGGT 4860
 TTTATTGATA TACATATTCA TGGTGGTTAT GGTCAAGATG CAATGGATGG GTCATACGAT 4920
 25 GGCTTAAAT ATCTATCCGA AAATTTGTTG TCTGAAGGGA CGACATCATA CTTGGCCACT 4980
 ACAATGACGC AATCGACTGA TAAATAGAT AATGCACTTA CAAATATTGC TAAATATGAA 5040
 GCGGAgCAAG ATGTTCAAA TGCAGCGGAA ATTGTAGGTA TACATTTAGA AGGACCATTT 5100
 30 ATATCTGAAA ATAAAGTTGG TGCTCAACAT CCGCAATACG TTGTACGCCC ATTTATCGAT 5160
 AAAATTAAC ATTTTCAAGA GACTGCTAAC GGATTAATAA AGATTATGAC GTTTGCACCT 5220
 GAAATTGAAG GTGCAAAAGA AGCGCTTGAA ACGTATAAAG ATGACATTAT TTTTCAATT 5280
 35 GGTACATACG TAGCAACATA CGAAGAAGCA GTTGAAGCTG TTGAGCGAGG AGCTAAACAT 5340
 GTCACGCATT TATATAATGC AGCGACGCCA TTCCAACATA GAGAACCAGG TGTTTTTGA 5400
 40 GCAGCATGGT TGAATGATGC TCTACATACC GAAATGATTG TTGATGGCAC TCATTCTCAT 5460
 CCGGCATCGG TTGCAATTGC TTACCGTATG AAAGGTAATG AACGTTTTTA TTTAATTACC 5520
 GATGCAATGC GTGCAAAAGG TATGCCTGAA GGAGAATATG ATTTGGGTGG ACAAAAAGTA 5580
 45 ACTGTTCAAT CGCAACAAGC ACGTCTTGCA AATGGTGCGC TTGCTGGTAG TATTTTAAAA 5640
 ATGAATCATG GGTTACGTAA CTTAATATCA TTTACAGGTG ATACATTAGA TCATTTATGG 5700
 50 CGAGTAACAA GTTTAAATCA AGCCATTGCA TTAGGTATCG ATGATAGAAA AGGTAGTATT 5760
 AAAGTAAATA AGGATGCAGA TCTTGTTATT CTAGATGATG ATATGAATGT AAAATCTACA 5820
 ATAAACAAG GCAAGGTTCA CACATTTAGC TAATAAATA TCATAATTAA ATGTATGCAA 5880

55

	TTTTCTGGGG GTGTCTAAAT GGGAAGGCCA TAACATGTAG TTGTAATTTA AGTCATAGTG	6000
	ATAAATTTGA ATGCGTGTTA CCCATGAGTG ACACATATAA CATGGAGGTG AATCCCTAGA	6060
5	AATAGGGAAT TAATTGGAAA CTTCGACCAT AATTAGTTTG ATTATATTTA TTCTATTAAT	6120
	TGCATTAACC ACTGTATTTG TTGGTTCAGA ATTTGCATTA GTAAAAATTA GAGCAACAAG	6180
	AATTGAACAG CTAGCAGATG AAGGAAATAA ACCTGCTAAA ATAGTAAAAA AGATGATTGC	6240
10	TAATCTAGAT TATTATCTTT CTGCTTGTC A GTTAGGTATA ACAGTAACAT CTTTAGGGTT	6300
	AGGTTGGCTT GGTGAACCAA CGTTTGAAAA GCTATTACAC CCAATATTTG AAGCAATCAA	6360
	TTTACCAACT GCATTACGA CGACGATTTT GTTTGCAGTG TCATTTATAA TCGTTACGTA	6420
15	TTTGCATGTA GTACTTGGTG AATTAGCGCC TAAATCTATA GCTATTCAAC ATACTGAAAA	6480
	GCTTGCTTTA GTATATGCAA GACCATTGTT CTATTTCGGT AACATTATGA AACCATTGAT	6540
20	TTGGCTGATG AATGGTTCTG CACGTGTTAT TATTAGAATG TTTGGTGTA ATCCTGATGC	6600
	CCAACTGAT GCAATGTCAG AAGAAGAAAT CAAAATTATT ATTAACAATA GTTATAATGG	6660
	TGGAGAAATC AACCAAATG AATTGGCATA TATGCAAAAT ATCTTTTCAT TCGATGAAAG	6720
25	ACATGCAAAA GATATAATGG TACCTAGAAC TCAAATGATT ACACTAAATG AACCTTTTAA	6780
	TGTAGACGAA TTAGTAGAAA CAATAAAGA ACATCAATTT ACGCGTTATC CAATTACTGA	6840
	TGATGGTGAT AAAGACCACA TTAAAGGATT TATTAACGTC AAAGAATTTT TAACTGAATA	6900
30	CGCTTCTGGA AAAACGATTA AATAGCAAAA CTATATACAT GAGTTGCCAA TGATTTCAGA	6960
	GACAACACGT ATCAGTGATG CATTAATTAG AATGCAACGT GAACATGTAC ATATGAGTCT	7020
	TATTATAGAT GAATATGGTG GAACGGCAGG TATTTTAACG ATGGAAGATA TTTTAGAAGA	7080
35	AATCGTTGGA GAAATTCGTG ATGAATTTGA TGATGATGAA GTGAATGATA TCGTTAAAAA	7140
	TGATGATAAG ACATTCCAAG TAAATGGCAG AGTACTATTG GATGATTTAA CTGAAGAGTT	7200
40	CGGTATAGAA TTTGATGACT CTGAGGATAT TGATACGATA GGTGGATGGT TACAATCTCG	7260
	TAATACCAAT TTACAAAAAG ATGATTACGT GGATACAACT TATGATCGCT GGGTTGTTTC	7320
	AGAAATCGAT AACCAACAAA TTATTGGGT GATATTAAAC TATGAATTTA ATGAAGCGAG	7380
45	ACCTACTATC GGACAGTCTG ATGAAGATGA AAAATCAGAA TAGATATTAA TATATAAACC	7440
	AACTAAGAAT GATTTAATTC ATTTTGGTT GGTATTTTTT TTGACTAAAA TTAAnGAAAA	7500
	GTGAAAAATAG TATTGGAAC CAATATCTTT AATGATTTAA TGAATAAnTT TTATTGAAAG	7560
50	CGA	7563

(2) INFORMATION FOR SEQ ID NO: 34:

55

(A) LENGTH: 3492 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

10	TTATATCAAC TTCATGGCGG AACCATTGAT GACCCATTAG ACGAAACAAT AAGCGCATT	60
	SATGAATTGA AACAGAAGG AATTATACGT GCTTACGGTA TTTCTTCTAT TCGCCCAAAT	120
	GTAATTGATT ATTATTTAAA ACATAGTCAA ATCGAAACGA TAATGTCTCA ATTCAATTTG	180
15	ATTGATAATC GTCCAGAATC ATTATTAGAT GCAATTCACA ACAATGATGT TAAAGTATTG	240
	GCAAGAGGAC CTGTGTCTAA AGGATTATTA ACTTCAAACA GTGTTAATGT GCTCGACAAT	300
	AAATTTAAAG ATGGTATTTT TGATTATTCT CATGATGAAT TGGGTGAAAC AATAGCCTCT	360
20	ATTAAAGAAA TTGAAAGTAA TTTATCTGCA TTGACATTTA GTTATTTAAC ATCACATGAC	420
	GTGCTTGGTT CCATCATTGT AGGTGCAAGT AGCGTCGACC AATTAAAAGA AAATATTGAA	480
	AACTATCATA CTAAAGTTAG TTTAGATCAG ATTTAAACAG CAAGAGCTCG TGTAAGGAT	540
25	TTGGAATATA CCAATCATTT AGTGTAGAAG TCATTTTCAG TAATAAAAAC AGCAGCATGA	600
	GGCGTTTCAT TATAAAAATG CCTTACTGCT GTTGTTTATG TACAATTCGC TATAATTTAT	660
30	GATTATGATT ACTCACTTAT GATAGAAATT AAAGCGTGTG CCTCACGCAT CAGTATTTAG	720
	TAATTTCGCC TTGCGGCATT GCCTTAAGCA AACTTCTGCC ACTTCATCTC TTAATAATTT	780
	TATTAAAACA TCTTTCTATA TTTCACTTCG CATGTTGATT CATCATTTAT AGTTATTATT	840
35	TGTACACCCA GCACATTTCC TTGCAACACA AGTAGTTTGA ATTTTTCACA AGTATAATAT	900
	AATGTACCGT CTGAAATTTG GTCTACAGAA ATATCGCCTA AAATATCCAG CACTGTAAAT	960
	TCTTCAAATA CTGATAGTTG TTCCGCATAT CGTACACAAA GTCTTACCAC ACTCTCCGAT	1020
40	TGACAGTTCA TTGCCATCCC ACCTATTTAT GCTTTATTTT TAAATAATTT AGGGAAACAT	1080
	CGTTCAAAAA ATCTAGCGCG AATTTGATAC ATTTTCAACG CATGaTGCAT CCATTTAGGC	1140
	CGATTAATTT CCAATTGTTT TGTTTTAATG CCATAAATGA TATCTTCTGC AAGCTGATTA	1200
45	GCATCAAGCA TAATTTCCCC CATCTTTTTA GCATACTTCA TTGATGGGTC GGCTTTTTGA	1260
	TGAAAAGGTG TATCAATCGG GCCAACATTA ACTGTCATGA TATGTAAGTT TGGTGACTCT	1320
50	AGTCTTAAAG CATTCATTAA TGCATAAAAC CCTGCTTTCG ATGCCCCATA ATGTGCAGCA	1380
	TTTGCTTG TGGAATGC AGCTTGACTT GAAATACCTA CAATATGTGC GTTAGATGTT	1440
	AAATATGGTC TCAACACAGT ATATAAAACA TTAAACTAA TTAAATTAAG CTGATACGTT	1500

55

	TAAATGAATC CATCGAATGA TGTATTGTCT TCAAATTGCA GTGCCTGTAT CGACTTCAAA	1620
	TCATTTAAGT CACAAGGAAT AACATTATA GTTTTCCCCA ATTCTGTTC AAAGATTCTA	1680
5	GTTGCTTTAT CAACATCACG CACCAACAAC GTTACATGCA CTTTATTTTC TAGTAACTTT	1740
	CGGACAATCG ATAAACCTAA ACCACTCGTA CCACCAGTCA CTATAAAATG TTGTCCTTTC	1800
10	ATCAATTAAC CTTCTTTTC AATTATATAG AATGCAATTT ATCAACTTTA CATAATTGAG	1860
	ACAAGTTGAT TATCTTTTCCT AATATATATA CAATAATAAG AAAATATAAC ATACAAATCA	1920
	AAAATAAAG GGATGTGACG TTAATGTAAC TCGTATTTTA TGGAGCTGGT AATATGGCAC	1980
15	AAGCTATATT TACAGGATT ATTAACTCMA GCAACTTAGA TGCCAATGAT ATATATTTAA	2040
	CAATATAATC TAATGAACAA GCTTTAAAG CATTGCTGA AAACTAGGT GTTAACTATA	2100
	GTTATGATGA TGCACATTA TTAAGATG CAGATATGT ATTTTAGGT ACCAAACCAC	2160
20	ATGACTTTGA TGCTCTAGCA ACACGCATCA AACCACATAT TACAAAAGWC AATTGCTTCA	2220
	TTTCAATTAT GGCAGGTATT CCGATTGATT ATATTAAACA ACAATTAGAA TGCCAAAATC	2280
	CaGTTGCTAG AATTATGCCA AACACAAATG CGCAAGTTGG ACACTCTGTT ACTGGCATT	2340
25	GTTTTTCAAA CAACTTTGAC CCTAAATCTA AAGATGAAAT TAACGATTTA GTTAAAGCAT	2400
	TTGGTTCTGT AATTGAAGTA TCAGAAGATC ATTTACATCA AGTAACAGCT ATCACCGGAA	2460
	GCGGCCAGC ATTTTATAT CATGTATTG AGCAATATGT TAAAGCTGGT aCsAAACTTG	2520
30	GTCTAGAAAA AGAACAAGTT GAAGAATCTA TACGCAACCT TATTATAGGT ACAAGTAAGA	2580
	TGATTGAACG TTCAGATTG AGCATGGCTC AATTAAGAAA AAATATTACC TCTAAAGGTG	2640
35	GTACGACACA AGCTGGCCTT GATACATTGT CACAATATGA TTTAGTATCT ATTTTCGAAG	2700
	ATTGTCTAAA CGCTGCCGTC GACCGTAGTA TTGAACCTTC TAATATAGAA GACCAATAAA	2760
	AACAACCCG CCAACACATG TATGCATCAT CGCAAGCACT GTGTTTGACG GGTATTTTTT	2820
40	ATAATTTATT GTTATTGGC AAGCATTGTT TATTACTTTG TCATTAGATT TTAAACTAT	2880
	CAAAATCTTT TACAAAATTA AAATTAGGTG TATCTTCATT TTGTATCAAT GTTTGATAAA	2940
	TTTCATTAT ATCTTCTGTA TTATAGCGAT TGCTCAAATG TGTAATCAAC GTACGTTTAA	3000
45	CATTGGCTTC TTTTATCAAT GCAAATACGT CTTCAATATG GCTATGATGA TAATTGTTGG	3060
	CTAAATGCTT TTCACCATCT ATATAGGTCG CTTCATGTAC CATCACATCA GCATCTCTAG	3120
	AAATCACACG TTCATTAGAA CATGGTTTTG TATCACCAA AATTGCTACA ACTGGACCCT	3180
50	GTTTGGACTC ACCTCTAAAA TCTTTGATT GATAAACTTG ACCATTATGT TCAAATGTAT	3240
	CATGAGATTT TACTTCTTGA TATTTAGGAC CTGGTTCAAG ACCAATGTTT TTTAACGCTT	3300
55		

CATGATTAAG TAAATGCGCC TCTACAGTAA AACCATCCAT GATGATATGT CAGATGATCA 3420
 TCGATTTCAA TATATGLAAT TGGATAGTTT AAATGTGACT CTGATAAATT CATAGACATT 3480
 5 TCCACATATG CT 3492

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1973 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

ATCTAGCGGT ACAAGCGTCT TGGAGGCTAG TATGTTGAAC ATTGTAAACC CTGAAGATCA 60
 20 CTTGCTGTGC ATTGTTTCAG GTGCCCTTGG TAACCGATTT AACCAAATTG CACAAACTTA 120
 TTACAAAAAT GTGCATATTT ATGACGTAAC ATGGGGAGAA GCTGTAGATG TCAAAGATTT 180
 CATCAATTTT CTTTCAACTT TAAATGTTGA AGTTAAAGCA GTATTTAGTC AATATTGCGA 240
 25 AACATCTACG ACAGTGCTAC ACCCTATTCA CGAGTTAGGA AATGCCATTA ATCAATTTAA 300
 TAGTAATATT TATTTTGTAG TTGACGGCGT AAGTtGCATT GGTGCTGTTG ATGTTGACAT 360
 TAACAAAGAT AAAATTGATG TACTTGTTC TGGTAGTCAA AAAGCAATTA TGTACCTCC 420
 30 AGGATTAGCT TTTGTAGCTT ATAGCCACCG TGCAAAAGAA CATTTCAAAG AAGTAACTAC 480
 GCCAAAATTT TATCTAGACT TAAATAAATA CATTTCTGTC CAAGCTGACA ATTCTACACC 540
 GTTCACACCA AATGTGTCTT TATTTAGAGG TGTAATGCA TACGTTGAAA CCGTAAAAGC 600
 35 AGAAGGTTTC AATCACGTAA TAGCAGACA CTATGCAATT AGAAATGCAT TAAGAAGCGC 660
 CTTAAAGCA TTAGATTTAA CTTTATTAGT CAATGATAAA GATGCATCTC CAACGGTTAC 720
 40 AGCATTCAA CTAATACAA ATGATGAAGT GAAATAATC mAAGATGAAC TTAAAAATnG 780
 CTTTAAAATA ACAATTGcng GTGGTCAAGG CCATCTTAAA GGTCAAATTT TnAGAATTGG 840
 TCATATGGGG AAAATTAGTC CTTTCGATAT TTTATCGGTA GTATCTGCTT TAGAAATTAT 900
 45 TTTAACTGAA CACCGTAAAG TTAATATAT CGGTAAAGGT ATATCAAAT ATATGGAGGT 960
 TATTCATGAA GCAATTTAAT GTACTCGTTG CAGATCCCAT ATCAAAGAT GGTATCAAAG 1020
 CATTATTAGA TCACGAACAA TTCAATGTAG ATATTCAAAC TGGCTTGTC GAAGAAGCAT 1080
 50 TAATCAAAT TATACCTTCA TACCATGCTT TAATCGTTTC TAGTCAAAC ACGGTTACTG 1140
 AAAATATCAT AAATGCTGCT GATTCTTTAA AAGTAATCGC ACGCGCCGGT GTTGGTGTAG 1200

5
 10
 15
 20

GTAATACGAT	TTCAGCTACT	GAACATACAC	TGGCAATGTT	ATTATCAATG	GCACGAAATA	1320
TTCCGCAAGC	ACACCAATCA	CTTACAAATA	AAGAATGGAA	TCGAAATGCA	TTTAAAGGTA	1380
CTGAGCTTTA	TCATAAAACA	TTAGGTGTCA	TTGGTGCTGG	TAGAATTGGT	TTAGGTGTTG	1440
CTAAACGTGC	GCAAAGTTTC	GGAATGAAAA	TACTAGCTTT	TGACCCTTAC	TTAACGGATG	1500
AAAAAGCAAA	ATCTTTAAGC	ATTACGAAGG	CAACAGTTGA	TGAGATTGCC	CAACATTCTG	1560
ATTTGTTTAC	ATTACATACA	CCACTAACAC	CTAAAACAAA	AGGCTTAATT	AATGCTGTCT	1620
TTTTTGCCAA	AGCAAAACCT	AGTTTGCAAA	TAATCAATGT	GGCACGTGGT	GGTATTATTG	1680
ATGAAAAGGC	GCTAATAAAA	GCATTAGACG	AAGGACAAAT	TAGTCGGGCA	GCTATCGATG	1740
TGTTTGAACA	TGAACCTGCA	ACTGACTCGC	CTCTTGTTGC	ACATGATAAA	ATTATTGTTA	1800
CACCTCATT	GGGTGCTTCA	ACAGTCGAAG	CTCAAGAAAA	AGTGGCAATT	TCTGTTTCAA	1860
ATGAAATCAT	CGAAATTTTA	ATTGATGGTA	CTGTAACGCA	TGCAGTGAAT	GCACCTAAAA	1920
TGGACTTAAG	CAATATAGAT	GATACTGTAA	AATCATTTCAT	CAATTTAAGC	CAA	1973

(2) INFORMATION FOR SEQ ID NO: 36:

25
 30

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7620 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

35
 40
 45
 50
 55

GGTGTTCAG	ATGTCACTGG	TTGATTTTAA	ATTGTAGACG	GGTATTTTGG	GCTTTCGCCA	60
TATTTATTTG	CCGGCTTACT	GTCAAAGCAT	AGGAATACTA	TCATAACAAT	TGTTAGGCCT	120
AAATGAACAA	AATAAGAAG	TACTAACAAA	ATATTAAGAC	CCATCGGCAT	TAATGTAAAA	180
TCCTGTTCAT	AATAACTATC	GATAATCTGT	AATACTATAT	AAAATATAAT	ACTGAATACT	240
GTCATAATCA	TTGGAAATAA	CATTGTTCTT	GATATATCGT	GAAATCTTCG	AACGCACAAC	300
GCTAAATTTG	GAATAAACGT	TGCCAAACTA	TAGACAAAAG	TATACACAGA	TGTAAGGATA	360
ATCATCAATA	TACTCATAAC	TATTAATGTT	TCGTTATCCG	CCGCTATAGA	AATAAAGAAT	420
AGAAATAGGT	TTATTATTAG	CACACACACA	GCTGGAACCA	TAAGTATCAA	ATGCCATAGT	480
GCCATATACC	AATATTCACT	ACGTCTTGAT	CTCCCCTTAA	AATTTACATA	ATTTTTCCAA	540
AATAAAACGA	ATGATTTTCAT	AAAACCTACT	TGAGGTAATT	GTTCCATTGT	AATCTCCCTT	600
TCGTTAATCA	TATTTATATT	TTTAATTATT	GTTACCGTTA	TAATTTACAA	GATTCATTAT	660

	GTAAAATGAA AACCCGCTAC AAGTACACAT CTATATGGAG ACTCATTGA AAGTCAACGC	780
	TTCGTAACT ATACTAAAAA TATGTCATAC TGCAATGTTT ACGTTTAAAA GAGTCTCAAT	840
5	CTATGCAAAT AAAATATTCC ATAACAAAGT ATATACTTTA CATTTTTATA ATTCTTAACA	900
	ATACTATTTT ATCAAACATT TACCACAATA AAAATATCTT TTTCAATTTT ATTTAAATTA	960
10	ATCATATAAT TCGGAGGAGA ATATTATGGA TTTCGTAAAT AATGATACAA GACAAATTGC	1020
	TAAAAACTTA TTAGGTGTCA AAGTGATTTA TCAGGATACC ACTCAAACGT ATACAGGCTA	1080
	CATCGTGGAA ACGGAAGCTT ACTTAGGTTT GAATGATCGT GCGGCTCATG GCTATGGCGG	1140
15	TAAAATAACA CCTAAAGTCA CGTCATTATA TAAACGTGGT GGTACAATTT ATGCACATGT	1200
	CATGCATACG CATTTACTCA TTAATTTTGT AACAAATCT GAAGGTATAC CTGAAGGCGT	1260
	ACTTATCCGC GCAATTGAAC CAGAAGAAGG TTTATCCGCT ATGTTCCGTA ACAGAGGTAA	1320
20	GAAAGGCTAC GAGGTAACGA ATGGCCCAGG AAAATGGACT AAGGCATTTA ACATTCCACG	1380
	GGCTATCGAT GCGGCTACGT TAAATGACTG TAGATTGTCT ATTGATACTA AGAATCGTAA	1440
	ATATCCTAAA GATATTATTG CTAGTCCACG AATCGGTATT CCAAATAAAG GTGATTGGAC	1500
25	ACATAAATCT TTACGTTACA CAGTGAAAGG TAATCCATTT GTGTCTCGCA TGCGTAAATC	1560
	AGATTGTATG TTTCCCGAAG ATACTTGGA AATAATGCCA TCTTTCATTG ATTACTATCA	1620
30	TGAAAATGAA ATCTATCTCC TTATAAGTCA ATCAATCGTG CCGTCAACAT GCGGATGGGT	1680
	TGATTGTTTT TCTTTGTATC CATCATATTT TTTGATTGAT CTCCTCTTAT TGAAC TTGT	1740
	CTTAATTATA AAATATAACA ATAGAATTAT TTATAATTAT TAAATTTAGA TGCATTAATA	1800
35	TTATTGATAT TATTTTCAAA AACTAGAAAT ATTGATTGTT TGCATGTATA ATGTTAAAAG	1860
	CGCCCTTTTA TAACGCTTAC ATATAAAAGC TTATTTAGGG AGAGGGATAT TCAACAAGGG	1920
	GGATTGAAA ATGATAGAAC TTAATGCAAT TACAACATTA TGTTTAGCTT GTATCCTTTA	1980
40	TTTACTTGGT AAGGCTATCG TTAATCACGT TAATTTTTTA AAACGTATTT GTATACCAGC	2040
	ACCAAGTATT GCGGCTTAA TCTTTGCTAT TTTAGTTGCG GCTTTGGATT CATTTGGCAT	2100
	GGTTAAGATT AAATTAGATG CTTCAATCAT TCAAGATTTT TCCATGTTAG CATTCCTTAC	2160
45	GACAAATCGGT CTGGTGCAT CATTGAAATT ATTTAAATTA GGTGGCAAAG TCTTGCTATT	2220
	ATACTTTATG TTTTGTGCTA TCATTTCACT CATTCAAAAC ATAGTTGGTG TATCACTAGC	2280
	AAAAGTATTA AATATTAAAC CTTTGTTAGG ATTAACAGCA GGTTCATGT CTATGGAAGG	2340
50	CGGTCATGGT AATGCTGCTG CTTATGGTAA GACAATTCAA GATTTAGGTA TTGATTCGGC	2400
	ACTGACAGCG GCTCTTGCAG CTGCAACTTT AGGTCTTGTA TTTGGAGGGC TTATCGGTGG	2460

55

	ATTTAAAGAT TATAGCCAAG TAGCATATAA CGAACATTTA CATAGTAAAT TTAATGCCAC	2580
	TGAAGTATTC TTCATTCAAT TTACAATCGT TGTATTCTGT ATGGCAGTTG GAAGTTATTT	2640
5	CAGTCATTTG TTTACAGCTC AAACAGGGAT TAATGTTCCA ATTTACGTTG GCTCATTATT	2700
	TGTAGCTGTT ATTGTCCGAA ATATCTCTGA AAGTTTTAAT TTTAATATTG TAGATTTAAA	2760
10	AATTACTAAT CAAATTGGCG ATGTCGCATT AGGTATTTTC TTATCTCTTG CGCTAATGAG	2820
	CATTCAATTA ATCGAAATTT ATAAACTTGC TATACCTCTT ATTATTATCG TTTTAGTTCA	2880
	AGTTGTGCTT ATGATTTTAT TTGCTGTTTT AATTTTATTT AGAGGTTTAG GAAAAGATTA	2940
15	TGATGCTGCA GTAATGGTAG GTGGTTTTAT CGGTCATGGG CTTGGTGCAC GCCAAATGCC	3000
	ATGGCAAATT TAGATGTTAT TACTAAAAA TATGGAACT CACCTAAAGC ATATTTAGTT	3060
	GTACCTATTG TTGGTGCATT CTTAATCGAT TTAATTGGTG TTATAGTCAT TATGGGATTC	3120
20	ATACAATGGT TTAGTTAAAC ACCAACTCA TAAATAAAG AGGAGGCCTT CGCCTCCTCT	3180
	TTTATTTATC CTCGATGTAT ATTCAAGTTA CGTTGTTCTA TCCATGACAA TATTTCCGGA	3240
	CTAAATACGA TTTGTTTTTG TGTAAAGTCG TCAATATTTT TAGCATCTAA CATCGTCATT	3300
25	ATTGATTTCA TGTGTTCAAT AAATGATTCT ACATAAGCTA CTGTATGTGC AATGCCATTA	3360
	TTTTCAACTT GATTTAAAAA CGGACGTGAC ATACCAAGTTG CCTTTGCACC AAGTGCTAAA	3420
	CTTTTAATTG CATCGAGTGG TGTACGTAAA CCACCACTCG CGAAAACTGA AATTTGCTT	3480
30	TGATAAGCCG TTGTTTCAAG TAATGACTCA ACTGTAGACT GTCCCATGA TGATAAGTAA	3540
	TCCATATCTT TATTTGCACG ACGTTCATTT TCAATATCTA CAAAGTTAGT ACCACCTTTG	3600
35	CCACTAACAT CGACATACTT GACGCCTATT TGTGTGAAGT CATGCATTAA TTCTTTGCTC	3660
	ATACCAAATC CAACTTCTTT TATAATGACT GGAACAGACA CTCGTGATAC AATCGACGCT	3720
	ATAATATCTA ACCAAGTCAC AAATTCACGA TTCCCTTCAG GCATAACTAA TTCTTGAGGA	3780
40	GAATTAACAT GGATTTGTAA CGCTTGCGC TCAAGTAATT CAACTGCTTC CAAAGCCTTT	3840
	TCTACTGGTA CGTCCGCACC AACATTGCTA AAAATCATGC CTTCAGGATT CATTTTTCGC	3900
	GCAATCGTAA ACGTCTCAGC CATGCGTGGT TTTCTCAATG CCGCATGTGT TGATCCAACT	3960
45	GCCATCGCTA AGCCAGTTTC TCTTGCAACT ACAGCTAGCT TTTCATTGAT GTTTTTCGTC	4020
	CACTCGCTAC CACCCGTCAT TGCATTAATA TAAACCGGAT ATGCCATCGT TAAGTCAGGC	4080
	GTCTGTGATG TCAAATCGAT ATCATTTACA TTAATTGATG GGATAGAATG ATGCACAAAA	4140
50	CGCATCTTAT CAAAATCTGA ATGCATTGCG TCAGATTGGG CCATTGCTAT TTCAACATGT	4200
	TCATTTTTTC TCTGTTCTCT TTGAAAATCA CTCATGATTA AACCTACCTT TTCGTCATTT	4260
55		

ATTACAGCTA AGCAAATATA ATATCCATAA TGTAAATGTA ATGCCGGCAT ATTTACAAAG 4380
 TTCATACCAT AAATCCCAGC TATGAATGTT AACGGTGAAA ATATAACTGA TACTAATGTC 4440
 5 AGTACTTGCA TAATACTATT CATTCTAAAT GACGTGTATG ACTCAAAATT TTCTCGTATT 4500
 TCGTTTGTC TTTCTTGAGC AGTACGAATG ATATTACGTT GCTTAATCAA GTGGTCATCG 4560
 10 ATATGTTGAA TGTATAGCGA ATGTTTATTA TCTATAATCA AATCACCATT TTGTTTCATT 4620
 GTATCAATTA GCTCTTGCA AGGAAACAGT ACACGTTTTA CTTTAATCAA ATCCGAACGT 4680
 AACTTAAAGA CACTATCCAT GACCATTTTA TTAAAGCGAT CATCTACATG GCGGTCTTCA 4740
 15 AAATGATAAA CACTATCTTC AAGTGCATAT ACAAAGTTGA AATATTTATC AACCATCATA 4800
 TCTAAAATTA ATATGACGAC ATCTGCACAA TCTAATTCTG CATCTAATGT ATTATATAC 4860
 TTATAGACTA CTTTATTTAA TGATTCCAAC GTTTGATGAT GATATGTTAC TAATACATTG 4920
 20 TCTTGATATA AAATATTTAG TGCTATTGGT GAATAGTTTG ACCCCATAAT ACTATGGAAT 4980
 ACTAAGTATT GATAATCTTT ATAAGATTTA TATTTAGCTC GTGGCATACC GTTAATTGCA 5040
 TCATCCACTT CTAAATCATT AAAATTAAAA TGTGCTTTAA ACCATTCATT TTCTTGTTCA 5100
 25 TTCGGTTCAT CAAAATCATA CCAAACAATA GTCGCATCTT TTGGTATCTC TTGATATCA 5160
 TCAACTACTT TAAACGGTTC ATATGTAGTT TGATACCGTA TCTTTAAAGC CATCGATACT 5220
 CCCCCTAAAT AACGAATTCT CTATTATTTT ATCATGAATT AAATAACGTG TATGTCTTAA 5280
 30 TTTATTTTAG TATGATAGTC ACTAAGGAGA TGGTTATTAT CAAACAACTT TTTACACATA 5340
 CTCAAACCGT AACATCTGAA TTCATTGACC ATAACAATCA TATGCATGAT GCAAATTATA 5400
 35 ATATCATTTT TAGTGACGTC GTGAATCGTT TTAATTACAG CCACGGTCTT TCTTTAAAAG 5460
 AACGCGAAAA TTTAGCATAT ACGCTATTTA CACTAGAAGA ACATACGACA TACCTCTCAG 5520
 AATTGTCTCT TGGCGATGTA TTTACTGTTA CTTTATATAT TTATGATTAC GATTATAAGC 5580
 40 GGTGTCATTT ATTTTAAACA TTAATAAAG AAGATGGTAC ACTAGCATCA ACAAATGAAG 5640
 TAATGATGAT GGAATTAAT CAGCACACAC GTCGTTCTGA TGCTTTTCCT GAATCATTTT 5700
 CAACACAAAT AGCACACTAT TATAAAAATC AATCAACTAT CACTTGGCCT GAACAATTAG 5760
 45 GACATAAAAT AGCAATTCCA CACAAAGGAG CATTAAATG ACAGATGCAT TACAACAAAA 5820
 GATTCATATC GAATTACTAG ATTTATTAGA TGATGTTAAG TTTGAATTAA CAGAATTAAA 5880
 TGCACAAAAA GGGTTATACA TTAACGGACC AGCAAATCAG CACTTAAGC GTGGCGTGCA 5940
 50 TATGGCTTAT GTTCAAGGAC AAAAGCAAGC CATCGATAAT ATTATGACTA TTGTGGAACA 6000
 ACAGCTTGAA AGATCAACAT TTCCTAGAAC ATTATGATAA ATTTCAAAT GAGGTTGCTC 6060

55

5 ATAATTTTTT AGATCAATTT TATCAAATTA AAGGGCAATA CTTTATCATC ACACATATCA 6180
 ATACACTTAT TGGTGATTTT CACTCAGAAG CTCATTAACA ATTAGTCTAT ATAACCCTTG 6240
 CTATATTTTC AAAAACAAAA CCCAATTACG TTTTCATGTC AAATATCATC TTGCATGAAA 6300
 TCGTAACTGG GTCATTTATA TGTATTAGT TATTTTGTGT TACATCCTCA TCTATCGATT 6360
 10 TGGCAATTTG TTTAATAGCT TTATGTGATT GTCTAATTGG ATAAATTGGA AAATCATGTA 6420
 CCATCTTAGG ATAATCATAA AACTCAATGT ATTGATGATG TTGCAACATC ATTTGTTCAA 6480
 ATAGCTTCAT ATCAGGATGT GTCATTTTAC GTCCACCACC AAACATATAA ACTGGTGGCA 6540
 15 ATCCTTCTAT TGTGCCATTA ATTGGCGATA TGCCTTATC TGTTAATGGT AGGCCATTCTG 6600
 CCCATTTTTT CATAATCTCA TTGACACCAA ACTGACTTAG aACCGCATCT TGTTCGATTA 6660
 AGGCGTCCGA AATATCTTTA TTAGATAGTG TTGCATCTAA AATTGGTGAG ATTAAATACA 6720
 20 ATTTATTCGG TAATGGCTGT TGATTAKCTA AAAGAGATTG TACAAAGGAT AATGCCAGTG 6780
 CACCACCTGA ACCATCACCC ATGACTACGA CATTTTGTATG TCCTACTTCA GATACTAATT 6840
 GATCATAAAC ACGTTGTATC GCTTGGA AAA GTATCGTCAa TATGnaAACT CTGGTGTCTT 6900
 25 TGGATAGATA GGCAGTACAA CCTCATATAA TGtACTTAAA GTGATTTTAT CCCAACAAATC 6960
 TCCAATGGAA CGGTGATGGT TGTAGTGCAT TGAATCCACC GTGAATATAT AAAATTTTCT 7020
 TATCAATTTG ATGTCTGAAA TTAAAGCGAA AGACTTGCAT ATCATCTAAT GACAATTTTT 7080
 30 CTAAATTTGC TTTAACATTT AATGTTGAAG GCTGCTTATG TTTTTTCTA TTTTCAATTT 7140
 CTCTTTTATA AAAAAATCTT TCAACATCTT GATCATTTTT AAACATAATC GAGCGATTGT 7200
 35 GAAGCAAATA TTTATTGACA ACGCTATTCA TAACACGGT TCTAATCAAT GTCTTAACCT 7260
 ACCTTTATAT ATTTTATGTA TCCAATGATk GTCTATCCCC TACATTCTTT GCCAAAAAAA 7320
 GTATATAATG TAGAAGATAT TTTCTTTTTC ACTTTCAAAT TTAAGACTAC AATTGAACAG 7380
 40 TGATTTTTCA TCATTATAAC AGACAAGTAG ACATATTGAT AAGTAAAGAA AAGAACTTTA 7440
 TACGGAGGTA CCTTGCATGA CAAATCCAAA TCAACGATTA GAACCATTG ATGAGACATT 7500
 TCAACAACCG AATATTCATC GTGGTAAGCG ATATGGTAAG AAAAAACGTT CATTGGTAAG 7560
 45 CATGATTATT CAAATCATTG TTGTWATATT AACCACCATC GCTGGAATAC AGCATGGTGG 7620

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 9834 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

	GTCATTACCG amTTTCTTAG AaTCATTAA AGATGATAAA TATACAAACG TTGGTAATTT	60
5	AAAAGAAGTG AATTTTGATA AAATTGCTGC GACGAAACCC GAAGTAATCT TTATCTCTGG	120
	ACGTACAGCT AATCAAAAGA ATTTAGATGA ATTCAAAAAA GCTGCACCTA AAGCGAAAT	180
	TGTTTATGTT GGTGCAGATG AAAAGAACTT AATTGGTTCA ATGAAACAAA AACTGAAAA	240
10	TATCGGAAAA ATTTACGATA AAGAAGATAA AGCTAAAGAA TTAAATAAAG ATTTAGATAA	300
	CAAAATTGCT TCAATGAAAG ATAAAACGAA AACTTCAAT AAACTGTTA TGTATTTACT	360
	AGTTAACGAA GGTGAATTAT CAACATTTGG ACCTAAAGGT CGTTTTGGTG GATTAGTTTA	420
15	CGATACATTA GGATTCAATG CAGTTGATAA AAAAGTAAGT AATAGCAATC ATGGACAAAA	480
	TGTTTCTAAC GAATATGTTA ATAAAGAAAA TCCAGATGTT ATTTTAGCGA TGGATAGAGG	540
20	TCAAGCGATA AGTGGTAAAT CAACTGCGAA ACAAGCATT AATAATCCTG TATTAAAAA	600
	TGTTAAAGCA ATTAAGAAG ACAAGTATA TAATTTAGAT CCTAAATTAT GGTACTTTGC	660
	AGCTGGATCA ACTACAATA CAATTAAACA AATTGAGGAA CTTGATAAAG TTGTAAAAA	720
25	ATTTTAAAG AGGGGAACAA TGGTTAAAGG TCTTAATCAT TGCTCCCCTC TTTTCTTTAA	780
	AAAAGGAAAT CTGGGACGTC AATCAATGTC CTAGACTCTA AAATGTTCTG TTGTCAGTCG	840
	TTGGTTGAAT GAACATGTAC TTGTAACAAG TTCATTTCAA TACTAGTGGG CTCCAAACAT	900
30	AGAGAAATTT GATTTTCAAT TTCTACTGAC AATGCAAGTT GGCGGGGCCC AACATAGAG	960
	AATTTCAAAA AGGAATTCTA CAGAAGTGGT GCTTTATCAT GTCTGACCCA CTCCCTATAA	1020
	TGTTTTGACT ATGTTGTTTA AATTTCAAAA TAAATATGAT AGTGATATTT ACAGCGATTG	1080
35	TTAAACCGAG ATTGGCAATT TGGACAACGC TCTACCATCA TATATTCATT GATTGTAAAT	1140
	TCGTSTTTGC ATACACCGCA TAAGATTGCT TTTTCGTTAA ATGAAGGCTC AGACCAACGC	1200
40	TTAATGGCGT GCTTTTCAAA CTCATTATGG CACTTATAGC ATGGATAGTA TTTATTACAA	1260
	CATTTAAATT TAATAGCAAT AATATCTTCT TCGGTAAAT AATGGCGACA scgTGTTC	1320
	GTATCGATTA ATGAACCATA AACTTTAGGC ATAGACAAAG CTCCTTAACT TACGATTCCT	1380
45	TTGGATGTTT ACCAATAATG CGAACTTCAC GATTTAATTC AATGCCAAAT TTTCTTTGA	1440
	CGGTCTTTTG TACATAATGA ATAAGGTTTT CATAATCTGT AGCAGTTCCA TTGTCTACAT	1500
	TTACCATAAA ACCAGCGTGT TTGGTTGAAA CTTCAACGCC GCCAATACGG TGACCTTGCA	1560
50	AATTAGAATC TTGTATCAAT TTACCTGCAA AATGACCAGG CGGTCTTTGG AATACACTAC	1620
	CACATGAAGG ATACTCTAAA GGTGTTTAG ATTCTCTACG TTCTGTAAAT TCATCCATTT	1680

55

	AGTGTTCCTTT TTGAATAATG CTATTACGAT AATCTAACTC TAATTCTTTT GTTGTAAGTT	1800
	TAATTAACGA GCCTTGTTTCG TTTACGCAAA GCGCATAGTC TATACAATCT TTAACCTCGC	1860
5	CACCATAAGC GCCAGCATTG ATATACACTG CACCACCAAT TGAACCTGGA ATACCACATG	1920
	CAAATTCAAG GCCAGTAAGT GCGTAATCAC GAGCAACACG TGAGACATCA ATAATTGCAG	1980
	CGCCGCTACC GGCTATTATC GCATCATCAG ATACTTCGAT ATGATCTAGT GATAATAAAC	2040
10	TAATTACAAT ACOGCGAATA CCACCTTCAC GGATAATAAT ATTTGAGCCA TTTCTTAAAT	2100
	ATGTAACAGG AATCTCATTT TGaTAGGCAT ATTTAACAAC TGCTTGACT TCTTCATTTT	2160
	TAGTAGGGGT AATGTAAAAG TCGGCATTAC CACCTGTTTT AGTATAAGTG TATCGTTTTA	2220
15	AAGGTTCAATC AACTTTAATT TTTTCATTTG GGATAAGTTG TTGTAAAGCT TGATAGATGT	2280
	CTTTATTTAT CACTTCTCAG TACATCCTTT CTCATGTCTT TAATATCATA TAGTATTATA	2340
20	CCAATTTTAA AATTCATTTG CGAAAATTGA AAAGAAAGTA TTAGAATTAG TATAATTATA	2400
	AAATACGGCA TTATTGTCGT TATAAGTATT TTTTACATAG TTTTTCAAAG TATTGTTGCT	2460
	TTGTCATCTC ATATTGTCTA ATTGTTAAGC TATGTTGCAA TATTGTTGT TTTTTGTAT	2520
25	TGAATTGCAA AGCAATATCA TCATTAGTTG ATAAGAGGTA ATCAAGTGCA AGATAAGATT	2580
	CAATGTTTG GGTATTCATT TGAATGATAT GTAGACGCAC CTGTTGTTTT AGTTCATGAA	2640
	AATTGTTAAA CTTCGCCATC ATAACCTTCT TAGTATATTT ATGATGCAAA CGATAAAACC	2700
30	CTACATAATT TAAGCGTTTT TCATCTAAGG ATGTAATATC ATGCAAATTT TCTACACCTA	2760
	CTAAAATATC TAAAATTGGC TCTGTTGAAT ATTTAAAATG aTGctACCGC CAATATGTTT	2820
	TGTATATTTT ACTGGGCTGT CTAAGAGGTT GAATAATAAT GATTCAATTT CAGTGTATTG	2880
35	TGATTGAAAA CAATTAGTTA AATCACTATT AATGAATGGT TGAACATTTG AATACATGAT	2940
	AAAGTcCTTT GATATTGAAA ATTAATTTAA TCACGATAAA GTCTGGAATA CTATAACATA	3000
40	ATTCATTTTC ATAATAACA TGTTTTTGTA TAATGAATCT GTTAAGGAGT GCAATCATGA	3060
	AAAAAATTGT TATTATCGCT GTTTTAGCGA TTTTATTTGT AGTAATAAGT GCTTGTGGTA	3120
	ATAAAGAAAA AGAGGCACAA CATCAATTTA CTAAGCAATT TAAAGATGTT GAGCAAAAAC	3180
45	AAAAAGAATT ACAACATGTC ATGGATAATA TACATTTGAA AGAAATTGAT CATCTAAGTA	3240
	AAACTGATAC AACTGATAAA AATAGTAAAG AATTTAAGGC ACTACAAGAA GATGTTAAAA	3300
	ACCATCTCAT ACCTAAATTT GAAGCATATT ATAAGTCAGC AAAAAATTTG CCTGATGATA	3360
50	CAATGAAAGT TAAGAAATTA AAAAAAGAAT ATATGACGCT TGCAAATGAG AAGAAGGATG	3420
	CGATATATCA ATTAAAAAA TTCATAGGTT TATGTAATCA ATCTATCAAG TATAACGAAG	3480

55

	AATTAGCTGA TAATAAAAGT GAAGCAACTA ATCTTACGAC AAAATTAGAA CATAATAATA	3600
	AAGCGTTAAG AGATACTGCG AAGAAGAACC TAGATGATAG TAAAGAAAAT GAAGTAAAAG	3660
5	GCGCGATTAA AAATCACATT ATGCCAATGA TTGAAAAGCA AATTACCGAT ATTAACCAAA	3720
	CTAATATTAG TGATAAGCAT GTTAATAATG CAAGGAAAAA CGCAATAGAA ATGTATTACA	3780
	GTCTGCAGAA CTATTATAAT ACACGTATTG AAACAATAAA GGTAGTGAG AAGTTATCAm	3840
10	AAGTCGATGT AGATAAGTTG CCGAAAAAGG GTATAGATAT AACTCACGGC GATAAAGCCT	3900
	TTGAAAAAAA GCTTGAAAAA TTAGAAGAAA AATAACTATA ATCATTTTTC AAAGTTAAAA	3960
	ATTTTGAATT TATGGTTAAC ATGTCAACTT ACTATGTGTA TAATGGTAAA CATTGATATT	4020
15	AACTATATGT ATAAAAATGT CACGCAGATG CTATTTAAAT GTGATAAATA TTTTLAGAGG	4080
	TGAATAGAGT GGCTATAAAG CTAAGTTCAA TTGACCAATT TGAACAGGTT ATTGAGGAAA	4140
20	ATAAATATGT TTTTGTATTA AAACATAGTG AAAGTTGTCC AATATCGGCA AATGCGTACG	4200
	ATCAATTTAA TAAATTTTTA TATGAACGCG ATATGGACGG TTATTATTTG ATTGTCCAAC	4260
	AAGAACGCGA TTTGTCAGAT TATATTGCTA AAAAAACGAA CGTTAAACAT GAATCACCTC	4320
25	AAGCATTTTA TTTTGTAAAT GGTGAAATGG TTTGGAATCG AGACCACGGT GATATCAATG	4380
	TGTCGTCATT AGCACAAGCA GAAGAATAAT GAAACTATAG GGTGGAACA TTTTGCCTTA	4440
	CACTACTAGA CGTGAATAGC ACAACTTAAA TTCGTGTGAA TCAGAGTAGT TTGGCTATAA	4500
30	TGATGTTCTG ACCTTTTATT TTATGTCACC TTTAGAAGCA GTTAAGTTAG TACTTTTTTA	4560
	CAAACATATG TATAATATAT TCGAGTATTT TTATTGAAAa tATTTTGGAA AACGACGAAT	4620
	CCAATAAGAA AATTTAAACA TGATTTGTAA GTTAGTTTAA TAGGAAATAT ATGCTAAACC	4680
35	AAAAGAAGCA TATTGTTATT TACTGGAATA ATTAATAATC ATGTCATGTT AAATGTTAGC	4740
	ATATAATCAC GAGATAAAAT CTAAATTTA AGATTAACTT TTTATGAATA AAAACGTAT	4800
40	CACAACAAAT AATAAAGTAA GGTGGTCAAG GTTATGAAAG TATTAGTAGC CATGGATGAG	4860
	TTTCATGGAA TTATTTCAAG TTATCAAGCT AATAGATATG TTGAAGAGGC AGTTGCAAGC	4920
	CAAATTGAAA CTGCAGATGT AGTTCAAGTA CCATTGTTTA ATGGAAGACA TGAATTATTA	4980
45	GATTCTGTAT TTTTATGGcm ATCTGGGcaA AAGTATCGTA TACCAGTACA TGATGCAGAT	5040
	ATGAATGAAG TTGAAGGTGT TTACGGACAA ACTGATACAG GGATGACCGT TATCGAGGGG	5100
	AATTTATTTT TAAAAGGTAA AAAACCAATT GTTGAACGAA CAAGTTATGG TTTAGGAGAA	5160
50	ATGATTAAAC ATGCATTAGA TAACGACGCA AAACATGTTG TAATTTCACT AGGTGGGATT	5220
	GATAGTTTTG ATGCTGGTGC AGGTATGTTA CAAGCATTAG GTGCTCAATT CTATGATGAC	5280

55

	GATATGTCGA ACTTACACCC TAAAATGGAA ACAGCAAGAA TTCAAGTAAT GTCGGATTTT	5400
	TCAAGTCGAT TATATGGTAA GCAAAGTGAA ATCATGCAAA CTTATGATGC GCATCAGTTG	5460
5	AATCATAATC AAGCAGCAGA AATCGATAAT TTAATTTGGT ATTTTAGTGA GTTATTTAAA	5520
	AGTGAATTGA AAATTGCAAT TGGTCCAGTT GAACGTGGTG GTGCTGGTGG TGGAAATTGCA	5580
	GCAGTCTTGA ATGGACTGTA TCAAGCTGAA ATATTAAACCA GTCATGCATT AGTAGACCAA	5640
10	CTAACACATT TAGAAAATTT AGTTGAACAA GCGGATTTAA TTATTTTGGG AGAAGGATTA	5700
	AATGAAAATG ATCAGTTGCT AGAAACGACA ACATTGCGTA TTGCAGAACT TTGTCATAAA	5760
	CATCAAAAGG TTGCCATTGC AATTTGTGCA ACTGCTGAAA AGTTTGATTT ATTTGAATCA	5820
15	CAAGGGGTTA CAGCAATGTT TAATACATTT ATCGATATGC CAGAACTTA TACTGACTTT	5880
	AAAATGGGtT ACAAATTAGG CATTATACGG TTCAGTCTTT AAAACTGTTG AAAACACATT	5940
20	TTAATGTTGA GGTTTAGTAA AGAAGGACTA AATTGGTGAT GCTGTCATGA TGGTTAATAA	6000
	CATTTATGAT GGTTAGCAAA ACGAATTAGA AGATCGAAAG TATACGTAAA AAATATGAAA	6060
	AATCAGCCTA TCATTGCACT GAATGTTAGC GTGATTTTAA TATATTAATT AAGCCTGAGT	6120
25	TGAACTAGTA TATAATCGTT GGTTTTTAGT GATTTTCAGC GATATCTTCT ACAATTCCAA	6180
	TGATTACTTG TACTGCTTTT TCCaTAACAT CAATGGATGC aTATTCATAT GGGCCGTGGA	6240
	AGTTACCGCA ACCTGTAAAG ATGTTTGGAG TTGGTAACCC CATAAATGAC AATTGTGAAC	6300
30	CATCTGTACC ACCGCGAATA GGTTCAGTGT TTGCTGGAAT ATCTAATTTG GCAAAGACAC	6360
	GTTTAGGTAT ATCAATAATA TGAGGCAATG GTAATATTTT TTCTGCCATA TTGAAATATT	6420
	GATCCGATAT ATCAACTTTA ACTGGATAAT TTTCAAAATG GGCATTGATA TCGTCACGTA	6480
35	TTTCTAAAAT ACGTTTCTTA CGCAATTCGA ATTGTTTTTT ATCATGATCA CGAATAATGT	6540
	ATTGCAAAGT TGCTTTTTCA ACAGTTCCTT CAAAGTTCAT TAAGTGATAA AAGCCTTCGT	6600
40	ATCCTTCTGT TCGCTCCGGA ACTTCACTAT CAGGTAGCAA ACTATCGAAT TGTTCACCTA	6660
	AACGTATTGC GTTTACCATT GCATTTTTAG CTGAACCAGG ATGAACATTT ACACCGTGGC	6720
	ATGTAATAAC CGCTTCAGCA GCGTTAAAGC TTTCATATTG TAATTCTCCA TATTGACTAC	6780
45	CATCCATAGT ATAAGCAAAA TCAGCATTGA AGCGGTCAAC ATCAAATTTA TGTGGACCAC	6840
	GACCGATTTC TTCGTCTGGT GTAAATCCAA TGCGAATGGT ACCATGTTTA ATTTCTGGAT	6900
	GTTCTTGTA ATAACAAATA GCTTCCATAA TTTCCACAAT ACCCGCTTTA TCGTCTGCAC	6960
50	CTAGTAACGA TGTACCATCA GTTACCATTA ATGTATGACC AACTAAACTG TTAAGTTCTG	7020
	GAAATACTTT AGGATCTAAG ACACGTTTAG TATTGCCTAG TTTGTATGGC TTACCATCAT	7080

55

	GCGCCAAAAA TCCAACTGTT GGGACGTCGA CATCGATGTT ACTTTCTAAT GTAGCAAATA	7200
	AGTAGCCATT TTCATCTAAA TCAGTTGGCA ATCCTAATTG TTGTAATTCT TTTTCTAATA	7260
5	AATGTAACAA ATCCCATGTC TTTTCAGTTG AAGGTGTTGT TGTAGATTTT GGATCAGATT	7320
	GCGTATCAAT TGTGCTATAT CTTGTTAATC TATCTATCAA TTGGTTCTTC ATTATATTCTG	7380
	ACCCCTTAAA CTCTATTATT CATGTTGTAA GATTTTTTAT ATGTCTTACC TTTGATTTTA	7440
10	CCATACAGTT GTTTGATACG TGTGTATAGG TAATATAGAA TTTCAGAAAC TAATATACCG	7500
	AAAGCAATCG CACCTGAAAT CAGTGTAATT CTAAAAATGT ATTTACAGCA CTTGTATAAT	7560
15	CATTTGATAC TAAAAAACGA GTCGCTTGAT AAGCTGCACC ACCAGGTACT AATGGTATAA	7620
	TGCCTGGCAC TATGAATATA ATTACCGGTC GTTTATATCT GCGACTCATA GTATGACTCA	7680
	TTAAGCCTAA AATTAAGCTT CCCAAAAATG AAGCGCCAAC TTTTCCAAAC TCTAAATCTA	7740
20	CCGTTAATTG GTAAATCGTC CATGCAATGG CACCCACAAA TCCACATGCT ACTAAGAGGC	7800
	GTTTGGGTGC ATTGAAAATG ATAGAGAAAA GTACTGTTGA TATAAAGCTG ATTGTAAAAT	7860
	GAAATAATA AAATAGCATG CTTTAACAGT CCTTCCTTAA ATGATTAATA AAACGATTGC	7920
25	GACACCAGCA CCGATTGCGA ATGCTGTTAA TGCAGCTTCA ACACCGCGAG ACATACCTGC	7980
	AAGTAATTCA CCCGCTAATA AATCTCGAAT GGCATTGGTA ATTAATATAC CAGGGACAAG	8040
	TGGCATGACA CTGGCTATAG TAATGATATC TTGATTGGTT GCAATGCCTA ATTTAGTAAA	8100
30	TGTGGCTGCA ATGGATATGA CCACAGCGGC TGCAACAAAC TCTGAGAAAA ATTTAATTTG	8160
	TATATAGCGT TGCACAAAGC TGAATGTTAA AAATGCGGAT CCGCCAGCAA TGACTGCAAT	8220
	CCAACAATCT GATGCGACAC CACCAACAT AAATAGGAAG AAGCCACATG CAATGGCAGC	8280
35	TGCAAAGAAA TTCGTTAAAA AAGAATATTG TAATGATGCA TGCTGTAAAT GAATAAATTC	8340
	AGATTTAGCT TCATCAATTG TGAGTTCTTT ATTTGATATT TTACGTGAAA GACTATTCTG	8400
40	TAAAGCGATT TTCTCTAAAT CTGTTGTACG CTCTGTACA CGAATTAATC TTGTACTTGT	8460
	TCGATCGTTT AATGAAAAAA TAATTGCAGT TGAACGACA AACTATATG TATTATGAAG	8520
	ACCATACTA TGTGCGATAC GGTTCATTGT ATCTTCAACT CGATATGTTT CAGCACCTGA	8580
45	TTCaAGTAAA ATTCTACCTG CAATTAATAC AACATCAATC ACTTTGTTTT CATCTATAAT	8640
	TGTGATTGAA TCTGGCATAT CAATTCACCT CCAATGATAT GTGTTATTTA TTTGAACAAT	8700
	TGaAGTTTAC AACTTGTGTG TACAACCTTC AATAGTGAGA CTTTGTGTTA GTATGATGAA	8760
50	CTTGTATGGT TCAAATTTAA ATAAGAAAAA CTGTTAATCT TTGCTATTAT ACTATGATTT	8820
	AATAATAGCA AAGGATTAAC AGTTTGTGCG TTGTTATAAA TTGATAATAG GGTTAAACAT	8880

55

TTTACGCTGT GATTTTGGAT CGTCATCTGT TAAATAACCA ACACCGATAG ACACTGACAA 9000
 TTTAATAACT TCTTTGTTTG GTAAATGGAA TGATGATTTT TCAACACCCG AACGAATATT 9060
 5 TTCAGCTAAT TTAACACTTT GATCAAGTGA ATAATTGTGA ATGACAACCTG AGAACTCTTC 9120
 GCCACCATT CTAAAAATTT TAAATTGATT CGGCACATAG TTTTAAAGTA ATTGAGACAT 9180
 TTGTTTAAAT ACAGCATCAC CTGATTTGTG TGAGTAGGTA TCATTGacAT CTTTAAATCC 9240
 10 ATCGATATCG ATTAATAATA ATGCGATACT TTGATGTTCT TTTTCAGCTT TTCGTGAAAT 9300
 TTCATTTAAA TGTCTATCAA ATTCTTTTAC ATTACCTAAG CCTGTTAAGT AATCATATTT 9360
 ATCTTCGTTT TCATAACGAT TTACGAGTGA GAAGAAATGC CAAATATCGA CAAATGTTAT 9420
 15 CGCTGAAGCT AAAGTGATAA TTAATGAAAT TGGTATTAAA ATGATAACTT CCGATAGTGT 9480
 GTAAATAGGA CTCACCTAACG CGACACCAAA TAAATGATT ATTGTAACAA CATTAAAGTAT 9540
 TAATAATGAT AGCACATCAT TTTGTTTTAA AAATGGTCCA ATAGCACTTG TTAGTGCAGC 9600
 20 AATAACAATC AACGTAACAC CGTACATAAT CGAGTTGTTA AATACTACAA TTTCAACAAT 9660
 TGCTACAATT ACTGTGGCAG ATAATGTATA GACCATATTT GTAAATCTAC CTAAAAACAA 9720
 25 TAAAGGAACG AATGTTAAGT GAATTAAATA ATCTTCACGA TAAGGGATAG GGTAGACAGA 9780
 TAATAATAAT GATACGATTG TCATTAAAC AGTGACATAA GCCTTAGAAA AAAC 9834

(2) INFORMATION FOR SEQ ID NO: 38:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23439 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

40 TCTCAATCAG ATGAAAAATT GCATATCGTA GGTTTTACAG AAAGTGCAA ATATAATGCG 60
 TCATCAGTCA TTTTCACGAA TGACGCTACC ATTGCCAAGA TCAATCCTAG ATTGACTGGA 120
 GATAAAATTA ATGCAGTTGT TGTACGTGAT ACAAATTGGA AAGACAAAAA ATTAAACCAA 180
 45 GAGCTTGAAG CGGTAAGTAT TAATGACTTT ATTGAAAATT TACCAGGTTA TAAACCACAG 240
 AACTTAACAT TAACTTTTAT GATTTTCATT TTATTTGTCA TTTCAGCTAC AGTTATAGGC 300
 ATTTTCCTAT ATGTCATGAC ATTACAAAAG ACGAGTTTAT TTGGCATATT AAAAGCTCAA 360
 50 GGATTTACGA ATGGCTATTT GCGAATGTG GTAATTTGCG AGACGGTCAT ATTAGCACTA 420
 TTGGTACGG CATTTGGCTT ACTGTTAACA GCGTTACAG GTGCATTTT ACCTGATGCA 480

55

TCTGTATTAG GAAGTTTATT CTCCATTTTA ACAATTAGAA AAATAGATCC GTTAAAGGCC 600
 ATTGGGTAGG AGGTGTAGCA AATGTTGAAA TTTGAAAATG TAACAAAGTC ATTTAAAGAT 660
 5 GGGAAATCGTA ACATTGAAGC GGTAAAGAT ACAAATTTTG AGATAAATAA AGGTGATATT 720
 ATAGCATTGG TTGGACCTTC TGGCTCTGGT AAAAGTACAT TTCTAACTAT GGCAGGTGCT 780
 TTACAAACAC CGACATCTGG GCACATTTTA ATCAATAACC AAGATATTAC GACAATGAAG 840
 10 CAAAAAGCAT TGGCAAAAGT TAGAATGTCT GAAATAGGTT TTATTTTACA AGCTACAAAC 900
 CTGTACCAT TTTTAACGGT AAAGCAACAA TTTACATTAT TGAAAAAGAA AAATAAGAAT 960
 GTTATGTCTA ATGAAGACTA TCAGCAACTT ATGTCACAAAT TAGGTCTAAC TTCATTGCTT 1020
 15 AATAAGTTAC CTTCAGAAAT TTCAGGTGGT CAGAAACAAC GTGTGGCGAT AgCaAAGCGT 1080
 TATATACGAA TCCGTCGATT ATTTTAGCGG ATGAACCTAC CGCGGCGTTA GATACTGAAA 1140
 20 ATGCGATTGA AGTCATTAAA ATTCTACGTG ATCAAGCCAA ACAAAGAAAG AAAGCATGTA 1200
 TTATTGTTAC ACATGATGAA CGACTTAAAG CATATTGTGA TCGTTCATAT CATATGAAAG 1260
 ATGGCGTCCT TAATCTTGAA AATGAAACAG TAGAATAGTT TTATTAAGCC GGTACATCAT 1320
 25 GTGCCGGTAT TTTTATGTTT ATGTATTATT TGAATAAACT TTCACATTCA ATTAATAATA 1380
 ATTATTATCG AAAATCAGAA ATATTCCGTG AAATATAATA TTTTGTAG TAAAATGGCC 1440
 TCTAAGTATT CAATATTAA ATATGGGGAT TGAATATAAA ATTATCGTAA TGGGGGTCAA 1500
 30 TGGTTATGGA TTTATTGATA GGTACTTTAT TTTTATTTT GGTCTTAGTG ATTTTACAT 1560
 TATTTACATA TAAAGCGCCT AATGGTATGC GTGCCATGGG AGCATTAGCT AATGCAGCAA 1620
 TCGCAACATT TTTAGTGGA GCATTAAATA AATATGTTGG TGGCGAAGTA TTCGGTATTA 1680
 35 AATTTTTAGA AGAGCTAGGA GACGCTGCGG GAGGTCTAGG TGGTGTGCT GCCGCTGGAT 1740
 TAACAGCATT AGCTATCGGT GTGTCACCAG TATATGCATT AGTTATAGCA GCCGCTGCG 1800
 GTGGTATGGA TTTATTACCA GGTTCCTTTG CGGGTTATAT GATTGGATAT GTGATGAAAT 1860
 40 ATACAGAGAA ATATGTGCCG GATGGTGTG ACTTAATTGG ATCGATTGTC ATCTTAGCGC 1920
 CATTAGCTCG TCTTATTGCA GTATTATTAA CGCCAGTAGT GAATAGTACA TTGATTGAA 1980
 45 TTGGTGATAT TATCCAAAGT AGTACGAATA CGAATCCAAT TATCATGGGT ATCATTTTAG 2040
 GTGGTATTAT TACGGTTGTC GGCACAGCGC CATTGAGTTC AATGGCATTG ACAGCATTAT 2100
 TAGGTTTAAC GGGTGTACCT ATGGCTATTG GTGCCATGGC AGCATTTAGT TCGGCATTTA 2160
 50 TGAATGGGAC GCTATCCAT CGCTTAAAT TAGGTGATCG TAAGTCTACG ATTGCAGTAA 2220
 GTATTGAACC TTTATCACAA GCAGATATTG TATCAGCCAA TCCAATTCCA ATCTATATTA 2280

55

	ATGCGACAGG TACAGCTACA CCGATTGCAG GATTTTTAGT TATGTTTGGA TTTAATCATC	2400
	CGACGACAAT TGTGATTTAT GGTGTAGTAA TGGCGATTGT AGGTGCGCTT GCAGGTTATC	2460
5	TTGTTTCAAT TGTATTTAAA AAATATCCAA TTGTTACTAA GCAAGACATG ATTAATCGAG	2520
	GTGCAGTAGA CGCATAGCAT CATCATATTG AATAGTAAAA ACAATAAAA CATAGTAACG	2580
	TGATTCAGTC GATGTAACAG TCGATAATGA GTCACGTTTT TTTATAGAAA AATACAAGAC	2640
10	ATAAAAAATGT CATAATTTAT TGTGACAAA TATCATACTG TATAAACATT TATCATTTTC	2700
	TCAAGTACCT TTTACACGAT GGAATGAACT TACTTTTTAC GAAATTATGC GTATTTTATA	2760
	AACAAATATC ATTGATATAA CGGTAAATGT AAGCGTTTAC AACAGAAATA ACAGCATGCT	2820
15	ACGATATTTT TGTAATTCA CTGATTCAAG TATTTTAAAGT CAATATGAGG AGGGATGTTA	2880
	TGAGCGATTG TGAGAAAGAA ATTTTAAAAA GAATTAAAGA TAATCCGTTT ATTTCAACAC	2940
20	GTGAACTTGC TGAGGCAATT GGATTATCTA GACCCAGCGT AGCAAACATT ATTTCAGGAT	3000
	TAATACAAAA GGAATATGTT ATGGGAAAGG CATATGTTTT AAATGAAGAT TATCCTATTG	3060
	TTTGTATTGG CGCAGCGAAT GTAGATCGTA AGTTTTATGT GCATAAAAAT TTAGTTGCAG	3120
25	AAACATCAAA TCCTGTAACG TCAACACGCT CTATTGGTGG CGTAGCAAGA AATATTGCTG	3180
	AGAACTTAGG TAGGCTTGGC GAAACGGTCG CTTTTTTATC TGCTAGTGGA CAAGATAGTG	3240
	AATGGGAAAT GATTAAACGA TTGTCCACAC CATTTATGAA TTTGGATCAT GTTCAACAAT	3300
30	TTGAAAATGC GAGTACAGGT TCATATACAG CTTTAATTAG TAAAGAAGGC GACATGACAT	3360
	ATGGCTTAgC AGATATGGAA GTGTTTGACT ACATTACGCC TGAATTTTTA ATTAAGCGTT	3420
	CACACTTATT GAAAAAGGCT AAGTGCATTA TTGTAGATTT GAATTTAGGC AAAGAGGCAT	3480
35	TAAACTTCTT ATGTGCCTAT ACCACGAAAC ATCAAATCAA ATTAGTTATC ACCACGGTTT	3540
	CTTCCCCAAA AATGAAAAAT ATGCCTGATT CATTACATGC TATTGATTGG ATTATCACGA	3600
	ATAAAGATGA AACAGAAACA TACTTAAAT TAAAAATAGA ATCTACTGAT GATTTAAAAA	3660
40	TAGCTGCTAA ACGCTGGAAT GATTTAGGTG TTA AAAATGT TATTGTGACA AATGGCGTGA	3720
	AAGAACTCAT TTATCGAAGT GGTGAGGAAG AAATCATTA GTCAGTTATG CCATCAAATA	3780
45	GTGTGAAAGA TGTTACAGGT GCAGGCGATT CATTCTGTGC TGCAGTAGTG TATAGCTGGT	3840
	TAAATGGGAT GTCTACTGAA GATATATTAA TTGCTGGTAT GGTTAACGCA AAGAAAACGA	3900
	TAGAAACGAA ATATACAGTT AGGCAAAACC TAGATCAACA GCAACTTTAT CACGATATGG	3960
50	AGGATTATAA AAATGGCAAA TTTACAAAAG TATATTGAGT ATTCTCGAGA AGTTCAGCAA	4020
	GCACGGGAGA ACAATCAACC GATTGTAGCA TTAGAATCAA CAATTATTTT GCATGGTATG	4080

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

GCCATTCCAG CAACCATAGC CATTATAGAT GGCAAATTA AAATTGGTTT AGAAAGCGAA 4200
 GATTTAGAAA TACTGGCAAC TAGTAAAGAC GTTGCTAAAG TATCTAGAAG GGATTTAGCA 4260
 GAAGTTATTG CGATGAAGTG TGTGGTGCT ACTACTGTAG CGACGACGAT GATATGTGCT 4320
 GCAATGGCTG GTATTCAATT TTTGTTACA GGAGGTATTG GGGGCGTCCA TAAAGGTGCA 4380
 GAACATACGA TGGACATTTT AGCAGACTTA GAAGAACTGT CTAAAACAAA TGTCACGTGT 4440
 ATCTGTGCAG GTGCCAAATC AATTTTAGAC TTACCTAAGA CGATGGAGTA TTTAGAAACA 4500
 AAAGGCGTTC CAGTTATTGG ATATCAAACG AATGAATTGC CAGCATTCTT CACTCGCGAA 4560
 AGCGGTGTTA AGTTAAACAAG TTCGGTTGAA ACGCCAGAAC GACTTGCTGA CATTCAATTA 4620
 ACAAACAGC AGTTAAATCT TGAAGGTGGC ATTGTTGTTG CTAATCCAAT TCCATATGAG 4680
 CATGCCTTAT CAAAAGCATA TATTGAGGCA ATCATAAATG AAGCTGTTGT TGAAGCGGAA 4740
 AATCAAGGTA TTAAAGGTAA GGACGCCACA CCGTTCCTGT TAGGGAAAAT TGTAGAAAAA 4800
 ACGAATGGTA AAAGTTTAGC AGCAAATATA AAAGTTGTTG AAAACAATGC GGCCTTGGGT 4860
 GCTAAAATTG CTGTCGCTGT TAATAAATTA TTGTAGGTGA TGATACATGA ATATTTTATT 4920
 CGCTATCACA GGGATAGCAT TTGCACTATT TGTTCGCTTT TTATTCAGTT TTGATCGTAA 4980
 AAAAATAGAC TTCAAAAAGA CGTTAATAAT GATATTTATT CAAGTGTGA TCGTGTATT 5040
 TATGATGAAC ACAACGATTG GTTTGACAAT TTTAACTGCA CTAGGTTTAT TTTTGAAGG 5100
 GCTAATAAAT ATTAGTAAAG CAGGCATAAA TTTTGTCTTT GGAGATATAC AAAATAAAAA 5160
 TGGCTTTACG TTCTTTTTAA ACGTATTACT GCCATTAGTT TTTATTCTG TATTAATAGG 5220
 CATCTTTAAT TATATTAAGG TATTACCATT TATTATCAA TATGTAGGTA TCGCTATTAA 5280
 TAAAATAACT AGAATGGGGC GCTTAGAAAG TTATTTTGCT ATTTCAACAG CAATGTTTGG 5340
 GCAACCAGAA GTATATTTAA CAATAAAGA TATTATTCCA AGATTATCTA GAGCGAAATT 5400
 ATATACAATT GCGACGCTG GTATGAGTGC TGTTAGTATG GCAATGCTAG GTTCATATAT 5460
 GCAGATGATT GAACCCAAGT TCGTAGTTAC AGCAGTAATG TTAAATATTT TTAGTGCGCT 5520
 TATCATCGCC AGTGTAATCA ATCCCTATAA ATCTGATGAT ACTGATGTTG AAATTGATAA 5580
 CTTAACGAAA TCCACAGAAA CTAAAACATT GAATGGAAAA ACAGGAAAAC CTAAGAAAGT 5640
 TGCCTTTTTC CAAATGATTG GTGATAGTGC GATGGATGGG TTTAAAATCG CTGTTGTAGT 5700
 AGCCGTAATG TTGTTAGCAT TTATTTTATT AATGGAAGCA ATTAATATCA TGTTTGTTAG 5760
 TGTGTTTGTG AACTTTAAAC AGCTTATTGG CTATGTGTTT GCACCAATCG CATTCTTAAT 5820
 GGGGATTCCA TGGAGCGAAC TGTTCCAGCT GGCTCTTTAA TGGCGACTAA ATTAATTACA 5880

	CAAGGTATCA TTTCAAGTTTA CTTAGTAAGc TTCGCTAATT TTGGTACGGT TGGTATCATC	6000
	GTAGGTTCAA TTAAAGGCAT TAGTGATAAA CAAGGAGAAA AAGTTGCATC CTTTGCAATG	6060
5	AGGTTGCTAC TTGGTTCAAC TCTAGCTTCA ATCATTTCAG GATCAATCAT TGGCTTAGTA	6120
	TTGTAAATGA ATCGAAGTAC CTAAATTAAA TTCATGGCAA AGCTAAACCC CGTCACCAAG	6180
10	TTGGCGCAAC AGCGcATgca TAACTTAGTG ACGGGGTTTT ATCATAACAA TCTACTTTTT	6240
	CGTAGCCGTT TTTGAAATGT ATGTTGATGG TTTATCTTTT TCAAAAATTG TTAATCCCGT	6300
	TATATCTTTT TTATGTTTTG AAGGGACAAT GAAGCTAAGT ATATAAGCAA AGACAAAAGC	6360
15	AACTGTAAAT GAAATGGTAG ATACATAGAA AGGTGAGTTA CCTTTGCCAA CACCATTATA	6420
	GACATAAGCA AAGATGATAC CCAATATTAA TCCACAAATA ACACCGAATG TATTCGTACG	6480
	TTTAGTGAAG ATACCAACTG CAAATACACC AGCCAATGGA ACGCCGAATA ATCCAGTCAC	6540
20	AAACAAGAAT AAATCCCATAGT AGTCATTTGA ATTAGAAGCA ATTAAGTATA GTGACATTCC	6600
	AAAACCGAAA ATACCTGCAA TGATAATAAT GAAACGTGCA AAGTTAACTT CGTGTGCTC	6660
	GCTACCTTTT CCGAAGAAGC GTTGCTTAAT GTCGATTGAA ATACAAGCAG ATATAGAATT	6720
25	TAAACTAGAT GAAATGGTAG ACTGTGCAGC GCGGAAAATG GCTGCAATAA GTAATCCTGC	6780
	TACAAATGGT GGCATCTCAG TCAAAATGAA ATATGGCACT ACAGATGATG TATTGAAGCC	6840
	TTTTGGTAAA ACAGCTTCAT GTGTATAAAA TGAATACAGC ATTGTACCCA TACCATAAAA	6900
30	TAAGGGTGCT GAAATTAAAG CTAGGATACC ATTTGTCCAT AACGATTTAT TTGTTTCTTT	6960
	TAAACTATCA GAAGCTTGAT AACGCTGCAC GACGTCTTGA CTCGCTGTGT ATTGATACAA	7020
35	GTTGTTGAAA ATATTTCCCTA GGAAAATAAT TGGAAATGGCA GCTGCCGCAG TATTTAGTTT	7080
	CCAATTGTCT GCACTAATTA ATTTTTTG TGCTCAATCGCA TCTGCAAAGA CAGTGCCGAA	7140
	ACCGCCTTTA ATGTTTCAAA CACCTAGAAT AATAATAACT AAAGCGCCGC CTAATAAAAT	7200
40	GACGCCTTGA ATGAAATCAC TCCAAACCAC ACCTTCGAAA CCACCTAAAA ATGTATATAA	7260
	AATACATAGT AAACCAACGA GTGATGCAAC GATATAAGGG TTCATGTCTG ATACAGATGT	7320
	GATTGCTAAT GTTGGTAAGT AGATAACAAT TGCAACACGC CCTAAATGGT AAACGACAAA	7380
45	TAATAATGAG CCAATGACAC GTATGCTAGG GCCAAATCTA GCTTCTAAAT ATTCATATGC	7440
	AGATGTTACC TTTAACTTTT TAAAGAAAGG GACATAGAAA TAAATAAGTA ATGGAATAAT	7500
	TGCGACGATA GCAATGTTAC CAGCGATATA TGACCAATCT GTTAAAAATG CTTTCTCTGG	7560
50	TGTCGACATA AATGTAATCG CACTTAACGT AGTAGCATAA ATTGAAAAGC CAACTACCCA	7620
	AGATGGCAAG CGACCACTTG CGGTAAAGAA ACTATTGGTA CTTTGGCTCG CGCGCTTGGT	7680
55		

	TGTGCCAAAT CCAACTTCTT TCATGGGCAA CATCCCCTTT ACAATGTATT GATTCTTTGA	7800
	TGTCTATAAA TCGTATTTTG CAATGAGTTG ATCTAATGTT TGTGATGTG CTTCGTTAAA	7860
5	AGGTTTGAAA GGTCTTTTCG GTAATCCTGC ATCAATGCCA CGATGACGTA ATATTTCTTT	7920
	CAATGTTGGA TAAATCCCCA TTGATAACAC TGTTCGATA ATGTCGTTTG AATCATGTTG	7980
	CAGTTGGTAA GCTTCTTGAA TTTGACCTTG TCGTGCTAAG TCGAAGATT TTTCTGCACG	8040
10	GCGACCATTA ACGTTATATG TAGAACCAAT TGCACCATCT ACGCCAGAAA TCGTAGCTTG	8100
	AACTAACATT TCATCAAAGC CAGATAAGAT TAATTTGTCT GGAATGCTT TTCTAATACG	8160
	TTGAGTAGG AAGAAGTTTG GCGCTGTATA TTTAACACCA ACAATTTTTT CATGATTAAA	8220
15	TAGCTCGCTG AATTGTTCAA TAGAAATATT CACACCTGTT AAATCTGGTA TTGCATAAAT	8280
	AATCATATTG TTCTGAGTTG CTTGATAAT ATCGAAATAG TAATCTCTAA TTTCTTCAA	8340
20	AGTAAATGGA TAGTAGAATG GTGTTACGGC AGAAAGTGCA TCATAACCGA GTTCTGTGGC	8400
	ATATTTTCCA AGTTCAATGG CTTCAATTAA ATCTAACGAA CCTACTTGAG CAATCAATTT	8460
	CACTTTATCC CCAACTGCCT CTTTGCCAAC CTTGAAAAC TGCTTCTTCT GCTCTGTATT	8520
25	TAATAAAAAG TTTTCGCCTG AGCTACCATT TACATAAAGA CCGTCTAATT CTTGAGTTTC	8580
	AATGGCATTG TGAGCAATTT GTTTAAGTCC TTGTTCAATT ACTTGACCAT TTTCATCAA	8640
	AGGAACGAGT AACGCTGCAT ATAAACCTTT TAAATCTTTG TTCATTATGA AGTCCCTCCA	8700
30	AAAATCATTT GATAATATAG TTTACAGCTA TAATTGTAAA CGCTATCATA AAATGTAACA	8760
	ATATCTTTTT GAAAATTGTA GTCATATTTA TGTATAATTA ATGAAAATGT TTTTCAAAT	8820
	CAATAGAAAT GGAGTGAGTA AGGTGTATTA CATCGCAATC GATATTGGAG GCACTCAAAT	8880
35	TAAATCGGCA GTTATTGATA AGCAATTGAA TATGTTTGAC TATCAACAAA TATCAACGCC	8940
	GGACAACAAA AGTGAGCTTA TTAAGTACAA AGTATATGAG ATTGTAACAG GATATATGAA	9000
40	GCAATATCAG TTGATCCAAC CTGTCATAGG TATTTTCATCA GCAGGCGTTG TTGATGAACA	9060
	AAAAGGCGAA ATTGTATACG CAGGGCCAAC CATTCCGAAT TATAAAGGTA CTAATTTTAA	9120
	GCGATTATTA AAATCACTGT CTCCTTATGT CAAAGTAAAA AATGATGTAA ACGCTGCATT	9180
45	ACTAGGCGAA TTGAAATTAC ATCAATATCA AGCAGAACGG ATCTTTTGTA TGACGCTTGG	9240
	TACAGGCATT GGGGGTGCCT ACAAGAATAA TCAAGGTCAT ATTGATAATG GTGAGCTTCA	9300
	TAAGGCAAAT GAAGTTGGGT ATTTATTGTA TCGTCCAAC TAAAATACAA CGTTTGAGCA	9360
50	ACGTGCTGCA ACGAGTGCAT TGAAAAGCG CATGATTGCC GGAGGATTTA CGAGAAGCAC	9420
	ACATGTGCCA GTATTGTTTG AAGCAGCTGA AGAAGGTGAT GATATTGCAA AACAAATATT	9480

55

	AGGGCTTATA TTAATTGGGG GCGGTATATC TGAACAAGGA GATAATCTCA TTAAATATAT	9600
	CGAGCCGAAA GTTGCACACT ATTTACCAAA AGACTATGTT TATGCACCAA TACAAACGAC	9660
5	TAAGAGTAAA AATGATGCAG CATTATATGG CTGTTTGCAA TGATAGTTGA AAGAAGGAGT	9720
	CATTCTAAAA TAGAATTTGA AACCGTTACG AGAGATGAGA GCTGTTGTTA GTTCCACACA	9780
10	TCACACTCTA TCTAGGACCA ATCTAACTA TATCAACCAA CAGTGTGCCA CGGGCAAATT	9840
	AAATTGAAGA AGCTGAGATA TTAAATTTT AGAAAATGTA AAAAAATATT TGGTATTGAA	9900
	ATTAAAAAAG CACCTAGCAA CTCGTTGGGA CAATCACGAT GATTGTCTAC AGTTGCAGGT	9960
15	GGATTTGAAT ATACTACTAG TTATTTGTTG TCTAGGATAA TAGATTTAGT ATGTTGATAA	10020
	GTTTGACTCA GATTCGTATT TTCTAATAAA TGATAACTCA CGATATCGAT TAAAAAGAGT	10080
	GTCGCAATTT GTGTGTTGAT AAATTGATGG TCGGTATTAC GCGATTGATC CGTTGTTAAA	10140
20	AGTACTAAAT CTGCACAATC TGTAAGTTTA CTACCTTCAA AATTGTGTGAT GGCAACGACA	10200
	TATGCACCAT GAGATTTGGC GACTTCCGCT GCAGAAATTA ATTCCGAAGT ATTACCACTA	10260
	TTTGACATAG CAATAAACAT ATCCGAATGA GATAGTAGGG ATGCCGATAT TTTCATTAAA	10320
25	TGTGAATCGG TAGTAACATT ACCTTTTAGC CCCATACGAA TCATACGATA ATAAAATTCA	10380
	GTCGCTGATA AACCAGAGCT ACCTAGTCCA GCAAAGAGTA TATGTCGACT TGATTGAAGT	10440
	TTGTCGATAA AGGTTTGGAT AATGTCGTTA TCAATAAATT CACCAGTTTG TTGAATGATT	10500
30	TGTTGATGAT ATTTATGAAT TCTTTGAATA ATTGGGCTAT TTCAATAAC TGTCTCTGTC	10560
	ATTTCTTGTT GAATATTAAA TTTTAAATCT TGGAAATTCT CATAATCCAG CTTATGACTA	10620
	AAGCGTGTCA TCGTTGCTGG TGATGTACCA ATCGCATGGG CTAAGGAGTT AATCGTTGAA	10680
35	AAGGCATCGC TATAACCATT TTGTCTTATA TAATTGACGA TCGGTTTATC AGTTTTTGTA	10740
	AATAAATGTT GATAACGTTG AACACGATTC TCAAATTTCA TTGTGTCACC CCTTCATCTT	10800
40	AATGATTACT ATTATATATG AAAAAATTT TCAAGATAGT AAAAAGCATT GATAAAAATT	10860
	ATCTTAATGA TATATTGTAA ATGACTTTAC GTGAAAAAAC GACTTATGGA GTGAGGAATA	10920
	ATGTTACCAC ATGGATTAAT AGTATCTTGT CAGGCACTAC CAGATGAACC ATTGCATTCA	10980
45	TCTTTTATTA TGTCGAAAAT GGCATTAGCT GCGTATGAAG GTGGTGCTGT TGGTATTCGC	11040
	GCAAATACTA AGGAAGACAT TTTAGCAATT AAAGAAACGG TAGATTTACC AGTTATTGGC	11100
	ATTGTGAAAC GTGACTATGA TCACTCAGAT GTTTTCATTA CTGCAACGTC AAAAGAAGTT	11160
50	GATGAACTGA TAGAAAGCCA ATGTGAAGTC ATTGCATTGG ATGCAACGTT ACAGCAACGT	11220
	CCGAAAGAAA CGTTAGACGA ATTAGTATCA TATATTAGAA CACATGCACC GAACGTTGAA	11280

55

TATATTGGCA CGACGTTACA TGGCTATACT AGTTATACGC AAGGACAATT ACTTTATCAA 11400
 AATGACTTCC AATTTTTTAAA AGATGTACTA CAAAGTGTTG ATGCAAAAGT TATTGCGGAA 11460
 5 GGTAATGTCA TTACACCGGA TATGTATAAA CGTGTGATGG ACTTAGGCGT TCATTGTTCA 11520
 GTCGTTGGTG GTGCGATAAC ACGACCAAAA GAAATTACGA AACGTTTTGT TCAAATTATG 11580
 GAAGATTAAA TGATAACGAT AAAAAAACGA GATGACCATC ATTAATTAAA GGCACCTAAT 11640
 10 TATCTTAGGT GGCTGAATGA ATGTAATGGG TTCATCTCGT TTTGTTTGT TATGATAGTG 11700
 ATTTTATTTT CAACTTTATC CAAAAATAAG TAAAGCGACG GGGATGGTGA TTAATAGCGA 11760
 CAACGCCACG CGTAAAAACC AAATGATGAT GAGTTTCCAG ACAGGTATTT TAATTTTCAGT 11820
 15 TGCTAGTATA CATGGCACTA ATGCTGAGAA AAAGATAATG GCTGATACGC TTACTACACC 11880
 GACGACAAAT TTAGTACTCA TTGCAGCTTT AGTTACTAAC AAAGATGGTA GAAACATCTC 11940
 20 TACAATAGAA AaCTGACGC TTTTGCTAGT AAAGCCTGAT CAGCAATTGG GAAAATATAA 12000
 ATAAATGGAT AGAAGATATA GCCAAGCCAA TCAATGAATG GTGTATAGTT CGCTACAATC 12060
 AGTCCTAAAA AACCAATCGA TAATATAGAA GGTAAATAC CAACAGTCAT TTCTAAACCG 12120
 25 TCTTTCAAAT TGTCCCAAAC GTTCTTCACG AGAGATGGTG TTAATGCATT TTGTTTCATC 12180
 GCCTCTGCAT ATGCAGTTTT CAGTCTGCTT CCTTCAATAG CAACTTCTTG TTCTCCTTCT 12240
 TGTCCGTTAT AATATTCTGT TGATTCATTG CTGATTGGCG GTAGCCATGC AGTAATTGCA 12300
 30 GTCACGACAA ATGTGATGAC TAAAGTTATC CAAAAGTATA AATTCCAATG CGGCATTAAT 12360
 CCTAAAGTTT TAGCAACGAT AATCATAAAA GTTGCTGAAA CTGTTGAAAA GCCAGTCGCA 12420
 ATAATCGTGG CTTCTCGTTT GTTGTACATC CCTTGCTTAT AGACACGATT AGTAATCAAT 12480
 35 AATCCTAAGG AATAACTGCC GACAAACGAA GCCACTGCAT CGACAGCGGA TTTTCCTGGT 12540
 GTTITAAAAA TAGGTCTCAT AATAGGCTCC ATATAAACAC CGACAAATTC TAATAAGCCA 12600
 TAGCCCACTA ATAAAGAAAG cGcAATTGCA CCTACTGGAA TTAAGATACT TAATGGCATC 12660
 40 ATTAATTTTT CAAACAAAAA CGGACCATAG TTAGCTTTAA ATAGTATTGA TGGACCGATT 12720
 TTAAATACAT ACATTATACC GATCATTGCA CCTGCAACTT TAAATAATGT AATGACCAAG 12780
 TTTGTGATTG AAGTCATAAA AGTACGTCTC ACTATTGGTA ACGCTGTACC AATTAAAATC 12840
 45 ATAATCAGTG CAACATAGGG CATAAGTGGA CCTATGATTG AGCGAATGGC TAGATGAACA 12900
 TGATCGACGA AAATAGTGTT GTTACCATTA ATCGTAAAAG GAATAAAGAA ACATAGTATG 12960
 50 CCCACTAAAC TATAGACAAA AAAACGCCAT GCACTTGGTT GTTGTGCATT AGAATGATAT 13020
 TGATTCATTA AAGCAACCCC TTTGTTTAAA TGAATACACA AAAGTGTATG ATGCATCTTC 13080

55

	ATAGTTTGAA TTATTTTCAT ACCAATACAA ATTAACATA TATATATAGA TTGAAACTAT	13200
	ATTACTTAAT AAAATATTTA TCTTAAATGT TGTGTGTTG ATTCAACACC ACAACTAAAA	13260
5	GTGTTTATAA ATTATTTGGA AATACACATA TTTGTAAATG ATTAGTATCG ATTTAATATC	13320
	GTATTATTAA ATTTTATT TAATTTGTAGT CTTAATCMAA AAATAATATA TGTCATGTTA	13380
	TATTGAAGGT GCAGTTGTTT TTCATTCTCA AGAGGGGGTC AAAAAAATAC TTTTGAGGTG	13440
10	ATTATATGTT AAGAGGACAA GAAGAAAGAA AGTATAGTAT TAGAAAGTAT TCAATAGGCG	13500
	TGGTGTCACT GTTAGCGGCT ACAATGTTTG TTGTGTCACT ACATGAAGCA CAAGCCTCGG	13560
15	AAAAACATC AACTAATGCA GCGGCACAAA AAGAAACACT AAATCAACCG GGAGAACAAG	13620
	GGAAATGCGAT AACGTCACAT CAAATGCAGT CAGGAAAGCA ATTAGACGAT ATGCATAAAG	13680
	AGAATGGTAA AAGTGAACA GTGACAGAAG GTAAAGATAC GCTTCAATCA TCGAAGCATC	13740
20	AATCAACACA AAATAGTAAA ACAATCAGAA CGCAAAATGA TAATCAAGTA AAGCAAGATT	13800
	CTGAACGACA AGGTTCTAAA CAGTCACACC AAAATAATGC GACTAATAAT ACTGAACGTC	13860
	AAAATGATCA GGTTCAAAAT ACCCATCATG CTGAACGTAA TGGATCACA TCGACAACGT	13920
25	CACAATCGAA TGATGTTGAT AAATCACAAC CATCCATTCC GGCACAAAAG GTAATACCCA	13980
	ATCATGATAA AGCAGCACCA ACTTCAACTA CACCCCGTC TAATGATAAA ACTGCACCTA	14040
	AATCAACAAA AGCACAAGAT GCAACCACGG ACAAACATCC AAATCAACAA GATACACATC	14100
30	AACCTGCGCA TCAAATCATA GATGCAAAGC AAGATGATAC TGTTCGCCAA AGTGAACAGA	14160
	AACCACAAGT TGGCGATT TAAGTAAACATA TCGATGGTCA AAATTCCTCCA GAGAAACCGA	14220
	CAGATAAAAA TACTGATAAT AAACAACATA TCAAAGATGC GCTTCAAGCG CCTAAAACAC	14280
35	GTTTCGACTAC AAATGCAGCA GCAGATGCTA AAAAGGTTTC ACCACTTAAA GCGAATCAAG	14340
	TACAACCACT TAACAAATAT CCAGTTGTTT TTGTACATGG ATTTTATAGGA TTAGTAGGCG	14400
	ATAATGCACC TGCTTTATAT CCAAATTATT GGGGTGGAAT TAAATTTAAA GTTATCGAAG	14460
40	AATTGAGAAA GCAAGGCTAT AATGTACATC AAGCAAGTGT AAGTGCATT GGTAGTAACT	14520
	ATGATCGCGC TGTAGAACTT TATTATTACA TTAAAGGTGG TCGCGTAGAT TATGGCGCAG	14580
45	CACATGCAGC TAAATACGGA CATGAGCGCT ATGGTAAGAC TTATAAAGGA ATCATGCCTA	14640
	ATTGGGAACC TGGTAAAAAG GTACATCTTG TAGGGCATAG TATGGGTGGT CAAACAATTC	14700
	GTTTAATGGA AGAGTTTTTA AGAAATGGTA ACAAAGAAGA AATTGCCTAT CATAAAGCGC	14760
50	ATGGTGGAGA AATATCACCA TTATTCACCTG GTGGTCATAA CAATATGGTT GCATCAATCA	14820
	CAACATTAGC AACACCACAT AATGGTTCAC AAGCAGCTGA TAAGTTTGGA AATACAGAAG	14880

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATTTAGGATT AACGCAATGG GGCTTTAAAC AATTACCAAA TGAGAGTTAC ATTGACTATA 15000
 TAAACCGCGT TAGTAAAAGC AAAATTTGGA CATCAGACGA CAATGCTGCC TATGATTTAA 15060
 CGTTAGATGG CTCTGCAAAA TTGAACAACA TGACAAGTAT GAATCCTAAT ATTACGTATA 15120
 CGACTTATAC AGGTGTATCA TCTCATACTG GTCCATTAGG TTATGAAAAT CCTGATTTAG 15180
 GTACATTTTT CTTAATGGCT ACAACGAGTA GAATTATTGG TCATGATGCA AGAGAAGAAT 15240
 GCGTAAAAA TGATGGTGTC GTACCAAGTGA TTTCTGTCATT ACATCCGTCC AATCAACCAT 15300
 TTGTTAATGT TACGAATGAT GAACCTGCCA CACGCAGAGG TATCTGGCAA GTTAAACCAA 15360
 TCATACAAGG ATGGGATCAT GTCGATTTTA TCGGTGTGGA CTTCTCTGGAT TTCAAACGTA 15420
 AAGGTGCAGA ACTTGCCAAC TTCTATACAG GTATTATAAA TGACTTGTG CGTGTGAAG 15480
 CGACTGAAAG TAAAGGAACA CAATTGAAAG CAAGTTAAAT TCATCTTCTG AATTTAATAT 15540
 GCTATGTAAT TCGTGCTGTT ATCATGGCAC ATCAGATATA AGTAGCATCA CAGTGTGAA 15600
 TTTAAAAATA GTAAAGTGAA ATAAAGCGCC TGTCTCATT GCGAAACTA AAGGGACAGG 15660
 CGTATCTGTT TATGAGCTTA ATAAATTGTA TGAATAATAT GGTGATCGA ATAAGTGT 15720
 ATCATGATGA TAAATTGAGT TTTTAAAT AATGATATAT TACATCATTG TTATAGCGTT 15780
 TAAGAAATCA ACAACTTTAC GATAAATAGT GATTGCTTCG TCATTAGGTC TACGATCAAA 15840
 ATCATGCTCG TTTTATTCA CGCGTTCAA TGTGTAATGT GGAACATGAT TCATGATATG 15900
 TTCGCTTCC TCAACGGGAA CATCATAATC GCCATTACAA TCGCAATGA AAACAGGTGG 15960
 AAGTGTTTTA AGTTCATCTG GTGCAATATT ATATTTTGAA TTAGTATAAT CAGCAATGTT 16020
 AATCATATTT ATCCATTTAC CTGTGCCACG TGCATAAAG TAGATTAAAA AACGTTGTGC 16080
 GATTGATCT TGAACAACCG GTGTTGGTGA AGTGAGTTGT GCAATCATTG TTTCTTTAC 16140
 GCTTTGAGCT ATTTTTCGT AATAACTATT AGTTGTTTAA AAAGGTTGAG TGTGATGCG 16200
 ACTATAACCA TAAAAATCAA TAACACCATC AATATCTCTG TCTCGTGCAA TTAATAGACT 16260
 TAAATATGCA CCTGATGATC TGCCAAAGGT AAAAATAGGG CAATTAGAAT ATTGTGATTG 16320
 AATCGCATCG AATGATGCgn AGnACATCCT CAATAATGCA ATCGAGACTT ACTTCTGGTA 16380
 ATAAACGATA ACTTAGTTGA ATTAAATCGT AATGTTCCGT AAGATATCGA TATACTGTGG 16440
 GGATAAATCG TTAGCTTTAC CGAACATTAA TCCACCACCG TGGATGTAGA CAATAGCGCC 16500
 TTTTGTGGT TGATTTTTTG CTTAATAAT TGTGTAAGGT AATGCAAATG CATCTTTAGT 16560
 AATTACTTTA TCTTTAATTT CAGTCACGAT TTAATAGGCT CCTTATTTT GATATTGATG 16620
 TCATTATAAC ACTGTCTTAA ATTTCCATGA AAAATAGTCT TAAGACGATG AGTCATGATA 16680

CATCATTTTA ACAATATCTT TAAAAGCAGC ATGTGGAATG GCTAAATCTT CTAAATCTGC 16800
 CATAGAAAAT TCAAGATTGA TATCATGTGG TCGCTGTTCA GCAAGTTTAT GCACAAAGTC 16860
 5 AGGTTCTGTG ACAAAGGCG AAGACATGCC GACCATATCT GCATGTTGTA AAGCATCTAA 16920
 AGCAGACTCT GGAGAATTAA TCCCGCCACT TGCAATTAAA GGGATACGAC CTGCTAAATG 16980
 10 TTCATAGACA ATTTGGTTAA CTGGTCGACC GAAATGATCA CCTGGTGTAC GAGACGTATT 17040
 TTGATAAATA TGTCGACCCC AGCTAGCGAT TGCTAAGTAT TGGATGTTTG AAACGTCCAT 17100
 GACCCAATTG ATTAATTGGT TGAACGTC AATGGTATAT CCTAAATCAC TGCCTCTGGT 17160
 15 TTCTTCTGGC GTTGCTCGAA ATCCTAAAAT AAAATTGTCA GGTGCTTCTT TATCAATCAC 17220
 TTCTTGTAAC GCACGCATAA CTCTAAACA TAATCTTGCA CGATTTTTTA ATGAGTCGGC 17280
 ACCGTAATGG TCTGTACGTT TATTCGAAAA AGTTGAGAAA AATGTTTGAA TCAGCAAACG 17340
 20 TTGTGCAATC GAAATTTCCA CACCATCAAA ACCTGCTTTA ATCGCGCGTA ATGTAGCATC 17400
 GCGATACTGC TGAATGATGC TATTGATTTT CTCATGAGAC ATGGCGATAA CATCGTGTTT 17460
 AATCGGTGAA TGCAATGTCA TAGGGCTTGG TCCATACACC TTTCCAAAAT TAAAATGGC 17520
 25 TTGATTGAA AAACGACCAG CATGCGCTAG CTGGATAATA GCGAGGCTAC CATGTTGTTT 17580
 CATCGTAGAT GCCATGTTAG TTAATCCAGG GATACAAGCA TCATGATCAA TATTAAAGCC 17640
 ATATTCAAAC AATTGACCAT AAGGTTCAAT GTAAGCAGCG CCGGTGACTT GCATTCCAGC 17700
 30 TGAATTAGAG CGACGTGCAG CATAAGCCAA GTCTTCTTTT GTAATATAGC CTTCTTTTGT 17760
 TGATGTGTTT ACGGTCATTG GTGATAATAC AAAGCGATTG GAAATTTTGA TGCCATTAGG 17820
 TAAGTGGATT GATTGTAAAA GTGGTTTGTA TCGGTACATA CTATGATTCC TTTTCTATTC 17880
 35 AATATTGTTT TCAAAGTACC ATGGAAAGAA TGAATAATCA ATGATGAACA GTCTTGATAG 17940
 AATAGAATTG GTACATGGAA AGTATTTTTA AAATTAACT AATGAATGGC ATTTGTAGGT 18000
 40 CTGAAAATAT GAATATGAAA AAGAAAAATA AAGGCGAAAA GATATAAAAG TTAATTGAAA 18060
 AACGTTATCA TATACGTGGG TATATGAAGA GGAATGGTA TTAAGAACGC TAAAATGTTA 18120
 TGTCGGTTTG ACATGACAGG ATAAGTTTGG AGATGACGGA TTGGTTAAAT TAAGCGTATT 18180
 45 AGACTATGCC TTAATAGATG AAGGTAAGGA TGCACAAAAG GCATTGCAAG ATTCAGTGAC 18240
 ACTTGCAAAA TTAGCAGATC GACTTGCTT TAAGCGAATT TGGTTTACGG AACATCATAA 18300
 TGTACCAGCG TTTGCGTGTA GTAGTCCAGA ACTTTTGATG ATGCATACAT TGGCGCAGAC 18360
 50 AAATCACATA CGAGTTGGCT CTGGTGGTGT GATGCTGCCG CACTATCGAC CTTATAAAAT 18420
 TGCTGAGCAT TTTAGAATGA TGGCAGCGTT ATATCCAAAT CGTATTGATT TAGGTATTGG 18480

55

	TAGTTACGAT GAATCGATTT CGTTATTACG TGATTATCTT ACAATAAAGG ATAAACCAAG	18600
	TGCGCATACG TTAGGTGTCC AACCACACAT TGATCATTTT CCAGAAATGT GGTATTAAAG	18660
5	TAGTAGCGCA ACATCTGCCA AAATAGCTGC CGAAGTAGGT ATAGGGCTTT CTGTTGGAAC	18720
	ATTTTTGCTA CCAGATATAA ATGCGATACA TACAGCGAAG GATAACATTG ATATTTACAA	18780
	AAAACATTTT CAAGCATCAA CGATTAAAAT GGACGCAAAG GTGATGGCAT CTGTATTTGT	18840
10	CATTGTAGCT GATAACGAAG CGGAAGTAGC AGCATTACAA CATGCCTTAG ATGTTTGGTT	18900
	ATTAGGTAAA TTACAATTG CAGAATTTGA AGATTTTCCT TCAGTAGACA CAGCACAAAA	18960
	GTATAAGCTT AATGATCGAG ACAAAGAGAT GATTCAAGCA CATCAAGCAC GCATCATTGC	19020
15	AGGTACACAA GAAAAGGTTA AAGCACAATT AGATGATTTT ATTGCTACGT TTGAAGTTGA	19080
	TGAGGTGTTA GTAGCACC GC TTATTCCAGG TATTGAACAG CGTTGTAAAA CATTAAAATT	19140
20	ACTCGCGGAA ATTTATTTGT AGCATTTTAA ATAGAAGAGA AAGGATGAAG ATAAGATGAA	19200
	AAAGTTAGCC AATTATTTAT GGGTAGAAAA AGTAGGAGAT TTGTATGTGT TTAGTATGAC	19260
	ACCTGAATTG CAAGATGATA TTGGGACAGT AGGTTATGTT GAATTCGTAA GTCCAGATGA	19320
25	AGTTAAAGTG GATGATGAAA TTGTGAGTAT CGAAGCATCG AAAACGGTCA TTGATGTGCA	19380
	AACGCCATTG TCAGGAACGA TTATTGAGCG AAATACAAAA GCGGAAGAAG AACCGACAAT	19440
	TTTAACTCT GAAAAACCAG AAGAAAATTG GTTGTTCAAA TTGGATGATG TCGATAAAGA	19500
30	AGCATTCCTA GCATTACCGG AGGCTTAAAT GGAAACGTTA AAATCAAATA AAGCGAGACT	19560
	TGAATATTTA ATCAATGATA TGCAATGAGA GAGAAATGAC AATGACGTAT TGGTAATGCC	19620
	ATCTTCATTT GAAGATTTGT GGGGAATTATA TCGAGGCTTA GCAAATGTCA GACCGGCATT	19680
35	ACCTGTAAGT GATGAATATT TAGCTGTACA AGATGCTATG TTAAGTGATT TGAATCGTCA	19740
	ACATGTTACG GATTTGAAGG ATTTGAAGCC GATAAAAGGT GACAATATCT TTGTTTGGCA	19800
	AGGTGATATC ACGACGTTAA AAATCGATGC TATTGTTAAT GCTGCAAATA GTCGTTTTCT	19860
40	AGGATGTATG CAAGCTAATC ATGACTGCAT TGATAATATT ATTCATACAA AAGCGGGTGT	19920
	TCAAGTTCGA CTTGATTGTG CAGAGATCAT TCGACAACAA GGGCGCAATG AAGGTGTAGG	19980
45	TAAAGCCAAA ATAACACGTG GATATAATTT GCCAGCAAAG TATATAATTC ATACGGTTGG	20040
	TCCGCAAATA CGTCGATTGC CTGTTTCAA GATGAATCAG GACTTGTTAG CTAAATGTGA	20100
	TCTTAGCTGT CTAAATTGG CTGATCAACA TAGTTTAAAT CATGTCGCTT TTTGCTGTAT	20160
50	ATCTACAGGT GTATTTGCTT TTCCTCAAGA TGAAGCAGCA GAAATTGCTG TTCGAACAGT	20220
	AGAAAGCTAT CTCAAAGAAA CAAATTCAAC ATTGAAAGTC GTGTTCAATG TATTTACAGA	20280

55

	CAATGTCTCT GTTAATGGAT GACAAGACAA AGCAGGCTGA AGTATTGCGT ACTGCGATTG	20400
	ATGAAGCAGA TGCGATAGTG ATTGGAATTG GTGCAGGCAT GTCTGCATCT GACGGATTTA	20460
5	CATATGTAGG AGAGCGTTTT ACGGAAAATT TCCCAGATTT TATTGAAAAA TATCGCTTCT	20520
	TTGATATGTT GCAAGCGAGT TTACATCCTT ATGGCAGTTG GCAAGAGTAT TGGGCATTG	20580
	AGAGTCGTTT TATTACATTA AACTATTTAG ATCAACCTGT AGGTCAGTCT TACCTCGCTT	20640
10	TAAAATCCTT GGTGGAAGGT AAACAGTACC ACATTATAAC TACGAATGCA GATAATGCTT	20700
	TCGATGTAGC TGATTATGAT ATGACTCATG TATTTTCATAT ACAAGGGGAG TATATACTGC	20760
	AACAGTGTAG CTCAGCATTG TCATGCTCAA ACGTATCGCA ATGATGATTT AATTCGTAAA	20820
15	ATGGTTGTTG CGCAACAAGA TATGCTTATA CCTTGGGAGA TGATTCCAAG ATGTCCAAAA	20880
	TGTGATGCCC CAATGGAAGT GAATAAACGT AAAGCGGAAG TTGGGATGGT TGAAGATGCT	20940
20	GAATTCATG CGCAACTACA TCGTTATAAT GCTTTTCTAG AGCAACATCA AGATGATAAA	21000
	GTGTTGTATT TGGAAATTGG AATTGGTTAT ACTACACCAC AATTTGTGAA GCATCCTTTT	21060
	CAGCGTATGA CACGTAAAAA TGAAAATGCC CTTTATATGA CGATGAATAA AAAGGCATAT	21120
25	CGCATTCCGA ATTCAATTCA AGAACGTACC ATACATTTAA CTGAGGATAT CTCAACATTG	21180
	ATTACAGCAG CACTCCGGAA CGACAGCACA ACGAAAAATA ACAACATTGG AGAGACAGAA	21240
	GATGTACTTA ATAGAACCGA TTAGAAATGG AGAATATATT ACTGATGGTG CGATTGCACT	21300
30	CGCTATGCAA GTTTATGTTA ACCAGCATAT CTTTTTAGAT GAAGATATTT TATTCCTTA	21360
	TTATTGTGAT CCAAAAGTGG AAATTGGACG TTTTCAAAT ACTGCTATAG AAGTGAATCA	21420
	AGATTATATA GATAAACACA GTATTCAAGT AGTTCGCCGA GATACTGGTG GTGGCGCTGT	21480
35	GTATGTTGAT AAAGGTGCCG TTAATATGTG TTGTATTTTA GAACAAGACA CTTCAATTTA	21540
	TGGTGATTTT CAACGATTTT ATCAACCAGC TATAAAGGCG TTGCATACAT TAGGTGCAAC	21600
	AGATGTGGTA CAAAGCGGTA GAAATGATTT AACATTGAAT GGTAAAAAAG TGTCAGGCGC	21660
40	CGCAATGACA TTAATGAATA ATCGTATTTA TGGCGGTTAT TCGCTATTAC TTGATGTTAA	21720
	TTATGAAGCA ATGGATAAAG TGTAAAGCC TAATCGCAA AAGATTGCAT CGAAAGGGAT	21780
45	TAAATCTGTG CGCGCACGTG TTGGTCATCT TAGAGAAGCA CTGGATGAAA AGTATCGTGA	21840
	TATAACCATT GAAGAATTTA AAAATTTAAT GGTGACGCAG ATTTTGGGAA TCGATGACAT	21900
	TAAAGAGGCG AAACGATATG AATTAACGGA TGCAGATTGG GAAGCGATTG ATGAATTAGC	21960
50	TGATAAAAAG TATAAAAATT GGGATTGGAA TTATGGCAAG TCACCCAAAT ATGAATACAA	22020
	TCGAAGTGAA AGATTATCTT CAGGTACGGT AGACATAACA ATTTCTGTTG AACAAAATCG	22080

55

AGAAGCATT A CAAGGAACAA AAATGACAAG AGAAGATTTA ACGCATCAGT TAAAGCAATT 22200
 AGACATCGTT TATTATTTTG GCAATGTTAC GGTAGAAGCA TTAGTGGATA TGATTTTAAG 22260
 5 TTAATATTGT TATTTTATGT ATGCTGAATC AITGGAAGTG TTTGCTTGCT CTTGAAAAGG 22320
 TGACAATAGT GTTTGGTGAA GGTGGAACAT ATGAGTGGA AATTATGCCT TTAACATTC 22380
 AAAGTATGAT ATATATATGG TTTTGTTC TAAATGATTG GGTATTTGAA AATAGATGAG 22440
 10 TTTAATATTT TAAGGAATAT AATGATGTTT ACTTTTATAA TTCATATAGA ATATTAAGCA 22500
 ATATAAGTCT GTTGATATAT ACAAATATA ATGACTGCTA TAATGAGTAA TCAATAGACA 22560
 CAAAGAGGAG ATTATGTGAT GAATAATAAA GTATTAGTAA CCGGTGGTAC AGGGTTTGTT 22620
 15 GGCATGCGAA TTATTTACAG ATTATTAGAA CAAGGTTATG ACGTACAAAC GACGATACGT 22680
 GATTTAAGTA AAGCTGATAA AGTAATTAAC ACAATGCAAG ACAATGGCAT TTCCACAGAG 22740
 20 CGATTAAATGT TTGTCGAAGC GGATTATCA CAAGATGAAC ATTGGGATGA AGCAATGAAA 22800
 GATTGCAAGT ATGCTTGAG TGTAGCATCT CCGGTGTTTT TCGGTAAAAC AGACGATGCA 22860
 GAAGTGATGG CGAaCTGcAA TTGAAGGTAT ACAACGTATT TTAAGAGCTG CAGAACATGC 22920
 25 GGGTGTTAAA CGTGTGGTAA TGA CTGCAAA CTTTGGTGCA GTTGgTTTTA GTAATAAAGA 22980
 TAAAAATTCA ATCACAAATG AAAGTCATTG GACAAATGAA GATGAACCAG GCTTATCAGT 23040
 ATATGAAAAA TCAAAATTGT TAGCTGAAAA GGCAGCGTGG GATTTTGTG AGAATGAAAA 23100
 30 TACAACAGTA GAATTGCCA CAATCAATCC AGTTGCAATT TTTGGGCCAT CATTAGATGC 23160
 ACACGTTTCA GGAAGCTTTC ATTTATTAGA AAATTTATTG AATGGTTCAA TGAAACGTGT 23220
 ACCGCAAATT CCGTTAAATG TTGTTGATGT GAGAGACGTa GCTGAACTGC ACATTTTGGC 23280
 35 AATGACAAAT GAACAAGCTA ATGGCAAGCG ATTTATTGCG ACGGCTGATG GACmAATTwa 23340
 tTTGTTGGGA ATTGcCAAAt TAATTAAAGA AAAGGGCCTG GAAATAGCTC CAAAAGTTCC 23400
 40 TACTAAAAAA TTACCCAGCT TTATTTTGAG CnAnGnGCC 23439

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4522 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

CCCTTTGAGA GTATATCATC TAGTCAAATT ATGCCTGTCA TTAGAGCGAC TAGCTTTGAT 60

	TATTATGCAG TCGATTTAGG GAAATCATAT CGTCTAATTG ACGAAAGCAT GTTAGAGGAT	180
	TTGAAGTTAA CTGAACAACA AATAAGAGAA ATGTCTCTGT TTAATGTTAG AAAATTGTCA	240
5	AATTCATATA CGACTGATGA AGTAAAAGGT AATATTTTTT ATTTTATTAA CTCAAATGAC	300
	GGGTATGATG CAAGTAGGAT ACTAAATACT GCATTTTTTAA ATGAAATTGA GGCACAATGT	360
	CAAGGCGAAA TGCTCGTAGC AGTGCCACAC CAAGATGTGT TAATTATTGC AGATATACGC	420
10	AATAAACAG GATATGATGT GATGGCACAT TTAACAATGG AATTTTTCAC TAAAGGTCTA	480
	GTTCCAATTA CATCATTATC CTTTGGATAT AAACAGGGTC ATCTTGAACC GATATTTATT	540
	TTAGGTAAAA ATAATAAACA AAAAAGAGAT CCAAACGTGA TTCAGCGTTT AGAAGCAAAT	600
15	CGTCGTAAAT TTAATAAAGA TAAATAGAAA TAATTGGATA AGGAGTTTGT TCATAATGAA	660
	TTTATTTTAC AATCCTAAAT ATGTAGGAGA TGTCGCATT TACAAATTG AACCAGTTGA	720
	AGGTGAATTA AACTACAATA AAAAAGGTAA TGTTGTTGAA ATTACTAATG AAGGTAATGT	780
20	TGTAGGTTAT AATATTTTTG AAATTTCAAA AGATATAACA ATTGAAGAAA AAGGTCATAT	840
	TAAATTAACT GATGAACTTG TAAATGTATT CCAAAGCGT ATTTCAGAAG CTGGTTTTGA	900
25	TTATAAATTA AATGCTGATC TATCACCGAA ATTTGTAGTT GGCTACGTTG AACTAAAGA	960
	CAAACATCCT GATGCAGATA AATTAAGTGT ACTAAATGTA AACGTTGGAA ATGACACATT	1020
	ACAAATTGTA TGTGGCGCGC CTAACGTTGA AGCTGGACAG AAAGTTGTTG TTGCTAAAGT	1080
30	AGGTGCAGTG ATGCCTAGCG GTATGGTAAT TAAAGATGCT GAATTACGTG GTGTTGCCTC	1140
	AAGCGGTATG ATTTGTTCAA TGAAAGAATT GAATTTACCT AATGCACCTG AAGAAAAAGG	1200
	TATTATGGTA TTAATGACA GCTATGAAAT TGGACAAGCA TTtTTTGAAT AATTAAGGAA	1260
35	GGTAGTGAAA ATATGAGCTG GTTTGATAAA TTATTCGGCG AAGATAATGA TTCAAATGAT	1320
	GACTTGATTG ATAGAAAGAA AAAAAGACGT CAAGAATCAC AAAATATAGA TtACGATCAT	1380
	GACTCATTAC TGCCTCAAAA TAATGATATT TATAGTCGTC CGAGGGGAAA ATTCCGTTTT	1440
40	CCTATGAGCG TAGCTTATGA AAATGAAAAT GTTGAACAAT CTGCAGATAC TATTTAGAT	1500
	GAAAAAGAAC AATACCATCG AGACTATCGC AAACAAAGCC ACGATTCTCG TTCACAAAAA	1560
45	CGACATCGCC GTAGAAGAAA TCAAACAAC TGAAGAACAAA ATTATAGTGA ACAACGTGGG	1620
	AATTCTAAAA TATCACAGCA AAGTATAAAA TATAAAGATC ATTCACATTA CCATACGAAT	1680
	AAGCCAGGTA CATATGTTTC TGCAATTAAT GGTATTGAGA AGGAAACGCA CAAGCCAAAA	1740
50	ACACATAATA TGTATTCTAA TAATACAAAT CATCGTGCTA AAGATTCAAC TCCAGATTAT	1800
	CACAAAGAAA GTTTCAGAC TTCAGAGGTA CCGTCAGCTA TTTTGGCAC AATGAAACCT	1860

55

	AAACAAAAT ATGATAAATA TGTAGCTAAG ACGCAAACGT CTCAAAATAA ACAATTAGAA	1980
	CAAGAAAAAC AAAATGATAG TGTGTGCAAA CAAGGAACTG CATCTAAATC ATCTGATGAA	2040
5	AATGTATCAT CAACAACAAA ATCAATGCCT AATTATTCAA AAGTTGATAA TACTATCAAA	2100
	ATTGAAAATA TTTATGCTTC ACAAATTGTT GAAGAAATTA GACGTGAACG AGAACGTAAA	2160
	GTGCTTCAAA AGCGTCGATT TAAAAAGCG TTGCAACAAA AGCGTGAAGA ACATAAAAAC	2220
10	GAAGAGCAAG ATGCAATACA ACGTGCAATT GATGAAATGT ATGCTAAACA AGCGGAACgC	2280
	TATGTTGGTG ATAGTTCATT AAATGATGAT AGTGACTTAA CAGATAATAG TACAGATGCT	2340
	AGTCAGCTTC ATACAAATGG CATAGAGAAT GAAACTGTAT CAAATGATGA AAATAAACAA	2400
15	GCGTCAATAC AAAATGAAGA CACTAATGAC ACTCATGTAG ATGAAAGTCC ATACAATTAT	2460
	GAGGAAGTTA GTTTGAaTCA AGTATCGACA ACAAACAAT TGTCAGATGA TGAAGTTACG	2520
20	GTTTCGAATG TAACGTCTCA ACATCAATCA GCACTACAAC ATAACGTGA AGTAAATGAT	2580
	AAAGATGAAC TAAAAATCA ATCCAGATTA ATTGCTGATT CAGAAGAAGA TGGAGCAACG	2640
	aATAAGAAG AATATTCAGk AAGTCAAATC GATGATGCAG AATTTTATGA ATTAAATGAT	2700
25	ACAGAAGTAG ATGAGGATAC TACTTCAAAT ATCGAAGATA ATACCAATAG AAACGCGTCT	2760
	GAAATGCATG TAGACGCTCC TAAAACGCAA GAGTACGCAG TAACTGAATC TCAAGTAAAT	2820
	AATATCGATA AAACGGTTGA TAATGAAATT GAATTAGCAC CGCGTCATAA AAAAGATGAC	2880
30	CAAACAACT TAAGTGTCAA CTCATTGAAA ACGAATGATG TGAATGATAA TCATGTTGTG	2940
	GAAGATTCAA GCATGAATGA AATAGAAAAG AATAACGCAG AAATTACAGA AAATGTGCAA	3000
	AACGAAGCAG CTGAAAGTGA ACAAATGTC GAAGAGAAAA CTATTGAAAA CGTAAATCCA	3060
35	AAGAAACAGA CTGAAAAGGT TTCAACTTTA AGTAAAAGAC CATTTAATGT TGTCATGACG	3120
	CCATCTGATA AAAAGCGTAT GATGGATCGT AAAAAGCATT CAAAAGTCAA TGTGCCTGAA	3180
	TTAAAGCCTG TACAAAGTAA GCAAGCTGTG AGTGAAAGAA TGCCTGCGAG TCAAGCCACA	3240
40	CCATCATCAA GATCTGATTC ACAAGAGTCA AATACAAATG CATATAAAC AAATAATATG	3300
	ACATCAAACA ATGTTGaGAA CAATCAACTT ATTGGTCATG CAGAAACAGA AAATGATTAT	3360
	CAAAATGCAC AACAATATTC AGAGCAGAAA CCTTCTGTTG aTTCAACTCA AACGGAAATA	3420
45	TTTGAAGAAA GTCAAGATGA TAATCAATTG GAAAATGAGC AAGTTGATCA ATCAACTTCG	3480
	TCTTCAGTTT CAGAAGTAAG CGACATAACT GAAGAAAGCG AAGAAACAAC ACATCCAAAC	3540
50	AATACTAGTG GACAACAAGA TAATGATGAT CAACAAAAAG ATTTACAGTC ATCATTTTCA	3600
	AATAAAAATG AAGATACAGC TAATGAAAT AGACCTCGGA CGAACCAACA AGATGTTGCA	3660

55

CCAAGTGTTC CATTACTAGA AGAACCACAA GTTATTGAGT CGGACGAGGA CTGGATTACA 3780
 GATAAAAAGA AAGAACTGAA TGACGCATTA TTTTACTTTA ATGTACCTGC AGAAGTACAA 3840
 5 GATGTAAGTG AAGGTCCAAG TGTTACAAGA TTTGAATTAT CAGTTGAAA AGGTGTTAAA 3900
 GTTCCAAGAA TTACGGCATT ACAAGATGAC ATTAAAATGG CATTGGCAGC GAAAGATATT 3960
 CGTATAGAAG CGCCTATTCC AGGAACTAGT CGTGTGTGTA TTGAAGTTCC GAACCAAAAT 4020
 10 CCAACGACAG TCAACTTACG TTCTATTATT GAATCTCCaA GTTTTAAAAA TGCTGAATCT 4080
 AAATTAACAG TTGCGATGGG GTATAGAATT AATAATGAAC CATTACTTAT GGATATTGCT 4140
 AAAACGCCAC ACGCACTAAT TGCAGGTGCA ACTGGATCAG GGAAATCAGT TTGTATCAAT 4200
 15 AGTATTTTGA TGTCTTTACT ATATAAAAAT CATCCTGAGG AATTAAGATT ATTACTTATC 4260
 GATCCAAAAA TGGTTGAATT AGCTCCTTAT AATGGTTTGC CACATTTAGT TGCACCGGTA 4320
 ATTACAGATG TCAAAGCAGC TACACAGAGT TTAAATGGG CCGTAGAAGA AATGGAACGA 4380
 20 CGTTATAAGT TATTTGCACA TTACCCATGT ACGTantATA ACAGCATTTA ACnAAAAAGC 4440
 CCCATATGAT GAAAGAATGn CAAAAATTGT CATTGTaATT GATGAGTTGG CTGATTTAAT 4500
 25 GATGATGGTC CGCAAGAAGT TG 4522

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 751 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

35 TCAAGTTTAC GGATACGTAT ATATTTTGCA TGACATTTAG TGCAATAATA TTCATAATTT 60
 GCCCGTTGTT GATAGCTTTC AATGCTGTTA CAAAATCTAG GCGCTCCAAC CTGTTGGCTC 120
 40 AATCGTTTAA AATCTTGATC TTTATGTTGA TAACCTTTAC CAGCAATATG CAAGTGATAA 180
 TGACACAATT CGTGCAGTAT AATTTTACA ACAGCATCTT CTCCATAATG CTCATATTGT 240
 TTTGGATTAA TTTCAATATC ATGGGACTTT AAAAGATAAC GTCCGCCTGT TGTACGTAAC 300
 45 CTTTTATTAA AATATGCACA ATGTCGAAAC GTACGTCCAA ATTTTCTTC CGAAAGATTC 360
 TCAACCATTG GCTGAAGTTT GTCATTATTC ATGTGGATCA ATCATCGTTA ATGATACTTT 420
 GTCTTTATTT TTGTCAATAC TGTAATCCA AACGTCAACG ATATCACCAA CACTGACAAT 480
 50 ATCCATTGGA TTTTITACGA ACTTCTTAGA AAGTTTCGAA ACATGGACAA GTCCATCTTG 540

55

TTTCATTCTT TCTTGTAAT CTTCAATTGA TAGCACATCG GATTTAAGGA TTGGTGTTTC 660
 AAATCTGTC CTTGGATCTC GATTAGGTGC GTTCAAGGAT TTAATAATAT CCTCTAATGT 720
 5 AGGTACACCG ACTTGTAATT CAATCGCCAG T 751

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1076 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

TCTCCAGCTT TAACTTGATC TGGCACTTTA ACAATTGTCT GATCCATACA TACGCGACCA 60
 20 ATAATCTCGC ATTGATGACC ATTTACATTT ACAAAGCTAC CTTGCATTAT GCGTAAATGG 120
 CCATCTGCAT ATCCAATAGG TAACAATGCT ATTGTAGTTG GGTCACTAGC TGTATAAGTT 180
 GCACCATAAC TTACAGACTC ACCCGCTTGT AGCGTCTTTG TTTGAACTAC ATTAGCAATT 240
 25 AATTGCACAC TTGGTTTAAG GTGTACTTTA ACTTTTGTCT GTACATACTC TGATGGATAA 300
 TATCCATAAA GGGAAATTCC TGGTCTTATT GCATTACAGA ATTGGCAATC CATTAAATAGA 360
 GAGCCTGCTG AGTTCTGACA ATGTATATAT TCAGGTTTAA TTGCTTCATT GACCATATCT 420
 30 TTAAACGTT GATATTGTTT AGTTGTCATA TCTCCTGGTT CGTCAGCACA GGCAAAGTGT 480
 GTAAACACGC CTTCAAATAC AAGTTGCTCA TATTGTTGAA TGATTTCAAT CACTTCTTGA 540
 TACGTTTTAG TATCTTTAAT ACCTAAACGT CCCATTCTCTG TATCTAATTT AATGTGCAAC 600
 35 CATAACTTTT TCTCTTGCTC ACCAGAAATG TTTTAAATG CTTCTTTCAA CCACTGTTTA 660
 GACGGAACCG TTAAGGCAAC TCGGTGTTGT ATCGCTTTAT CAATATCTTT AGCTGGTAAC 720
 ACACCTAAGA CTAAAATTTT AGCAGTAATC CCATGCATTC TAAGTTCTAT CGCTTCATCT 780
 40 AACGTTGCTA CAGCAAAAAA TGTGGCGCCA TTTTCCATTA AATGACGTGC TACTTTAACA 840
 CTACCTAGTC CATAGGCATT GGCTTTAACG ACAGCCATCA CTGTTTTATT TGGATGCAAT 900
 GTACTGAATA CTTTGAAATT TGATGCAACA GCGTTTAAAT CTACATTCAT ATACGCAGAT 960
 45 CTATAATATT TATCCGACAT ATTACTTCCT CCTGTAATTC CCACACGTTT TAAACTAGA 1020
 TCTTAATTAT CATTGTATAA CAAATTTAAA ATGCTGACTT TTCTAAAACA ACTTGG 1076

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2930 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

	TGACCACAAT GCCCAATACA ACCATCCCAT GGTAAAGCCA AGAGATGAGT CAATAAAGCG	60
	TGTTGAATAA GAGCTGAATG AACCTGATAC TGGATAAAAT GTTGCCAACCT CTCCAATTGA	120
10	TGACATTAAAG AAATATAGCA TGACACCAAT AACAAGATAA GCGAGTATAG CGCCTCCAGG	180
	ACCAGCTTGA GAAATGATAT TACCAGTAGC TACAAATAGA CCAGTCCCAA TTGCACCACC	240
15	TATAGCAATC ATGGAAATGT GTCTTGAGTT AAGACTACGG TTCATTTTAT TATCTTCCAT	300
	ATTTAGTCTC CCATCTATTT AAATATACCC ATTATTGTAA GCTTTTTAAG TGTACTATTC	360
	AATACTATT TAGTACTGTA AAGCGAAAAA ATTAAATTT TCTGATTTTT TAATCATCTT	420
20	GAGCATGTTT AATTGTAATT TTGATGGGGT TAAATTATAA TATGTATTAA ATTATAATTA	480
	TnATAAATTG TGGAGGGaTG ACTATGTCAC AACAAGACAA AAAGTTAACT GGTGTTTTTG	540
	GGCATCCAGT ATCAGACCGA GAAAATAGTA TGACAGCAGG GCCTAGGGGA CCTCTTTTAA	600
25	TGCAAGATAT TTACTTTTTA GAGCAAATGT CTCAATTGTA TAGAGAAGTA ATACCAGAAC	660
	GTCGAATGCA TGCCAAAGGT TCTGGTGCAT TTGGGACATT TACTGTAACT AAAGATATAA	720
	CAAAATATAC GAATGCTAAA AaATTCTCTG AAATAGGTAA GCAAACCGAA ATGTTTGCCC	780
30	GTTTCTCTAC TGTAGCAGGA GAACGTGGTG CTGCTGATGC GGAcGTGACA TTCGAGGATT	840
	TGCGTTAAAG TTCTACACTG AAGAAGGGAA CTGGGaTTTA GTAGGGAATA ACACACCaGT	900
	ATTCTTCTTT AGAGATCCAA AGTTATTGT TAGTTTAAAT CGTGCGGTGA AACGAGATCC	960
35	TAGAACAAT ATGAGAGATG CACAAAATAA CTGGGATTTT TGGaCGGGTt TCCAGAAGCA	1020
	TTGCACCAAG TAACGATCTT AATGTCAGAT AGAGGGATTC CTAAAGATTT ACGTCATATG	1080
	CATGGGTTTG GTTCTCACAC ATACTCTATG TATAATGATT CTGGTGAACG TGTTTGGGTT	1140
40	AAATTCCATT TTAGAACGCA ACAAGGTATT GAAAACCTAA CTGATGAAGA AGCTGCTGAA	1200
	ATTATAGCTA CAGATCGTGA TTCATCTCAA CGCGATTTAT TCGAAGCCAT TGAAAAAGGT	1260
45	GATTATCCAA AATGGACAAT GTATATTCAA GTAATGACTG AGGAACAAGC TAAAAACCAT	1320
	AAAGATAATC CATTTGATTT AACAAAAGTA TGGTATCACG ATGAGTATCC TCTAATTGAA	1380
	GTTGGAGAGT TTGAATTAAA TAGAAATCCA GATAATTACT TTATGGATGT TGAACAAGCT	1440
50	GCGTTTGCAC CAACTAATAT TATTCCAGGA TTAGATTTTT CTCCAGACAA AATGCTGCAA	1500
	GGGCGTTTAT TCTCATATGG CGATGCGCAA AGATATCGAT TAGGAGTTAA TCATTGGCAG	1560

55

GGTCAAATGC GCGTAGTTGA CAATAACCAA GGTGGAGGAA CACATTATTA TCCAAATAAC 1680
CATGGTAAAT TTGATTCTCA ACCTGAATAT AAAAAGCCAC CATTCCCAAC TGATGGATAC 1740
5 GGCTATGAAT ATAATCAACG TCAAGATGAT GATAATTATT TTGAACAACC AGGTAAATTG 1800
TTTAGATTAC AATCAGAGGA CGCTAAAGAA AGAATTTTTA CAAATACAGC AAATGCAATG 1860
GAAGGCGTAA CGGATGATGT TAAACGACGT CATATTCGTC ATTGTTACAA AGCTGACCCA 1920
10 GAATATGGTA AAGGTGTTGC AAAAGCATTG GGTATTGATA TAAATTCTAT TGATCTTGAA 1980
ACTGAAAATG ATGAAACATA CGAAAACTTT GAAAAATAAA TTTGATATGT AGTTTCTATA 2040
TTGCGTAGTT GAGCAGTTTA TGATATCATA ATAAATCGTA AAGATTCCTA ACAAGAGAGG 2100
15 GTGTTTAAACG TCGCGTAAA CGTAACATTA GCATGCACAG AATGTGGCGA TCGTAACTAT 2160
ATCACTACTA AAAATAAACG TAATAATCCT GAGCGTATTG AAATGAAAAA ATATTGCCCA 2220
AGATTAAACA AATATACGTT ACATCGTGAA ACTAAGTAAT TCTTATCATT CAAATACGAC 2280
20 GATTTGAAAA TAAAGCGGGC TTACCTATTA TATTGGGGAG CTCGCTTTTT TATGAAATTT 2340
TTGTGAAGAG TGATTAATGG ATTGAGTTTC ATCGGTAGAA CAATATATGA TTATATTAGT 2400
TGTTACTTTA TTAAaTTTG AGAATATTTA TAGAAGGAAA TAGATTACTG ATTTTATAAA 2460
25 GTCACTTTGT TAGCGAATGC TTGAAAGAGT ATTTAATATA GTAGAATTTA AAATTTCAAA 2520
GCGGAATTTA ATAAGTACGA AGTAGTTCTG GGTATGTTTT ATAAATGTTC GATAATACAC 2580
TTTAATCTTA AATATGATGG TTTAGAAAAT GATTTAACAA AGAAATGAaA CTTTACTGTT 2640
30 GAATTATGTG AGGATTGTGT TATTATATAA ATCGTAATAA TTACGATTTG ATAAAAAGTG 2700
AGGTAACATAT ATATGGCTAA GAAATCTAAA ATAGCAAAAG AGAGAAAAAG AGAAGAGTTA 2760
35 GTAAATAAAT ATTACGAATT ACGTAAAGAG TTAAAAGCAA AAGGTGATTA CGAAGCGTTA 2820
AGAAAATTAC CAAGAGATTC ATCACCTACA CGTTTAACTA GAAGATGTAA AGTAACTGGA 2880
AGACCTAGAG GTGTATTACG TAAATTTGAA ATGTCTCGTA TTGCGTTTAG 2930

40 (2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3606 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

50 CTTCTTGCCA TGGCTCTCTT TATTTAAAAA TGCTTCCAAC TTGTCCATTT GATTGTTTCT 60

55

	TTATAAAAAA	CTAATTTTAC	AAATGCTTTT	GCGTTCTTAC	AAAAAATGCA	TTTGACTATT	180
	ATTATAATAA	GCGTATAATT	GTCGCATATT	ATTTTTTGTA	TTTTTGCCAA	TAACGAAGGA	240
5	GTATTTATGA	ATAAAGACAA	GCAATTGCAC	AACGACAAAA	TCAATCTATC	CCAATTAGTC	300
	TTATTAGGGT	TAGGCTCTTT	AATAGGATCT	GTTTGGCTAT	TTGGTGCGTG	GGAAGCATCA	360
	TCAATAGCTG	GACCAGCAGC	AATCATATCA	TGGGTTCTTG	GATTCCTAGT	CATTGGAACC	420
10	ATTGCCTATA	ACTACATTGA	AATCGGCACA	ATGTTTCCTC	AATCAGGTGG	CATGAGTAAC	480
	TATGCCCCAGT	ATACACATGG	CTCATTATTA	GGCTTTATTG	CTGCTTGGGC	GAATTGGGGT	540
	TCTTTGGTGA	CAATAATACC	TATCGAAGCT	GTGTCAGCTG	TTCAATATAT	GAGTTCTTGG	600
15	CCGTGGCATT	GGGCGAAACC	AATGAGATAT	TTAATGGAAA	ATGGCTCTAT	TAGCACATAC	660
	GGATTGCTAG	CTGTATATCT	CATCATTGTT	ATTTTTTCAT	TATTAAACTA	TTGGTCCGTA	720
	AAACTTTTAA	CATCATTTAC	GAGTTTAATT	TCTGTATTTA	AATTAGGCGT	ACCCATGTTA	780
20	ACCATCATCA	TGTTGATGCT	ATCAGGATTC	GACACTTCAA	ATTACGGCCA	TTCGGCAAGC	840
	ACATTTATGC	CTTACGGAAG	TGCACCGATT	TTTGCTGCAA	CAACAGCATC	AGGGATTATT	900
	TTTTCATTC	ATTCAATCCA	GACAATTATT	AATATGGGTT	CAGAAATTAA	AAATCCTGAA	960
25	AAAAATATCG	CAAGAGGCAT	CGCTATCTCA	CTGTCAATCA	GTGCAGTGTT	GTACATCATT	1020
	TTACAAAGTA	CGTTTATCAC	TTCTATGCCT	CAATCAATGT	TACAACATAG	TGGATGGAAT	1080
	GGCATCAACT	TCAATTCACC	ATTTGCTGAT	TTAGCTATCT	TATTAGGAAT	TAATTGGCTC	1140
30	GCAATTTTAC	TATACATTGA	AGCTTTTGTA	TCACCATTCG	GTACTGGCGT	GTCATTTGTC	1200
	GCCGTTACAG	GTCGAGTTTT	ACGAGCAATG	GAGAAAAATG	GACATATCCC	TAAATTTCTT	1260
	GGGAAGATGA	ATGAAAAATA	TCATATCCCA	CGTGTAGCAA	TCATCTTTAA	TGCCATCATT	1320
35	AGTATGATTA	TGGTTACATT	ATTTAGAGAT	TGGGGTACGC	TAGCAGCAGT	TATTTCTACT	1380
	GCAACTTTAG	TAGCCTATTT	AACTGGCCCA	ACGACAGTGA	TTGCATTAAG	AAAAATGGGA	1440
40	CCAACAATGA	CTCGTCCATT	TAGAGCAAAA	ATTTTAAAAG	TAATGGCACC	ATTATCATT	1500
	GTATTAGCTT	CATTAGCTAT	ATATTGGGCA	ATGTGGCCAA	CAACGGCTGA	AGTTAATTTA	1560
	ATCATTATAC	TTGGATTACC	AATCTACTTC	TTCTATGAAT	ATCGTATGAA	TTGGCGTAAT	1620
45	ACAAAGAAAC	AAATTGGTGG	TAGCTTATGG	ATTATTGTAT	ATTAAATCGT	GCTATCAATA	1680
	CTGTCATTTA	TAGGAAGCAA	AGAATTTAAA	GGCTTAAATA	TGATTCACTA	TCCATTTGAC	1740
	TTTATCGTTA	TTATTATTGT	GGCACTTATC	TTCTATTACA	TCGGTACAAC	GAGTTCATTT	1800
50	GAAAGCGTCT	ATTTCCGTCT	CGCAACACGA	ATCAATACGA	AGATGCGTGA	GTCACTAAAT	1860

55

EP 0 786 519 A2

	CACACACATT AACCAACCAT TGATTTC AAC ATCTTGGTTG GTTTTTTATT TTGAAAATCG	1980
	GTATAAATA ACTAACATAA CAAGATGATG ATCAGGCTGG GACATAAATC AATGTTCTAT	2040
5	GCTCTACGAA gTTATATTGG CAGTAGTTGA CTGAACGAAA ATGCGCTTGT AACAAAGCTTT	2100
	TTTCGATTCT AGTCAGGGGC CCCAACACAG AGAATTTCTGA AAAGAAATTC TACAGGCAAT	2160
	GCAAGTTGGG GTGGGACGAC GATAAAGAAA TACTTTTTCT ATAGAAATTA GTATyCTTA	2220
10	TGCATGAGTT TTA CTCTATGT ATTCATATTT TTAAGTACAC ATTAGCTGTG GCTAATGTAT	2280
	AAGAACCCT ACATAATAAA TCATTTGTGG CTCTTTATCA TTTCTGTCCC ACTCCCGTAG	2340
	AAGTACATCA TATAATGCTG AAAATGGTTT GAGTTAAAC AGATATCAAG CTCGTCTGAT	2400
15	TCAGTCACAA AATTGTCTTG TTATACTTGT CACCTATCAT CTATAGACCG TGGTATGATT	2460
	AAATTGGGGA TGATAAAGGA GGTTAATAAA TATGAAGATT AATACTACAG GTGGTCAAAT	2520
	TCATGGTATT ACACAAGATG GTTAGATAT CTTCTTAGGC ATTCCTTATG CAGAACCACC	2580
20	AGTTCATGAC AATCGCTTTA AACATTCTAC GTTAAACAA CAATGGTCAG AGCCAATTGA	2640
	TGCAACTGAA ATACAACCCA TCCCACCGCA ACCAGACAAC AAATTAGAAG ATTTTTTCTC	2700
	CTCACAATCT ACAACTTTTA CTGAACATGA AGACTGTTTA TATCTAAATA TTTGGAAACA	2760
25	ACATAATGAT CAGACGAAGA AACCTGTCAT CATTTATTTT TATGGTGGTA GTTTTGAAAA	2820
	TGGTCATGGT ACAGCCGAAC TCTATCAACC GGCACATTTA GTACAAAATA ACGACATTAT	2880
30	CGTTATTACA TGCAATTATC GTTTAGGCGC ATTAGGATAT TTAGACTGGT CATATTTTAA	2940
	TAAAGATTTT CATTCCAATA ATGGCCTTTC AGATCAAATC AATGTCATAA AATGGGTGCA	3000
	TCAATTTATT GAATCCTTCG GTGGCGACGC TAATAACATT ACTTTAATGG GTCAGTCTGC	3060
35	AGGCAGTATG AGCATTTTGA CTTTACTTAA AATACCTGAC ATTGAGCCAT ACTTCCATAA	3120
	AGTCGTTCTA CTAAGTGGCG CACTACGATT AGACACCCCTT GAGAGTGCAC GCAATAAAGC	3180
	ACAACATTTT CAAAAAATGA TGCTCGATTA TTTAGATACA GATGATGTTA CATCATTATC	3240
40	GACAAATGAT ATTCCTTATG TGATGGCGAA gCTAAACAA TCTCGAGGAC CTTCTAAAGG	3300
	GCTTGATTTA ATATATGCGC CTATTAAAC AGATTATATA CAAAATAATT ATCCAACAAC	3360
	GAAACCAATT TTTGCATGTT ATACAAAAGA TGAAGGCGAT ATTTATATTA CTAGTGAACA	3420
45	GAAAAAATTA TCGCCGCAAC GCTTTATCGA CATTATGGAA TTAAATGATA TTCCTTTAAA	3480
	ATACGAAGAT GTTCAGACGG CGAAGcAACA ATCTTTAGCG ATTACACATT GTTATTTCaA	3540
	ACAGCCGATG aAGCAATTTT TACmACmACT CAATATACmA GATTCCAACC GCACCAACTA	3600
50	TGGCTT	3606

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15109 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

10	GAAATTAAAA AAGCAATTGG AACAAGATGC AACAGTGTCA TTGTTTGATG AATTTGATAA	60
	AAAATTATAC ACTTACGGCG ATAACTGGGG TCGTGGTGA GAAGTATTAT ATCAAGCATT	120
	TGGTTTGAAA ATGCAACsAG AACAACAAAA GTTAACTGCA AAAGCAGGTT GGGCTGAAGT	180
15	GAAACAAGAA GAAATTGAAA AATATGCTGG TGATTACATT GTGAGTACAA GTGAAGGTAA	240
	ACCTACACCA GGATACGAAT CAACAAACAT GTGGaAGAAT TTGAAAGCTA CTAAAGAAGG	300
	ACATATTGTT AAAGTTGATG CTGGTACATA CTGGTACAAC GATCCTTATA CATTAGATTT	360
20	CATGCGTAAA GATTTAAAG AAAAAATTAAT TAAAGCTGCA AAATAATTCA GCTATATAAG	420
	TTAGTGAAAT GAGAGTCTGA AACATATCAA TCTTTTGATA TTGTATTAGG CTCITTATTTT	480
	TATAGCTAGA AAGTTAGATA TTTGTATTTT TTTAAATAAT AAGTGCCGTT GTTATCGTTC	540
25	AATTTAATTA ATGATAGATT AGTATTATTA TAGCTAAAGT AGTATACCTG AGAAAATAGC	600
	TCAATGTATC TCTTTATTAA TAAGTTATAT CATAATTATT TTAGTGCATA CTTTATGGAA	660
	GGGATATCAG GGAATGGCTT TCAATTAAAG AAGAGGTTTA AAAGGATTAC AACAGAATGT	720
30	TATGATTTTG TAGAAAGATA TATAACAACG TTTTATAAAA ACATAATATT GTTAATGGAA	780
	AATGAAATGT AAGGGGGATT TCGAGTGA CT AAGAAAGTTT ATTTTAACCA CGATGGTGGT	840
35	GTAGATGATT TAGTATCTCT ATTTTTATTA TTACAAATGG AAAACGTTCA ATTGATAGGG	900
	GTCAGTACAA TTGGTGCTGA TTGTTATTTA GAGCCATCTT TGAGCGCATC AGTAAAAATT	960
	ATTAATCGTT TTTCAAATGA AGATATTCAA GTTGCGCCAT CATATGAACG AGGAAAAAAT	1020
40	CCATTTCCCTA AAGAATGGCG TATGCATGCC TTTTTTATGG ACGCATTGCC AATTTTAAAT	1080
	GAGCCAGTCA AACATGTTGC TTCAAATGTG AGCGACAAAG AAGCCTTTGA AGACATTATT	1140
	CAAACTTTAA AGAGACAATC AGAAAAAGTA ACATTATTAT TTACAGGCCC GCTTACAGAT	1200
45	TTAGCAAAAG CACTACAAAA AGATTCATCT ATCGTTCAGT ATATAGAAAA ATTAGTTTGG	1260
	ATGGGTGGCA CCTTTTACC AAAAGGAAAT GTTGAAGAAC CTGAGCATGA TGGTTCTGCA	1320
	GAATGGAATG CATATTGGGA TCCAGAAGCG GTTAAAATTG TTTTGTATAG CGATATAGAG	1380
50	ATTGATATGG TTGCTTTAGA AAGTACGAAT CAAGTACCGC TAACGTTAGA TGTTAGACAA	1440

EP 0 786 519 A2

	GTACCACCAT TAACACACTT TATAACAAAT TCTACTTACT TTTTATGGGA TGTTTTAACG	1560
	ACTGCTTATA TTGGTAACAA GGACTTGGTT CATTCAATTG AGAAAAAGT CGATGTAATA	1620
5	AGTTATGGAC CAAGTCAAGG TAAGACATTT GAGTGTAAG ATGGGCGCAA AATTAATGTC	1680
	ATAAATCATG TAGATAACAA CGCATTTTTT GATTATATAA CTGCACTTGC TAAAAAGTA	1740
	AATTAACAGC TGTGTAGAAT AATTAAGGTT TTAATTTATA TAGAACAACT TATTGTAAAC	1800
10	TTTTCATTTT TTAAGTTTA CAATGGTGCT ATAATAATGG TCATGAAATA CGAAAGGAAG	1860
	TAAAAATGA CAACAAAACA GTTAGTATAT ACAGCTTTAA TGACAGCGAT TATCGCTATT	1920
	TTAGGATTGG TACCGGTAAT TCCACTACCA TTTTCTTCAG TACCAATTGT ACTTCAAAAC	1980
15	ATTGGTATTT TCTTAGCAGG TGCGATTTTA GGACGTAAAT ATGGCACATT AAGTGTATC	2040
	GTCTTTTTAT TATTAGTAGT TGCTGGCTTG CCATTGTTAT CAGGTGGTCG CGGTGGCATC	2100
	GGTGTAATCG CAGGTCCTTC AGCAGGGTTT TTAATTATAT ATCCAGTTGT AGCATTGATG	2160
20	ATTGGGGCGA TTCGAGATAG ATTCATCAAT GAAATTAATT TCTGGATTTT ATTCGTTGGT	2220
	ATTTTAGTTT TTGGTGTTAT AGCATTAGAT GTTATTGGTA CATTGATTAT GGGCATGATT	2280
	ATTAACATAC CATTACGAA AGCTATTTCA ATTTCAATTAG CTTATTTGCC TGGTGATATA	2340
25	TTAAAGCAA TTGTAGCAAG TTTGATTGGT ACAGCTTTAC TTAATCACTC GCAGTTTCGT	2400
	CAAAATTATG GAATAAAATA ATCATATTTA AGATAGTAAA GTAATTGAAT AAGTTGCTTT	2460
30	GAAATTTATA AAAGTGAAAG GAGTAGGTGT CAATGGCTAG TATAAGTATG TCAGATATAT	2520
	ATTGTAACGG CACTATATTT GAAATGACG ACGAGCAGTT GATTTATTTA ACGCCTTCTT	2580
	TTCCACAACG ATACACAAGT AACACATGGA TATATAAAAA GACGCCTACC CAAGAGCGAT	2640
35	GGCTGAAAGA CTTAGAACGT CAACATCAAT TACATACAAA TCAAGGTTCA AATCATTATG	2700
	CGTTTAGTTT CCCGGAAT GAACAACCTG ATAATCATTG GATGGCTATG TTTAAAGATA	2760
	TGAATTTTGA ACTAGGTATT ATGGAATTGT ATGCCATAGA AAGTGATGCG CTTGCCAATT	2820
40	TGCCGCGTAA CTCTGACGTT GAAATTGCCA TCGTTGACGA GTCGCATATA GATGCCTATT	2880
	TAAAAGTTGC ATATCAGTTT AGTTTGCCAT TTGGAAAAGA CTATGCAGAT GCACATGAAG	2940
	AAATGGTAAG GGAACATTAT CAAAAAGATG TGATTAAACG CTTAGTAGCT TATTTAAATA	3000
45	ATGAACCTAT TGGCGTTGTA GATGTCATTG AAAGTGAAAA TTACATTGAA TTAGATGGAT	3060
	TTGGTGATT AGAACATTT CGGCACCAAG GAATTGGATC TACAATTCAA TCGTTGATAG	3120
	GTGAATACGC CATATCAAAA AATCACAAAC CAATCATATT AGTTGCAGAT GGTGAAGATA	3180
50	CAGCAAAGA TATGTATGCA AAGCAAGGTT ATGTCTATCA ATCGTTTGT TATCAAATAT	3240

55

	TAAGCTGGTT TCGAGTAGAA ATCAACTTAC TGCTTTTTTAA ATTGTTTTGA GCTACTTATA	3360
	CTTATAAAAA TAGTGCCTTT AAATTGTTGA TTCATGTAGA ATATCGTTCA TTATGACACA	3420
5	CTATAATGAA TATGTTATTG TTCAGAATCA ATGATACGTT CTGGATGACT GTATATATTA	3480
	AAGCCACCAT TTCGAATAAA TCCAACGGCC GTAATATTTA GGTCAATAGC TAAGGTTACA	3540
	GCAAGCGTTG TCGGAGCTGA TTTAGATAAA ATGACGCCAA CACCAATTTT TCGCGCTTTA	3600
10	ATTAAAATTT CTGATGAAAT ACGTCCACTA AAAATTAATA CTTTATCTCG GACAGTAATA	3660
	TGTCGCTGAA TACAAAATCC ATATAATTTA TCTAGAGCGT TATGTCTACC AATGTCTTGT	3720
15	CGATGTACAA AAAATGTCAA ACCATCGCTT ATAGCAGCAT TATGTAAGCC ACCTGTTTCT	3780
	TGGTAAATAT GACTTGCACT TTGTAATCGA GTCATCATGT TAATAATTTG CATTGGAGTT	3840
	AAAGTGATTT TAGACATAGA TGTTTTAGCG ATAGCAGCAT CATTTTGAAA ATAAAACTCA	3900
20	CGACTCTTTC CGCAACAAGA TGCAATCATT CGTTTTGTGG AATATTGAAA GCGATCGCCT	3960
	AAATCTTTAT TAAGTTCAAC ATGGGCAAAA CCTTTACTAT CATCAATCAG TACAGATTTT	4020
	AATTCATCTC GCTTTAAAAT GGCACCTTCC GAAGCCAGAA ATCCAATGAC TAACTCCTCA	4080
25	AGGTTTGTG GACTGCATAT AACAGTCGCA AATTCCTCAC CATTCACCAT AATTGTAAGT	4140
	GGAAATTCTG TCACATATTG ATCTGTTGTA TTGAATAATT TTCCATCTTC ATATCTAACA	4200
	ATTGGTTGAC CTAAAGATAC ATCTTTGTTT ATTATCTAAC CCCTTTAATT AGCTTAAACT	4260
30	TTATTTTAAA GCAATTTGCT TAAAATTTTA ACATATTTGC TTAAGTTTGA AATTTGATTG	4320
	ATAAAAATTA ATAGCGAGCA ATCTGTTTGA TTTAAATTGA ATTCGAGAAT ATACATACTA	4380
35	GGGCATCAAT TAATAAATAT CAATCTTATG CAAATTGAC AATTGTTTGA ATCAATATAT	4440
	AAACAGGCAA CGGTTCTTTT CAAATATAAT AGTAAGTGTA TAATGAAAAT GTAAATATTA	4500
	TTAAAAATGG GGGTTCCTC AATGAAATTG AACGTTTAT TTGCTGTTGT GATTGCAATG	4560
40	CTTTTAGTAT TAGCTGGTTG CTCTAATTCT AACGATAATA ATGAAAGTAA AAAAGATGAC	4620
	GCAGACAATG GTAAGAAACA AGAGATTCAA GTTGCGAGCG CAGCAAGTTT AACAGATGTA	4680
	ACCAAGAAAT TAGCTTCAGA ATTTAAAAAA GAGCATAAAA ATGCTGATAT TAAATTTAAC	4740
45	TATGGTGGAT CAGGGGCATT AAGAAAACAA ATTGAATCAG GCGCACCTGT TGACGTATTT	4800
	ATGTCTGCAA ATACTAAAGA TGTAGATGCA TTTAAAGACA AGAATAAAGC GCATGATACA	4860
	TATAAATATG CGAAAAATAG TCTAGTATTA ATTGGTGATA AAGATTCAAA TTACACTTCA	4920
50	GTAAAGACT TAAAAGACAA TGATAAATTA GCATTAGGTG AAGTGAAAAC TGTACCAGCA	4980
	GGAAAATATG CGAAACAGTA TTTAGATAAC AATAACTTAT TTAAAGAAGT CGAAAGTAAA	5040
55		

CAAGGTTTTG TGTATAAAAC TGACTTATAT AAACAAAATA AAAAAATTGA TACTGTAAAA 5160
 GTAATTAAAG AAGTAGAACT TAAGAAGCCA ATCACATACG AAGCTGGTGC TACATCAGAT 5220
 5 AGTAAATTAG CAAAAGAGTG GATGGAATTC TTAAATCAG ATAAAGCTAA AGAAATACTA 5280
 AAAGAATACC ACTTTGCAGC ATAAGGAGTT GTAATCCATG CCTGACTTAA CACCTTTTTG 5340
 GATATCAATA CGAGTTGCTG TAATCAGTAC GATTATTGTA ACGGTTTTAG GTATTTTTAT 5400
 10 ATCTAAATGG TTGTATCGTC GTAAGGGTTC GTGGGTAAAG GTATTGAAA GTTTATTGAT 5460
 ATTACCTATT GTTTTGCCGC CAACGGTATT AGGTTTTATT CTATTAATCA TCTTCTCGCC 5520
 AAGAGGACCA ATCGGTCAAT TCTTTGCGAA TGTACTACAT TTACCTGTAG TGTTCACTTT 5580
 15 GACAGGTGCT GTGATAGCAT CTGTCATTGT TAGTTTTCCA CTAATGTATC AACATACTGT 5640
 GCAAGGCTTC AGAGGTATAG ACACGAAAAT GATTAATACA GCTAGAACGA TGGGAGCAAG 5700
 20 TGAAACGAAA ATTTTCCTCA AATTAATTTT ACCATTAGCT AAACGCTCTA TTTTAGCAGG 5760
 TATAATGATG AGTTTTGCTC GTGCATTAGG TGAGTTTGGT GCTACATTAA TGGTTGCAGG 5820
 ATATATTCCA AATAAACGA ATACACTACC TTTAGAAATA TACTTCTTAG TGGAACAAGG 5880
 25 TAGAGAAAAT GAAGCGTGGT TATGGGTATT AGTGCTAGTC GCATTCTCTA TTGTGGTTAT 5940
 ATCTACAATT AATTTATTGA ATAAAGATAA ATATAAGGAG GTCGACTAGA TGCTTAAAT 6000
 CAATGTGAAA TATCAATTAA AGAACACTTT AATTCGCATC AATATAGATG ATACTGAACC 6060
 30 AAAAAATTAT GCAGTTCGTG GTCCATCTGG CATTTGGTAAA ACTACTGTTT TAAATATGAT 6120
 TGCCGGATTA CGTAAAGCAG ATGAAGCTAT TATCGAAGTG AATGGGCAAT TACTTACTGA 6180
 TACGGCAAAA AACGTGAATG TTAAAATTCA ACAACGACGT ATTGGATATC TGTTTCAAGA 6240
 35 CTACCAATTG TTTCTAATA TGACGGTCTA TAAAAATATT ACTTTTATGG CTGAACCATC 6300
 TGAACACATC GATCAATTAA TTCAAACCTTT AAACATTGAT CATTTGATGA AACAATATCC 6360
 TATGACATTG TCAGGTGGAG AGGCACAACG TGTAGCACTT GCACGTGCAC TTAGCACAA 6420
 40 ACCAGATTTA ATTTTATTAG ATGAACCTTT TTCTAGTTTG GATGATACTA CAAAAGATGA 6480
 GAGTATTACA TTAGTTAAAC GTATTTTCAA CGAATGGCAA ATACCAATCA TATTTGTGAC 6540
 45 ACATTCAAAC TATGAAGCAG AACAAATGGC TCATGAAATT ATTACAATTG GGTAATCATT 6600
 TATTTGCCAT TAAAGAGTTT AGAACGTATT TAAATTTGTA GAAGTGAATG CTTCTATCAG 6660
 CATTTTAATG ATGTTTTAAA CTCTTTTTTA GGGGCAGTTT TTTTGAGAGA CATTGACGCG 6720
 50 CGTCATATAA TGAAAGTAAT GATAAAAAGA AAGGATAACT TAATGTGAGT CAAGAACGTT 6780
 ATTCAAGGCA AATTTTATTT AAACAAATAG GTGAAATAGG TCAAAGCAAA ATAAATCAAA 6840

55

	GAGCAGGCAT TGCCAAACTA ATCATTGTTG ATAGAGATTA TATTGAATTT AGTAATTTAC	6960
	AAAGACAAAC ATTGTTTACT GAAGAAGATG CTTTGAAAAT GATGCCTAAG GTGGTTGCAG	7020
5	CTAAAAAGCA TTTGCTAGCG TTACGTAGTG ATGTTGATAT TGATGATTAT ATTGCCCATG	7080
	TGGATTATTA TTTTTTGGAA ACACATGGAC AGGACGTTGA CGTTATTATT GATGCAACCG	7140
	ATAACTTTGA AACACGACAA CTGATTAAATG ATTTTGCATA TAAATATCGT ATACCTTGGA	7200
10	TTTATGGTGG TGTGTACAG AGTACATATA CAGAAGCTGC ATTTATACCT GGTAAACAC	7260
	CTTGCTTTAA CTGTTTGGTA CCACAATTGC CAGCATTAAA TTTAACATGT GATACAGTAG	7320
15	GGGTCATTCA ACCTGCCGTG ACGATGGCAA CAAGTTTACA ATTAAGAGAT GCGATGAAAG	7380
	TATTAACGGA ACAACCAATT GACACAAAAA TAACTTATGG CGATATTGG GAAGGTAGTC	7440
	ATTATTCATT TGGTTTCAGT AAAATGCAAC GTTCAGACTG TACAACCTGT GGAGATGTAC	7500
20	CAAGTTATCC GTATTTAAAC AAGAATGAAC AACGTTATGC AACATTGTGT GGTAGAGACA	7560
	CTGTACAGTA TGAAAATGCA TCAATTACAC ACGACATTCT TGTTCATTT TTAACAAC	7620
	ATCAGTTAAA TTATCGCAGT AATTCGTATA TGGTTATGTT TGAATTTAAA GGACACCGCA	7680
25	TTGTTGCTTT TAAAGGTGGA AGGTTTTTAA TACATGGCAT GACACGCACA TCAGATGCCA	7740
	CACATCTAAT GAATTTATTG TTTGGATAAA AAAAGATAAG ACAAAGGAG TGTAAATATTA	7800
	TGGGCGAACA TCAAAACGTT AAATTGAATC GTACAGTTAA AGCAGCCGTA CTAACGGTAT	7860
30	CAGATACTAG AGACTTTGAT ACAGATAAAG GTGGTCAATG CGTGCGCCAA CTATTACAAG	7920
	CAGATGACGT TGAAGTGAGT GACGCACATT ATACAATTGT GAAAGATGAA AAAGTAGCCA	7980
	TCACGACGCA GGTGAAGAAG TGGTTAGAAG AAGATATTGA TGTCAATCATT ACGACTGGTG	8040
35	GAACAGGTAT TGCACAACGT GATGTGACGA TTGAAGCAGT AAAACCACTT TTAACATAAG	8100
	AGATAGAAGG CTTTGGGGAA TTGTTTAGAT ATTTGAGTTA TGTGAAGAT GTTGGCACGC	8160
40	GTGCATTATT GTCTCGTGCT GTAGCAGGTA CAGTTAATAA TAAATTGATA TTTTCGATTC	8220
	CAGGATCAAC AGGCGCAGTT AAATTAGCAT TAGAAAAGCT CATTAAACCA GAATTAAATC	8280
	ATCTGATTCA TGAGCTTACA AAATAATTTA TTGATTGAT TGGCGTTGAA AATCTCCAGA	8340
45	TTTACCGCCA GACTTGCTTT CAAGGTAGGT TTCGCCAATA ATCATACCTT TATCAACTGC	8400
	TTTCGTCATG TCGTAAATGG TTAAAGCCGT TGCTGATGCA GCGGTTAAAG CTTCCATTTT	8460
	AACACCGGTT TTGCCAGTTG TAGAGACAGT TGTGTAATG TTTAAAGTAT AAAGGGGTGC	8520
50	ATTTGTTTCA TCCCAGCTGA AGTGAACATC TATGCCAGTC AATGGTAATG GATGGCACAT	8580
	CGGAATAAGT GTTGATGTAT TTTTGGCAGC CATAATACCA GCGATTTGAG CAGTGTTCAA	8640

55

	AATGCTTGAA TGAGCGACAG CAGTTCTTTT TGTAATTTGT TTGTCTGATA CATCGACCAT	8760
	TTTGGCGTGG CCTTGTTGAT TAATATGAGT AAACTCAGTC ATTTTACCCC TCCTAGTGCA	8820
5	TCTAGTATAT CATGAAAAAA TAAAAGTTTT GGAGATGATT TTTAATGGTA GTAGAAAAAA	8880
	GAAACCCAAT CCCAGTTAAA GAAGCAATTC AACGTATCGT TAATCAGCAG AGTTCAATGC	8940
10	CGGCAATTAC GGTAGCACTT GAAAAAAGTC TAAATCATAT CTTAGCAGAA GATATTGTAG	9000
	CTACTTATGA TATACCAAGG TTTGATAAAT CACCTTATGA TGGTTTTGCA ATTTCGAGTG	9060
	TTGATTCACA AGGGGCAAGT GGTGAGAATC GCATTGAGTT TAAAGTGATT GATCATATTG	9120
15	GTGCAGGTTT AGTTTCTGAT AAATTAGTTG GGGATCACGA AGCGGTGCGT ATTATGACTG	9180
	GAGCACAAAT ACCTAATGGC GCAGATGCTG TTGTTATGTT TGAACAAACG ATTGAACTAG	9240
	AAGATACATT TACAATTCGT AAACCATTTT CAAAAAATGA AAATATATCT TTAAAAGGTG	9300
20	AAGAAACAAA GACAGGCGAT GTTGTCTTAA AAAAAGGACA AGTAATTAAT CCAGGGGCTA	9360
	TCGCGGTCCT TGCAACATAT GGCTATGCAG AGGTTAAAGT TATTAAGCAA CCGAGTGTCTG	9420
	CTGTTATTGC AACAGGAAGC GAATTATTAG ATGTTAATGA TGTATTAGAA GATGGGAAAA	9480
25	TTGTAAGTCT TAATGGCCCA ATGATTCGTG CCTTAGCAGA AAAATTAGGT CTTGAAGTTG	9540
	GTATTTACAA AACACAAAAA GATGATTTAG ATAGTGGCAT CCAAGTCGTT AAAGAAGCTA	9600
	TGGAAAAACA TGATATCGTT ATTACAACGG GCGGAGTTTC TGTGGAGAT TTTGACTATT	9660
30	TACCTGAGAT TTATAAGGCT GTAAAGGCGG AAGTGTTATT TAATAAAGTA GCAATGCGTC	9720
	CTGGTAGCGT AACAAACGTT GCATTTGTAG ATGGaAAGTA TTTGTTTGGa TTATCTGGAA	9780
	ATCCATCAGC TTGTTTTACA GGATTGGAAC TATTTGTGAA nCCAGCTGTT AAACATATGT	9840
35	GTGGCGCACT AGAAGTCTTC CCGCAAATAA TTAAAGCAAC ATTAATGGAA GATTTTACCA	9900
	AGGCAAACCC ATTCACACGA TTTATACGTG CTAAAGCAAC GTTAACAAGT GCTGGAGCTA	9960
40	CTGTAGTACC TTCAGGATTC AATAAATCAG GTGCGGTTGT AGCGATTGCA CATGCTAACT	10020
	GTATGGTCAT GTTACCAGGA GGGTCACGTG GTTTTAAAGC GGGGCATACA GTAGATATTA	10080
	TATTGACTGA ATCTGACGCT GCTGAAGAGG AACTTCTTTT ATGATTTTAC AAATTGTAGG	10140
45	TTACAAAAAG TCTGGTAAGA CAACATTGAT GAGGCATATT GTCTCTTTCT TAAAGTCACA	10200
	TGGTTATACA GTTGCTACTA TTAAACATCA TGGGCATGGT AAGGAAGATA TTCAATTACA	10260
	GGATTCAGAC GTCGATCACA TGAAGCATTT TGAAGCGGGG GCAGATCAAA GTATTGTACA	10320
50	AGGTTTTCAA TATCAGCAAA CTGTAACACG TGTAGATAAT CAAAATCTTA CTCAAATTAT	10380
	TGAAAAATCT GTTACAATTG ACACCAATAT CGTATTAGTT GAAGGCTTTA AAAATGCTGA	10440

55

	GAATGTTTTGT TATAGCATT A ATGTAAGGGA GCATGAAGAT TTTACAGCAT TTGAGCAATG	10560
	GTTATTAAAT AAAATTAAAA ATGATTGTGA TACACAATTA ACATAGAGGA TTGAAATGAA	10620
5	TGAAACAATT TGAAATCGTG ACAGAACCGA TACAAACAGA ACAATATCGT GAATTCACCTA	10680
	TAAATGAATA TCAAGGTGCA GTAGTTGTTT TTACCGGTCA TGTTCCGGAA TGGACTAAAG	10740
	GCGTCAAAAC GGAATATTTA GAATATGAAG CGTATATTCC AATGGCTGAA AAGAAATTGG	10800
10	CACAAATTGG AGATGAAATA AATGAAAAAT GGCCTGGAAC GATAACGAGT ATTGTTTATA	10860
	GAATAGGGCC ATTACAAATT TCAGATATCG CTGTATTAAT TGCGGTTTCT TCACCGCATC	10920
	GTAAAGATGC CTATCGAGCA AATGAATATG CAATTGAGCG TATAAAAGAA ATTGTTCCGA	10980
15	TTTGAAAAA AGAAATTGGG GAAGATGGTT CAAAATGGCA AGGGCATCAA AAAGGGAATT	11040
	ATGAAGAAGC AAAGAGGGAG GAATAAGAGA GATGAAGGTA CTTTACTTCG CAGAAATTAA	11100
20	AGATATATTA CAAAAGCAC AGGAAGATAT TGTGCTTGAA CAAGCATTGA CTGTACAACA	11160
	ATTTGAAGAT TTATTGTTTG AACGTTATCC GCAATCAAT AATAAAAAGT TTCAAGTTGC	11220
	TGTAAATGAG GAATTTGTAC AAAAATCGGA TTTCATTCAA CCTAATGATA CTGTTGCATT	11280
25	AATTCCACCG GTTAGTGGAG GTTAAGGGAG CATGAAAGCA ATAATTCTTG CAGGTGGTCA	11340
	TTCAAGTGGC TTTGGTAAGC CCAAAGCTTT TGCGGAAGTG AACGGTGAGA CCTTTTATAG	11400
	TAGAGTAATT AAGACATTAG AATCAACAAA TATGTTCAAT GAAATTATTA TTAGTACAAA	11460
30	TGCGCAATTG GCAACGCAAT TTAATATATCC AAATGTTGTT ATAGATGATG AGAATCATAA	11520
	TGATAAAGGT CCATTAGCAG GAATTTATAC AATCATGAAG CAACATCCTG AAGAAGAATT	11580
	GTTTTTTGTC GTTTCTGTTG ATACACCAAT GATTACTGGT AAAGCTGTAA GCACGTTGTA	11640
35	TCAGTTTTTA GTTTCTCATC TTATTGAAAA TCATTAGAT GTCGCAGCTT TTAAAGAAGA	11700
	TGGAGGTTTT ATTCCAACAA TTGCATTTTA TAGTCCGAAT GCATTAGGCG CTATAACTAA	11760
	AGCACTACAT TCTGATAATT ACAGTTTTAA AAATGTATAT CATGAATTAT CAACGGATTA	11820
40	TTTGATGTA AGGGATGTAG ATGCGCCCTC ATATTGGTAC AAAAATATAA ATTATCAGCA	11880
	TGATTTGGAC GCTTTAATTC AAAAATTGTA AGCTGTTAGG AGGTCCACAA ATGGTAGAAC	11940
45	AAATAAAGA TAACTAGGA CGTCCCATCC GTGACTTACG GTTATCTGTG ACAGATCGGT	12000
	GTAACTTTAG GTGTGATTAT TGCATGCCTA AAGAGGTATT TGGAGATGAT TTCGTATTTT	12060
	TACCTAAAAA TGAACTTTAA ACGTTTGATG AAATGGCTAG AATCGCTAAG GSTATGCGAG	12120
50	AATTAGGTGT AAAAAAATA CGCATTACAG GTGGAGAACC ATTGATGCCA CGGGATTTAG	12180
	ATGTACTTAT AGCTAAATTA AATCAATCG ATGGTATTGA AGATATTGGT TTGACTACAA	12240

55

	ATGTCAGTTT GGATGCTATT GATGATACGC TATTTCAATC AATCAATAAT CGTAATATTA	12360
	AAGCGACTAC GATTTTAGAA CAAATTGATT ACGCGACGTC TATTGGTTTG AATGTAAAAG	12420
5	TAAATGTTGT TATACAAAAA GGTATTAACG ATGATCAAAT CATACCAATG CTTGAATATT	12480
	TTAAAGATAA ACATATAGAG ATTCGATTTA TAGAATTTAT GGATGTTGGT AATGATAATG	12540
	GATGGGATTT CAGTAAAGTT GTAAGTAAAG ATGAAATGCT TACAATGATA GAGCAGCACT	12600
10	TTGAAATCGA TCCTGTAGAA CCAAATATT TTGGGGAAGT AGCAAAATAT TATCGCCATA	12660
	AGGATAATGG TGTTCATTTT GTTTTGATTA CAAGTGTTC ACAATCATTT TGTCTACAT	12720
	GTACACGCGC AAGGCTGTCA TCAGATGGGA AGTTTTACGG ATGTTTTATTT GCAACTGTCTG	12780
15	ATGGATTTAA CGTTAAAGCG TTTATTCGTT CTGGCGTGAC CGACGAAGAA TTAAAAGAAC	12840
	AATTTAAAGC TTTATGGCAA ATAAGAGATG ATCGATATTC AGATGAGAGA ACTGCTCAAA	12900
20	CAGTTGCCAA TCGTCAACGT AAAAAGATAA ACATGAATTA TATTGGTGGT TAATGTGTAG	12960
	GGACCACTAC ATATTAAATC ATTAGAGATG TTTTAATATT TCTGTCTTAC TCCCTAAAAT	13020
	ACAATATTAT TTATTAAAGT AAAACGGTC ATATCTATGC CAGATTTAAT AGAAATGATC	13080
25	GTTTTTAAAG TTTTACAAAG TTGGCGGGGC CCCAACACAG AAGCTGACAG AAAGTCAGCT	13140
	TACAATAATG TGCAAGTTGG CGGGGCCCCA ACATAGAGAA TTTCAAAAAG AAATTCTACA	13200
	GACAAATGCAA GTTGGGGAAC GGGGCCCCAA CACAGAAGGT GACGAAAAGT CAGCATACAA	13260
30	TAATGTGCAA GTTGGCGGGG CCCCACATA GAGAATTTCA AAAGAAATTC TACAGACAAT	13320
	GCAAGTTGGG GATCAACGAA ATAAATTTTA TGAGAATATC ATTTCTATCC CACTCTTAAG	13380
	AATCACTACA TAATAAATCT TTAGTGGTTC TTTAACATTG ATGTCACACT CCATGCCATT	13440
35	GAGTTGTAAT ATATCTTTTT TAGGTATAAA TGTTGTGCAA TAAACAACAA GTTGTCCAAA	13500
	AGATATAAAT CTAAACAAGA TATAGCCAGC AATTTAATAT TTGTAATAGA TAAAATGCTA	13560
	AGTTTGATAT ATAATAAATT TAAGTAATTG TATAATAATA TGAATTACAA ACATCTAAGA	13620
40	AGAAACATAG GAGGCATCAT ATTATGAGTA ATAAAGTTCA ACGTTTTATA GAAGCAGAAA	13680
	GGGAGTTAAG TCAGTTAAAG CACTGGTTAA AAACAACACA TAAGATTTC AATTGAAGAAT	13740
	TTGTAGTCCT TTTTAAAGTG TATGAAGCTG AAAAGATTAG CGGTAAAGAA TTGAGGGATm	13800
45	CATTACATTT TGAAATGCTA TGGGATACAA GTAAAATCGA TGTGATTATC CGTAAAaTCT	13860
	ATAAAAAAGA GCTTATTTCT AAATTGCGTT CTGAAACGGA TGAAAGACAA GTATTCTATT	13920
50	TCTATAGTAC TTCTCAAAAG AAATTGTTAG ATAAATTAC TAAAGAAATA GAAGTGTTAA	13980
	GCGTTACAAA CTAAAACTT aaaaagcaTG CCAATCTCTA TTCATCATAA TTGCGTCTTG	14040

55

GTTCATGGCA TTTCTAGTTA CATGACGTCC ATGAATTAAG AAGTAAACAA GCATAGTAAT 14160
 GATTGCTAAA GCGGCCATAA AGCCGAAGAT TTCACTATAT GAAAACATAT GAGTAAATAA 14220
 5 CCCAAGGAAT GATGGACCGA AGCCGACACC TGCATCTAGA CCAACGTAAA AAGTAGATGT 14280
 CGCGATACCA TATTTAATCG GGGGTGAGAC TTTTATCGCA ATAGATTGCA TTGCAGATGA 14340
 TAAATTTCCA TACCCTAAAC CTAGGCAAGC ACCAGCAAGT AATATTAACC AGCTTTGATA 14400
 10 GCTTGAAATT AAGCATACAA ATGAAAGGAA AAGCATGATA AATGCTGGGT AGACAATAAT 14460
 ATTTTCATTT TTATCATCCA TCAATCTACC AGCAATAGGT CTAGTAATTA ACGATGCTAT 14520
 AGCATAGCAA ATAAAGAAAT AGCTTGCTGC AGTGACTAGG TGTCGCTCTA AAGCAAATGC 14580
 15 TTGTAAATAA GTTAGGATGG ACGCATAGGT AACGCCAATT AAAAGCATAA TTACAGCAAC 14640
 AGGAATGGCC TCTTTTGCAA TAAATTGATG AATACTAAAT CTTGGTTTAT CAATGACATT 14700
 AGTTTCAGTT TTGTTATTTG TTAATTCGAA ATCAACTTTT ATAAATAATG AGATAATGAG 14760
 20 TCCGAGTATG CCTAATATGA CACAAATAAT AAACAGTAAG TCAATTGCGT ATTTTGTAAT 14820
 AAGTAACATG CCTAGAAATG GGCCAATCGC TGTACCTAAT ACTAACTTA AGGAAAATAA 14880
 ACTGATGCCT TCACTTTTTT TATTAACAGG GGTACGTAT GCCGCAATAG TACCTGTTGC 14940
 25 AGTTGTCACA ACTGCAGTTG CGATACCGTT TATGAGACGT ACAAAGATTA AAAAAGCTAA 15000
 AGATCCATCA ATAAATAAAA GTAATTGCGT GATAATTAAA GCAATTAAAC CAATAAATAA 15060
 30 TAATCGTTTA GGTCCrATTT sATTTACAAA TTTACCTGTA GCAAATCGA 15109

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9072 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

GAGAGTCAAT GGCAAGAAGA ATATAAATAT TTGAGAGCGT TAATCTTTAA TGAAACAGAA 60
 TTAGAGGAAG CGTATAAATG GATGCATCCT TGTACACGT TGAATAATAA AAATGTAGTA 120
 45 CTTATCCATG GCTTCAAAAA TTATGTTGCA CTATTATTTT ATAAAGGTGC CATTTTGGAG 180
 GATAAATATC ATACACTCAT TCAACAGACT GAAAAGGTGC AAGCAGCTCG TCAGTTACGA 240
 TTTGAAAATT TAACAGAGAT TCAAGCACGT ACCGAAGAAA TTAAATATTA TCTAGCCGAA 300
 50 GCAATTAAAG CTGAAAAAGC TGGTAAAAAA GTTGAAATGA AGAAAACAGA GGAATATGTT 360

	AAATTAACGC CAGGCAGACA ACATCAATAT ATATATCATA TTGGACAAGC TAAACGCagT	480
	GgAACAAAGAC AAAAGCGTGT TGAAAAGTAT ATTAACCAAA TACTAGAAGG TAAAGGGATG	540
5	CATGATAAGT AATTAATGAG TAAAGCATAC CGGTTATACA ACAACATACA AGATGACACG	600
	AAACAACCAA TGGCTCATGC TGTGTTGTGT TTTTITAGGT GTGTCTGTCA TGGGCAACAC	660
	TTTGACGTTG GAATTCGGT ACAGGCTTGG GAGTAGAAAA TGTTAGCAAA AGGCAAGGGT	720
10	GTCTACAATG AATGATGAAG ATATTAAAAT ATAAGGATGA CTTTGTGAGT GCGGATGGG	780
	CGGTGTCCG TCTGTAACAA TGGATGCGTG TGCATTATTA CAAAAATTCG ACTTTTGTAA	840
	TAATATTTCA CATTTTCGAC ACTTTTTTGC TATAAAACAA CCAATTGAGC GATAATAAAT	900
15	TCGCTTTTAA AAAATATGAG TTATCTATTT AGTTGCCAAA GATAAAATAA TAATGTTTAA	960
	TAACATCATA TAGAGTATGT TAGTTTTAAA TGTCGAATAT ACGAATGTGc AAACAAAGTA	1020
20	ATCGGTAGAA ATTCAACATA CATAGCGCCG TTTACTGTTA AGTATTCACA TTACAGATGA	1080
	AAAAATATAA ATTCTACATA ATCAAGACCA TGATGTGTAC TTGTTTAACT TATGACTCTA	1140
	TTTGTTTAAC AATTGCGATA ATGGTCTTTT TATTTTATGC GTATCATTCTG TCATATTTTT	1200
25	TATGAGGAAG GAGAAATGAT TATGTTAAGT ATTAAGCATT TAACGAAAAT TTATTCTGGT	1260
	AATAAAAAGG CAGTAGATGA CATCTCTTTA GATATTCAAT CTGGGGAATT TATCGCATT	1320
	ATTGGAACCA GTGGAAGTGG CAAAACGACT GCTTTAAGAA TGATAAACCG TATGATTGAA	1380
30	GCGACAGAAG GACAAATTGA AATTGATGGT AAAGATGTTT GGAGTATGAA TCCTGTCGAA	1440
	TTGCGTAGAA ATATTGGCTA TGTATTCAA CAAATTGGCT TAATGCCTCA TATGACGATT	1500
	AAAGAGAATA TTGTGTTGGT ACCCAAATTG TTGAAATGGA CTAAAGAGGA AAAGGATAAA	1560
35	CGTGCAAAGG AATTAATTAA ACTTGTGGAT TTACCGGAGT CATTTTTAGA GCGTTATCCA	1620
	GCAQA ACT TAT CAGGTGGGCA ACAACAACGT ATCGGTGTTG TAAGAGCACT TGCGGCCGAA	1680
	CAAGATATTA TTTTAAATGGA TGAACCTTTT GGTGCATTGG ATCCTATTAC GAGAGATACG	1740
40	TTACAAGATT TAGTTAAAAC GTTACAACGA AAATTAGGCA AGACGTTTAT CTTTGTAAAC	1800
	CATGATATGG ATGAAGCGAT TAAATTAGCA GACAAAATTT GTATTATGTC AGAAGGTAAG	1860
	GTGGTGCAAT TTGATACGCC AGACAATATT TTAAGACATC CCGCAAATGA TTTTGTACGT	1920
45	GATTTTATAG GACAAAATAG ACTGATTCAA GACCGTCCCA ATGACAAGAC TGTAGAAGGT	1980
	GTAATGATTA AACCAATCAC GATACAAGCA GAAGCAACAC TGAATGACGC CGTTCATATT	2040
50	ATGAGACAAA AACGTGTTGA TACTATTTTT GTAGTAGATA GTAATAACCA TTTACTAGGT	2100
	TTCTTAGACA TTGAAGATAT AAATCAGGGT ATACGTGGAC ACAAAGTTT ACGAGACACC	2160

	ATTTTAAAAA GAAACGTTAG GAATGTACCT GTCGTAGATG ATCAACAGCG TTTAGTAGGA	2280
	CTGATTACGC GTGCCAATGT TGTGATATT GTATATGACA CGATTTGGGG CGATAGTGAG	2340
5	GATACAGTGC AAACAGAACA TGTGGGGGAA GACAcTGCCT CCTCAAAAGT GCATGAGCAA	2400
	CACACTACTA ATGTCAAAGT ACGTGACATA GGAGATGATA AATCATGATT GAGTTCCTAC	2460
	ATGAACATGG TGGACAGTTG ATGTCGAAAA CACTGGAACA TTTCTATATT TCTATAGTGG	2520
10	CATTATTACT TGCCATCATT GTTGACAGTAC CTATAGGCAT TTTATTATCA AAAACAAAGC	2580
	GAAGTGCCAA TATTGTATTA ACTGTGGCAG GTGTCTTACA AACTATTCCA AACTAGCTG	2640
	TACTTGCTAT TATGATACCG ATTTTGGTG TTGGTAAAC GCCTGCAATT GTAGCGCTAT	2700
15	TTATTTATGT ATTATTACCT ATTTTAAATA ACACGGTACT CGGTGTTCAA AATATTGATA	2760
	GCAACATTAA AGAAGCTGGA AAAAGTATGG GAATGACACA ATTTCAATTG ATGAAGGATG	2820
20	TTGAATTGCC GTTAGCATTG CCGCTTATCA TTGGTGGCAT TCGTTTGTC TCTGTGTATG	2880
	TAATTAGTTG GGCTACACTT GCAAGTTATG TAGGTGCGGG TGGATTAGGT GATTTCATTT	2940
	TCAATGGTTT AAATTTATAT GATCCACTGA TGATTGTAAC TGCAACGGTA CTCGTTACTG	3000
25	CACTAGCATT AGGTGTTGAT GCCTTATTAG CTTTAGTTGA AAAATGGGTA GTTCCCAAAG	3060
	GCTTAAAGT ATCTGGATAA TTAGGAGGCT AAGATAATGA AGAAAATTAA ATATATACTT	3120
	GTCGTGTTTG TCTTATCGCT TACCGTATTA TCTGGATGTA GTTTGCCCCG ACTAGGTAGT	3180
30	AAGAGCACGA AAAATGATGT CAAAATTACA GCATTATCAA CAAGCGAATC GCAAATTATT	3240
	TCACATATGT TACGGTTGTT AATAGAGCAT GATACACACG GTAAGATAAA GCCAACATTA	3300
	GTAAATAATT TAGGGTCAAG TACGATTCAA CATAATGCCT TAATTAATGG GGATGCTAAT	3360
35	ATATCAGGTG TTAGATATAA TGGCACAGAT TTAACGGGAG CTTTGAAGGA AGCACCAATT	3420
	AAAAATCCTA AGAAAGCAAT GATAGCAACA CAACAAGGAT TTAAAAAGAA ATTTGATCAA	3480
	ACGTTTTTTG ATTCGTATGG TTTTGCGAAT ACGTATGCAT TCATGGTAAC GAAGGAAACC	3540
40	GCTAAAAAAT ATCATTTAGA GACAGTTTCA GATTTAGCAA AGCATAGTAA AGATTTACGT	3600
	TTAGGTATGG ATAGTTCATG GATGAATCGT AAAGGCGATG GCTATGAAGG ATTTAAAAA	3660
	GAGTATGGTT TTGACTTTGG TACAGTGAGA CCAATGCAAA TAGGTCTAGT CTACGACGCA	3720
45	TTAAACTCAG AGAAGTTAGA CGTTGCATTA GGTTATTCTA CAGATGGTCG AATTGCGGCG	3780
	TATGATTTGA AAGTACTTAA AGATGATAAA CAATTTTCC CACCTTATGC TCGAGTGCT	3840
50	GTTGCAACAA ATGAATTATT ACGGCAACAC CCAGAACTTA AAACGACGAT TAATAAGTTG	3900
	ACAGGAAAGA TTTCGACTTC AGAGATGCAA CGCTTGAATT ATGAAGCGGA TGGTAAAGGT	3960

55

	AAAGGTGGTC ATAAGTAATG GAAGGTAATT TATTACAGCA ATTATTCAAT TATTATGTTA	4080
	CGAACTTTGG TTATCTATGG GATTTATTTT TCAAACACTT ATTAATGTCT GTCTATGGTG	4140
5	TGCTGTTTGC AgCTTTAATT GGTATTCCAT TGGGAATCTT GCTTGCaAGA TACACAAAAC	4200
	TTTCTGGATT TGTAATTACA ATTGCAAATA TAATTCAAAC AGTTCAGTC ATTGCAATGT	4260
	TAGCTATTTT AATGTTAGTC ATGGGCTTAG GTTCAGAAAC AGTAGTTTTA ACAGTGTTTT	4320
10	TATATGCGTT ACTTCCAATT ATAAAAACA CTTATACTGG TATAGCTAGT GTTGATGCGA	4380
	ATATTAAGGA TGCTGGCAAA GGTATGGGAA TGACACGCAA TCAAGTGCTA CGAATGATTG	4440
	AATTACCGTT ATCTGTTTCG GTTATTATCG GTGGCATTCTG TATTGCCTTG GTTGTTGCGA	4500
15	TAGGTGTTGT TGCCGTTGGA TCATTTATAG GAGCACCTAC GCTTGGTGAC ATTGTGATTC	4560
	GTGGTACAAA TGCGACGGAT GGCACAACGT TTATTTTAGC AGGTGCGATT CCGATTGCTA	4620
20	TCATTGCAAT CGTCATTGAT GTACTATTAA GATTTTTAGA AAAACGATTA GACCCAACAA	4680
	CACGACATCG TAAAAATCAA TCTAATCATC GGCCGCAAAG TATTAATATG TAATAGTAGA	4740
	AGATGTTTAT AATTTAGCGA TTTCGTTTCA TGATTTATAA AAAATGAGGC TACTCAAGGA	4800
25	GCTCAAATAA TCTTTGAGTA GCCTTTTTAT AGGTTGTGTT TGTATGCGTT TACACTAAAA	4860
	TAGCAATTAT TATCATGAAA GTTTTGGAT AAAAAGCGTT AATTATTGTA AAAATACTAA	4920
	AAAATGAGAT GTTTTATTTA TAATTTCTG CAAATTTATG ATATTGTTTC TTAATATATC	4980
30	ATATTAAAAA TTTGTTTTTC TTAAACATAG GAGGCTTATC TAATTCATGG ACACATCAAA	5040
	ACAATTTAGA GGTGACAACC GATTGCTTTT GGGTATCGTT TTAGGGGTTA TTACCTTTTG	5100
	GCTATTCGCG CAGTCACTTG TTAATCTTGT TGTCCCATTA CAATCAACAT ATAGTAGTGA	5160
35	CGTTGGAACG ATAAATATCG CTGTTAGCTT ATCTGCCTTA TTTGCTGGTT TGTTTATCGT	5220
	AGGTGCTGGT GATGTTGCTG ATAAATTTGG TCGCGTCAAA ATTACTTATG TAGGATTGAT	5280
	ATTAAATGTT GTAGGTTTCAT TACTCATCAT CATTACACCT TTGCCAGCAT TTTTAATTAT	5340
40	AGGTAGAATA ATTCAAGGTT TGTCTGCAGC ATGTATTATG CCATCAACAC TTGCTATTAT	5400
	TAACGAATAT TATATTGGTA CAAGAAGACA ACGTGCCTTA AGCTATTGGT CTATTGGTTC	5460
	TTGGGGTGGT AGTGGTATTT GTACGTTGTT TGGTGGCTTA ATGGCTACAT ATATAGGTTG	5520
45	GCGTTCAATA TTTGTTGTTT CAATTCTATT AACATTATTA GCAATGTACT TAATCAAACA	5580
	TGCACCTGAG ACTAAAGCAG AACCAATCAA AGGTATGAAA GCAGAAGCTA AAAAGTTTGA	5640
50	CGTTATTGGT TTAGTCATTT TAGTAGTGAC GATGTTAAGT TTAAATGTAA TCATCACACA	5700
	GACGTCTCAT TTTGGTTTAG TTTCACCGTT AATTCTAGGT TTAATTGTTG TGTTTATCTG	5760

	AATTTTAA	AATAGAGGAT	ACAGTGGTGC	AACTATTTCA	AACTTCTTAT	TAAATGGTGT	5880
	AGCAGGTGGT	GCACTTATCG	TTATTAACAC	GTATTATCAA	CAACAATTAG	GATTTAATTC	5940
5	TTCGCAAACG	GGTTATATTT	CATTAAACGTA	TTTAATAACA	GTGTTGTCAA	TGATTTCGTGT	6000
	AGGTGAAAAG	ATTTTATCTC	AACATGGTCC	GAAGCGCCCA	CTATTACTAG	GAAGTGGCTT	6060
	TACAGTGATT	GGGTTAATCT	TATTGTCGTT	AACATTTTTA	CCAGAAGTGT	GGTATATCAT	6120
10	ATCTAGTATA	GTTGGATATT	TATTGTTTGG	TACTGGTTTA	GGATTATATG	CTACACCATC	6180
	AACTGATACA	GCAGTTGCTA	GTGCGCCAGA	TGATAAGTCG	GGTGTGCTT	CAGGTGTGTA	6240
	TAAAATGGCG	TCATCATTAG	GAAATGCATT	TGGAGTAGCA	GTATCTGGTA	CGGTTTATAC	6300
15	TGTGTTAGCA	GCTAATTTAA	ATTTGAACCT	AGGTGGTTTC	ACAGGTATGA	TGTTTAATGC	6360
	CTTGCTAGCA	ATTGTTGCAT	TTTAGTCAT	TTTACTATTA	GTTCCCTAAA	ATCAAACGAA	6420
	TTTGTAAGAC	TGAAATGAAA	GCAAGTTATT	ATGTAGGGAT	TTTAAAGGAA	ATTTTGTGAA	6480
20	AGTAAGTTTA	TCATACACAC	TTAATGTTGC	GTATTGACGT	TTAATGTTAG	GTGTGTTCTT	6540
	TTATAGACGA	TAAAAGCTGT	GTGCATATTA	AGCGAATGAT	TTTCAAATTG	ACGCTAATAT	6600
25	GCGAAGTAG	TATTTTAA	ATGAACAACA	ACGATGAAGA	GGGGTTTATA	GGATGAAAAT	6660
	TGCAATTGCT	GGATCGGGTG	CATTAGGTAG	TGGCTTTGGT	GCCAACTAT	TTCAAGCAGG	6720
	ATATGATGTC	ACACTTATTG	ACGGATATAC	ATCTCATGTT	GAAGCGGTTA	AGCAACATGG	6780
30	ATTAAATATA	ACGATTAATG	GAGAGGCATT	CGAGTTAAAC	ATTCCGATGT	ATCATTTTAA	6840
	TGATCAACCG	GACGAAAGCA	TTACGATGT	TGTCTTTCTA	TTTCCAAAGT	CTATGCAATT	6900
	AAAAGAAGTG	ATGGAAGATA	TGAAGCCACA	TATTGATAAT	GAAACGATCG	TCGTATGTAC	6960
35	GATGAATGGT	CTGAAGCATG	AAGAAGTCAT	TGCGCAGTAT	GTTGCTCAAT	CACAAATTGT	7020
	CAGAGGTGTT	ACGACTTGGA	CGGCAGGTCT	TGAAAGCCCT	GGACACAGTC	ATTTACTTGG	7080
	TAGTGGACCA	GTTGAAATAG	GTGAACTAGT	GGATGAAGGT	AAAGAAAATG	TTATAAAAGT	7140
40	TGCTGATTTA	CTTAACGAAG	CGGAATTGAA	TGGTGTCAAT	AGTAAAGATT	TATACCAATC	7200
	GATTTGAAA	AAGATTTGTG	TTAATGGTAC	GGCAAATGCA	TTAAGCACAG	TGTTGGAGTG	7260
	TAATATGGCA	TCGCTGAATG	AAAGTAGTTA	TGCGAAGTGT	TTGATTTATA	AATTAACGCA	7320
45	AGAAATAGTG	CATGTAGCGA	CGATTGATAA	TGTTTCATTTA	AATGTTGATG	AAGTATTTGA	7380
	ATATTTAGTT	GATTTAAATG	AAaAGTTGG	TGCGCATTAT	CCATCCATGT	ATCAAGATTT	7440
50	AATTGTTAAT	AATAGAAAAA	CTGAAATTGA	TTATATTAAT	GGCGCAGTTG	CAACATTAGG	7500
	TAAACAACGT	CaTATTGAAG	CGCCAGTCAA	TCGCTTTATT	ACTGATTTAA	TTCATACTAA	7560

55

	CAATCACGTG ATATTACGGT CATTATTAAG ATTGAAATGT AATAAATAAA GAACAGCAGT	7680
	AAGGTACTTT CAAATTGAAA TGATCTTGGT GCTGTTTTTC TTGATTGATC TTCGTCATAA	7740
5	TTCAGATTTG TCATAGGCTA CGACATACTA TTAGTATTTA CTAGACAGTT TTTACGACGA	7800
	CACTTTGAAA AATTTTGAGG CAAATCATTT GGAAGTCTCA CGTGAATTTT GTAAACTCAT	7860
	CAAGCAAGTA ATTATATTAA AAAGACAAAT AGAGAAAAGG TGTTTATAAT GAGTAAAATT	7920
10	TTTGTAAGTG GTGCAACGGG CCTTATTGGC ATTAAATTAG TTCAAAGACT AAAAGAAGAG	7980
	GGGCATGAGG TTGCTGGTTT TACTACATCT GAGAATGGTC AACAAAAGCT AGCTGCTGTT	8040
	AATGTAAAAG CATATATTGG TGATATATTA AAAGCTGATA CTATTGATCA AGCGTTAGCA	8100
15	GATTTTAAAC CAGAAATCAT TATCAATCAA ATTACGGATT TAAAAAATGT TGATATGGCA	8160
	GCAAATACGA AAGTACGTAT TGAAGGTTCT AAAAACCTAA TTGATGCGGC GAAAAAGCAT	8220
20	GACGTTAAGA AAGTAATTGC CCAAAGTATT GCCTTTATGT ATGAACCTGG CGAAGGATTA	8280
	GCAAATGAGG AAACCTCACT TGATTTTAAAC TCAACTGGCG ATAGAAAAGT AACGGTTGAT	8340
	GGTGTGGTTG GTTTAGAAGA AGAAACGGCT CGTATGGATG AATACGTTGT TTTACGTTTT	8400
25	GGCTGGTTAT ATGGCCCAGG TACTTGGTAC GGAAAAGATG GCATGATTTA TAATCAATTT	8460
	ATGGATGGTC AAGTGACACT TTCAGATGGC GTAACATCAT TTGTGCATCT TGATGATGCA	8520
	GTTGAAACAT CTATTCAAGC TATTCATTTT GAAATGGTA TCTATAATGT AGCAGATGAT	8580
30	GCACCTGTTA AAGGTTCTGA ATTTGCAGAA TGGTATAAAG AACAACTTGG TGTTGAACCA	8640
	AATATTGATA TTCAACCTGC GCAACCATTT GAACGTGGCG TAAGCAATGA GAAGTTTAA	8700
	GCGCAAGGTG GTACTCTGAT TTATCAAACCT TGGAAAGATG GCATGAATCC AATTAAATA	8760
35	TAATTTATCC GTTTAATATA CAAAGAATAA AGACTTGGTC GAATCGTGGA TGATATATTA	8820
	TCAAACGCAC GGCTCGAACA AGTCTTTTTT ATTATGTCTT CGTTATCTTT GTATGAAGGA	8880
	ATAACAGAAT TACAATTAAT GTACTGAATA ATGCAATTAA TGTGTGATT AGTGCTAATT	8940
40	TAATTTCTAT TGGTAGCCAA GTCAGTACAA AAGACCAATT ATTGCTACCG AGAATGAGAT	9000
	ATGGTAATGC ATATAATATG AGCGCTAAAG CGATACATAT ACATAATGAT AACCAACTCA	9060
45	ATACAGCAAT CC	9072

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16826 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

	GTGGAACAGC TGTAACATA TCATTTCTTT CAACATTTAT TGGGAAAATG TTAGCTACAT	60
5	TTCTATATCC GATTAATAAT GTAGTACTTT CATATATnTC TGTAATGAA AGTGACAATA	120
	TAAAGAAGCA ATATTTGaaa ACTAATCTAA TTGCTATAGC TGCCCTATGT TTAGTCATGA	180
	TTATATGTTA TCCAATTACA ATAATTATTG TCTCTTTACT GTATAACATT GATTCAAGTT	240
10	TATATTCGAA GTTTATTATT TTAGGTAATA TAGGTGTTTT ATTCAATGCA GTGAGTATTA	300
	TGATCCAAAC TTAAATACA AAACACGCAT CAATAACATT ACAAGCGAAT TATATGACGC	360
	TTCACACGAT TACATTTATA TTCATAACTA TTTTAATGAC AATTGCGTTT GGTCTAAATG	420
15	GATTCTTTTG GACAACGCTG TTCAGCAACA TTATTAAGTA TGTGATTTTA AATATTATAG	480
	GTTTAAAGTC TAAATTCATT AATAAAAAGG ACGTCGATTA GATGAGTGAA AAAAAGATTT	540
20	TGATTTTATG TCAGTATTTT TATCCGGAAT ATGTATCTTC TCGACGTTA CCAACTCAAT	600
	TGGCGGAAGA TTTAATTGCG AATCACATTA ATGTCGATGT CATGTGTGGA TGGCCATATG	660
	AATATAGTAA TCATAACAG GTTCTTAAAA CCGAGATGCA TCGTGGTATT CGCATTGAC	720
25	GTCTCAAGTA TTCGAGGTTT AATAACAAAA GTAAGGTGG AAGGATCATC AATTCTTTA	780
	GTTTATTTTC AAAATTCGTG ATTAATATAC CTAAAATGTT GAAATATGAT CAGATTCTTG	840
	TTTACTCTAA TCCACCAATC TTGCCATTAA TACCAGACGT TTTACACAGA CTGCTTAAGA	900
30	AAAAATATTC TTTTGTGGTG TATGATATAG CACCTGATAA TCGATTAAAG ACAGGTGCAA	960
	CTCGTCCAGG TAGCATGATT GATAAGCTGA TCGGTTACAT TAATAGACAT GTCTACAAGA	1020
	ATGCTGAAAA TGTCATTGTC CTTGGTACGG AAATGAAAA CTACTTACTA AATCATCAAA	1080
35	TTTCTAAAAA TGCTGACAAAT ATCCATGTGA TTCCTAACTG GTATGACATG CGTCAATTAC	1140
	AAGATCAATCG TATCTATAAT GACACATTTA AAGCTTACCG TGAGCAATAC GACAAAATTT	1200
	TATTGTATAG CGGTAATATG GGCAGTTAC AGGATATGGA GACACTTATC TCATTTTAA	1260
40	AATTAAATAA GGATCAGTCT CAAACGTTAA CAATACTTTG TGGTCATGGT AAGAAATTTG	1320
	CAGATGTCAA AACGGCAATA GaAGACCATC GTATTGAAAA TGTTAAAATG TTTGAGTTTT	1380
45	TAACAGGTAC AGACTATGCT GACGTATTAA AAATTGCGGA TGTATGTATT GCATCGCTGA	1440
	TTAAAGAAGG CGTCGGTTTA GCGTGCCGA GCAAGAATTA TGGCTATCTT GCAGCTAAGA	1500
	AAGCGTTGGT ACTCATCATG GATAAGCAAT CTGATATCGT TCAACATGTT GAACAATATG	1560
50	ATGCGGGTAT CCAAATTGAT AATGGCGATG CACATGCCAT TTATAACTTC ATCAACACTC	1620
	ACTCGAGTAA GGAATTGCAC GAGATGGGTG AGCGCGCACA TCAACTGTTT AAAGATAAAT	1680

55

	AAGCGATTAT TCGATGTAGT GAGTTCAATA TATGGTTTAG TAGTTTTAAG TCCGATTCTG	1800
	TTAATTACAG CATTACTAAT TAAAATGGAA TCACCTGGAC CAGCCATTTT CAAACAAAAA	1860
5	AGACCGACGA TTAATAATGA ATTGTTTAAT ATTTATAAGT TTAGATCAAT GAAAATAGAC	1920
	ACACCTAATG TTGCAACTGA TTTAATGGAT TCAACATCGT ATATAACAAA GACAGGGAAG	1980
	GTCATTGTA AGACCTCTAT TGATGAATTG CCACAATTAT TGAATGTTTT AAAAGGAGAA	2040
10	ATGTCAATTG TAGGTCCTAG ACCAGCGCTT TATAATCAAT ACGAATTAAT CGAAAAACGT	2100
	ACAAAAGCGA ACGTGCATAC GATTAGACCA GGTGTGACAG GACTAGCTCA AGTGATGGGG	2160
	AGAGATGATA TCACTGATGA TCAAAAAGTA GCGTATGATC ATTATTACTT AACACATCAA	2220
15	TCTATGATGC TTGATATGTA TATCATATAT AAAACAATTA AAAATATCGT TACTTCAGAA	2280
	GGTGTGCATC ACTAATGAGA AAAAATATTT TAATTACAGG CGTACATGGA TATATCGGTA	2340
20	ATGCTTTAAA AGATAAGCTT ATTGAACAAG GACATCAAGT AGATCAAATT AATGTTAGGA	2400
	ATCAATTATG GAAGTCGACC TCGTTCAAAG ATTATGATGT TTTAATTCAT ACAGCAGCTT	2460
	TGGTTCACAA CAATTCACCT CAAGCAAGGC TATCTGATTA TATGCAAGTG AATATGTTGC	2520
25	TGACGAAACA ATTGGCACAA AAGGCTAAAG CTGAAGACGT TAAACAATTT ATTTTTATGA	2580
	GTACTATGGC AGTTTATGGA AAAGAAGGTC ATGTTGGTAA ATCAGATCAA GTTGATACAC	2640
	AAACACCAAT GAACCCTACG ACCAACTATG GTATTTCCAA AAAGTTCGCT GAACAAGCAT	2700
30	TACAAGAATT GATTAGTGAT TCGTTTAAAG TAGCAATTGT GAGACCACCA ATGATTTATG	2760
	GTGCACATTG CCCAGGAAAT TTCCAACGGT TAATGCAATT GTCAAAGCGA TTGCCAATCA	2820
	TTCCCAATAT TAACAATCAG CGCAGTGCAT TATATATTAA ACATCTGACA GCATTTATTG	2880
35	ATCAATTAAT ATCATTAGAA GTGACAGGTG TGTACCATCC TCAAGATAGT TTTTACTTTG	2940
	ATACATCGTC AGTAATGTAT GAAATACGTC GCCAATCACA TCGTAAAACG GTATTGATCA	3000
	ACATGCCTTC AATGCTAAAT AAGTATTTTA ATAAGTTGTC GGTCTTTAGA AAATTATTCTG	3060
40	GCAATTTAAT ATACAGCAAT ACGTTATATG AAAATAATAA TGCATTGAA ATTATTCCTG	3120
	GAAAAATGTC ACTTGTTATT GCGGACATCA TGGATGAAAC GACAACCAAA GATAAGGCAT	3180
45	AAGTCATCTA TTAAATAAAA TCAACATACA AATCGTTTTA TTTGGAGGTT ATAGTATGAA	3240
	GTTAACAGTA GTTGGCTTAG GTTATATTGG TTTACCAACA TCAATTATGT TTGCAAAACA	3300
	TGGcGTCGAT GTGCTTGGTG TTGATATTAA TCAGCAAACG ATTGATAAGT TACAAAGTGG	3360
50	TCAAATTAGT ATTGAAGAAC CTGGATTACA AGAGGTTTAT GAAGAGGTAC TGTCATCGGG	3420
	AAAATTGAAG GTATCTACAA CGCCAGATGC ATCTGATGTT TTTATCATTG CCGTTCCGAC	3480

	TAGTATTTTA TCATTTTGTAG AAAAAGGAAA TACCATTATT GTAGAGTCGA CAATTGCGCC	3600
	TAAAACGATG GATGATTTTG TAAAACCAGT CATTGAAAT TTAGGGTTTA CAATAGGTGA	3660
5	AGATATTTAT TTAGTGCATT GTCCAGAACG TGTACTGCCA GGAAAAATTT TAGAAGAATT	3720
	AGTTCATAAC AATCGTATCA TTGGCGGTGT GACTGAAGCT TGTATTGAAG CGGGTAAACG	3780
	TGTCTATCGC ACATTCGTTT AGGGAGAAAT GATTGAAACA GATGCACGTA CTGCTGAAAT	3840
10	GAGTAAGCTA ATGGAAAACA CATATAGAGA CGTGAACATT GCTTTAGCTA ATGAATTAAC	3900
	AAAAATTTGC AATAACTTAA ATATTAATGT ATTAGATGTG ATTGAAATGG CAAACAAACA	3960
	TCCGCGTGT AACATCCATC AGCCTGGTCC AGGTGTAGGC GGTCAATTGTT TAGCTGTTGA	4020
15	TCCGTACTTT ATTATTGCTA AAGACCCTGA AAATGCAAAG TTAATTCAAA CTGGACGTGA	4080
	AATTAATAAT TCAATGCCGG CCTATGTTGT TGATACAACG AAGCAAATCA TCAAAGTGT	4140
20	GAGCGGGAAT AAAGTCACAG TATTTGGTTT AACTTATAAA GGTGATGTTG ATGATATAAG	4200
	AGAATCACCA GCATTTGATA TTTATGAGCT ATTAAATCAA GAACCAGACA TAGAAGTATG	4260
	TGCTTATGAT CCACATGTTG AATTAGATTT TGTGGAACAT GATATGTCAC ATGCTGTCAA	4320
25	AGACGCATCG CTAGTATTGA TTTAAGTGA CCACTCAGAA TTTAAAAATT TATCGGACAG	4380
	TCATTTTGAT AAAATGAAGC ATAAAGTGAT TTTTGATACA AAAAATGTTG TGAAATCATC	4440
	ATTTGAAGAT GTATCGTATT ATAATTATGG CAATATATTT AATTTTATCG ACAAATAAAA	4500
30	TGTGTCAAAC TAGGGCATAAC ATGATTAAGG AAAGATAAGC TGTCATGTGT TTGAACCTCA	4560
	GAGAGGATAA TGTATGAAA AAAATTATGG TTATTTTCGG TACGAGACCC GAAGCAATAA	4620
	AAATGGCACC ATTAGTAAAA GAAATTGATC ATAATGGGAA CTTTGAAGCG AACATTGTGA	4680
35	TTACAGCACA ACATAGAGAT ATGTTAGATA GTGTGTTAAG TATATTGAT ATTCAAGCTG	4740
	ATCATGATTT AAATATTATG CAAGATCAAC AAACATTAGC AGGCCTTACG GCGAATGCAC	4800
	TTGCTAAACT TGATAGCATC ATTAATGAGG AACAACCGGA TATGATTTTA GTACATGGTG	4860
40	ATACTACAAC GACTTTTGTA GGAAGTTTGG CAGCATTTTA TCATCAAATT CCGGTGCGAC	4920
	ATGTAGAAGC TGGACTTCGA ACACATCAGA AATACTCACC ATTTCTGAA GAGTTAAATC	4980
45	GAGTCATGGT AAGTAATATT GCTGAATTGA ATTTTGCGCC AACAGTAATT GCAGCTAAAA	5040
	ATTTACTTTT TGAAAACAAA GACAAAGAGC GTATCTTTAT TACTGGAAAT ACAGTTATTG	5100
	ACGCATTGTC AACAACAGTT CAAAATGATT TTGTTTCAAC GATTATTAAT AAACATAAAG	5160
50	GCAAGAAAGT TGTTTTACTA ACAGCGCATC GTCGTGAAAA TATTGGGGAA CCGATGCATC	5220
	AGATTTTAA AGCAGTAAGA GATTTGGCAG ATGAATATAA AGATGTTGTC TTCATTTATC	5280

55

	GGATTGAATT AATTGAGCCA TTAGATGCGA TTGAGTTCCA TAATTTTACA AATCAATCGT	5400
	ACCTCGTGCT GACAGATTCT GGTGGTATTC AAGAGGAGGC TCCTACATTG GGAAAACCTG	5460
5	TGTTGGTATT AAGGAATCAT ACAGAGCGTC CCGAAGGCGT TGAGGCGGGA ACATCGAGAG	5520
	TAATTGGCAC AGATTATGAC AATATTGTTC GAAATGTGAA ACAATTGATT GAGGATGATG	5580
	AAGCGTATCA ACGTATGAGT CAAGCGAATA ATCCATATGG TGATGGACAA GCATCAGGAC	5640
10	GTATTTGTGA AGCAATAGAA TATTATTTTG GATTGCGCAC AGACAAGCCG GATGAATTCTG	5700
	TACCTTTACG TCACAAATAA TAAAAACCC CTAATCATGA AGTTGGTTTA GACAACCAGC	5760
	GGTGACTAGG GGTTTTAAAT ATATTATTTT TTGATAGTGG TAGCCAATAT CATATTTGAA	5820
15	TACTTTATTT GATAATATTG GACTTTGCTG TCCATCGTCA TCACTTTTAA AACGTACATT	5880
	TTTATGAGCT TCTTTAAATA CATCGGAATT CAACCAATTA TTAAAGCTAT CTTTCAGATT	5940
20	CCAAATAGTT AAGATTTTAA CTTCTGTCTG ATCCTCGGTA TTTAATGTTT TAGTGACAAA	6000
	CATTTGTTGG AAGCCTTCAA TAGTTTCAAT ACCTTGTCTA TTGTAAAAAC GTTCAATCGT	6060
	TTCTTCCGCA CTGCCTTTTT GTAATTGTAA TCTATTTTCT GCCATAAACA TGGGCAATCA	6120
25	CTCCTCTATT TTATGATTGG ATTTGGGTAA TGTTTTTACA AATGTAAAGA GTACAGCGGT	6180
	TTGTATGATA ACCATTATGA TTAATCCTAC ACGGACTGCA AGAACATCCA CCATATAAAT	6240
	TGAAAAACCT ATTACAATGT ATAAGCTAAT TAAATTTTAA ATTTTCTGTT GTAGCGTGTA	6300
30	GCCTCGATGT AAATAAAGT TTTCTACATA TTCTTTATAA ATTTTGTGAT TAATAAGCCA	6360
	ATTGTAAAG CGATCTGAAC TTCGAGCAAA GCAAAAACT GCTACGAGTA AAAAAGGGGT	6420
	CGTTGGCAGT AAAGGTAATA CGGCACCTGC AATACCAAGC GCTGTAAATA TTAAGCCAAT	6480
35	GACGATTAAA ATAAGTCGCA TTGAAAAAAC TCCATTCTAG TACTAATGCG CATGTAATAT	6540
	TGTTTITAGTA ATATAACTCA TGCTAAATAT AATGTGTATG ATAAGTGCAA TGACTCAGTA	6600
	AAATGAAACG ATGTTGAATT ATCCTTGTCA CATTACGCA TTTTAAGCGC GACTTTCATA	6660
40	ACAACCAAAC TATTTAATGA GAATTATTCT CAAGTATTAT AGTTATATTA TGTGTTTTAT	6720
	TTTTGAAAAG TGCAATATGT TTTGAAAAAT AAGATTATTT TTATGTGCAA AAACGACGCA	6780
	AAAGTTTAA AAATGAGACT TCTGTGAGCT GATTATTTTA TAAATGTAA ACGCTTACTA	6840
45	TATAATGTGA ATCATATCGT TTAAAAGCAT TATTAAATAT GATGCTAAGA GATTATATT	6900
	ATAGCCAATA AACAAAGGAG AGATAATATG GCAGTAAACG TTCGAGATTA TATTGCAGAG	6960
50	AATTATGGTT TATTTATCAA TGGGGAATTT GTTAAAGGTA GCAGTGACGA AACAATCGAA	7020
	GTGACTAATC CAGCAACTGG AGAAACACTA TCACATATTA CAAGAGCAAA AGATAAAGAT	7080

55

	TCAGAACGTG	CACAAATGTT	GCGTGATATT	GGTGATAAAT	TAATGGCACA	AAAAGATAAA	7200
	ATTGCAATGA	TTGAAACATT	AAATAATGGT	AAACCGATTG	GTGAGACAAC	AGCAATTGAT	7260
5	ATTCCATTTG	CTGCAAGACA	TTTCCATTAT	TTCCGAAGTG	TTATTGAAAC	AGAAGAAGGT	7320
	ACAGTGAATG	ATATCGATAA	AGACACAATG	AGTATCGTAC	GACATGAGCC	GATTGGCGTC	7380
	GTAGGTGCTG	TTGTTGCTTG	GAACCTCCCA	ATGCTATTAG	CTGCATGGAA	GATTGCGCCA	7440
10	GCCATTGCTG	CAGGTAATAC	AATTGTGATT	CAACCTTCGT	CTTCAACACC	ATTAAGTTTA	7500
	TTGGAAGTTG	CTAAAATTTT	CCAAGAGGTA	TTACCTAAAG	GTGTTGTCAA	TATACTAACG	7560
	GGTAAAGGTT	CAGAATCAGG	TAATGCAATT	TTCAATCATG	ATGGTGTAGA	TAAATTATCA	7620
15	TTTACGGGCT	CAACTGATGT	AGGTTATCAA	GTTGCCGAAG	CTGCAGCAAA	ACATCTAGTA	7680
	CCCGCTACAT	TAGAGCTTGG	TGGTAAAAGC	GCCAATATCA	TATTAGATGA	TGCTAATTTA	7740
	GACCTTGCAG	TTGAAGGTAT	TCAGTTAGGT	ATTTTATTCA	ACCAAGGTGA	AGTATGTAGT	7800
20	GCAGGTTCTC	GATTATTAGT	TCATGAAAAA	ATTTATGATC	AATTGGTGCC	ACGTTTACAA	7860
	GAGGCATTTT	CAAATATTAA	AGTTGGAAT	CCACAAGATG	AAGCTACACA	AATGGGTAGT	7920
25	CAAACCTGGTA	AGGATCAATT	AGATAAAATT	CAATCATATA	TTGATGCAGC	AAAAGAATCA	7980
	GATGCACAAA	TTTTAGCAGG	CGGTCATCGC	TTAACTGAAA	ATGGATTAGA	TAAAGGGTTC	8040
	TTCTTTGAGC	CGACATTAAT	TGCTGTGCCA	GACAATCATC	ACAAATTAGC	ACAAGAAGAA	8100
30	ATATTTGGAC	CAGTGTTAAC	AGTGATTAAA	GTGAAGGACG	ATCAAGAAGC	AATTGATATA	8160
	GCTAATGATT	CTGAGTATGG	TTTAGCAGGC	GGTGTATTTT	CTCAAAATAT	CACACGTGCA	8220
	TTAAATATTG	CTAAAGCTGT	ACGTACAGGA	CGTATTTGGA	TTAACTCTTA	CAACCAAGTA	8280
35	CCAGAAGGCG	CACCATTTGG	TGGTTATAAA	AAATCAGGTA	TCGGTCGAGA	AACTTATAAA	8340
	GGTGCGTTAA	GTAACATATCA	ACAAGTTAAA	AATATTTATA	TTGATACAAG	CAATGCTTTA	8400
	AAAGGTTTGT	ACTAGAATAA	ATATCGTTTT	TGAAGCGTGT	TTGTAGGTCA	GTCTAGCGGT	8460
40	AAGTCTTAAC	ATTTAACGGC	GTTGTTTAGA	TTTTAAGCAA	AACAAAATAT	ATAGGAACAC	8520
	GTATCATGAT	ATTAGGATAT	AATGACTAAA	ATAATAGCAG	TAGGATGGTT	TTTAATTGCA	8580
45	AATCATCTTA	CTGCTGTTTT	TAATTATGCT	AATTTGCGAT	GCGGCTATTA	TAAGGACAGA	8640
	GTTGTTTATT	AATTATGGTG	ATTTAGAAAT	ATGAAGTTCA	ATATGCAAAG	TCATCGTTTG	8700
	TTTTAATATG	CGGAACAATC	ATTAAAGTTA	TTGCGATTTT	TTGAACTTAA	TGAAACTAAA	8760
50	CAATAAATTT	GAGATACTTT	TTTGTCATTT	TTATGTAACT	AACACAATAA	TCTCGTACAT	8820
	TATTAAAATT	TTCTATATGA	TAGGAATAAA	GCAAAGCGCG	AGTGTGCTGT	AAAAGTTTTT	8880

55

	GATGATGTAT AAATCATGGT TAATTACGGA AGCATTAAATA TTAACCTGAG AAGCTATAAA	9000
	GAATTATTTT TAAAAGCGAC AATATTAAAT ACGACGCATT TATTTAGGAG TGGCAAACGT	9060
5	ATGAATGGGA AAAAGGCGAA TACGATAAAC AGATACAAAT ATTTTCATCA TGTCAATCAT	9120
	CAAAAAATTC AACAAAGTTC TAAAAAGACG CTGTGGGCAT CACTAATCAT CACATTGTTA	9180
	TTTACAGTGA TTGAATTGT CGGAGGTTTA GTATCTAAT CATTGGCATT ACTGTCAGAT	9240
10	TCATTTTATA TGCTTAGTGA TGTATTAGCA CTGGGTTTAT CTATGTTGGC CATTTATTTT	9300
	GCAAGTAAAA AGCCGACTGC ACGATACACA TTTGGATATT TAAGATTGA GATATTAGCT	9360
	GCATTTTAA ATGGTTTAGC ATTAATTGTA ATTTCAATCT GGATTTTATA TGAAGCTATT	9420
15	GTACGTATTA TTTATCCGCA ACCAATTGAA AGTGGCATT TGTATTATGAT TGCTAGTATT	9480
	GGTTTACTCG TCAATATTAT TTTGACTGTT ATCCTTGTA GGTCTTTAAA ACAAGAAGAC	9540
	AATATCAATA TTCAAAGTGC ATTATGGCAT TTCATGGGAG ACTTATTGAA CTCTATTGGT	9600
20	GTCATCGTTG CAGTTGTATT GATTTACTTT ACAGGATGGC GCATCATCGA CCCAATCATT	9660
	AGTATTGTAA TTTCACTCAT CATTTTACGT GGTGGTTATA AAATTACGCG TAATGCGTGG	9720
	CTAATTTTAA TGGAAAGTGT GCCTCAACAT TTGGATACTG ATCAAATTAT GGCAGATATT	9780
25	AAAAACATAG ATGGCATATT AGATGTACAT GAATTTTCAAT TGTGGAGTAT TACAACAGAG	9840
	CATTATTCAT TAAGTGCCCA TGTGTGTGTA GATAAAAAAT ATGAGGGTGA TGATTATCAA	9900
30	GCGATTGATC AAGTATCATC ATTGTTGAAA GAAAAATATG GCATTGCACA TTCAACGTTG	9960
	CAAATTGAAA ACTTGCAATT GAATCCATTA GATGAGCCAT ACTTCGACAA ATTAACATAA	10020
	ATAAACATT GTAGCGCCTA AAACATTAAT CTATGTCATA GGCGCACGTT TCGTTTTATA	10080
35	CTTATGTTGC ATCATTTAAA TGATTTTCGT CAATTTCTTT GATGCTATCT ACATCTAACA	10140
	CGACATCTTT AGGTTTCAAA ATATGAATAT GTTTTTCATC ATTTGTATGT AAAATGCGTT	10200
	CTATGATGTA CCTTGACCG GCCATTGTTT CTACAGCAAT CTTTTTGTTT CTAGCTAAAC	10260
40	TTGCTACGAC AGATTCTTTA TCCATAATGA TAGCCCCCTA TATATATGTT TATTTACTTA	10320
	TACCCTAACA TGATTTTTAT ACTCTTTGAA AATATATTTT ACAGAATTTT ATCTAAATAT	10380
	TTAAAAAAT ATCTTAATAT CCTTGTAATC CGATAAGAAT TATAGTAATA TTTTTC AAC	10440
45	CATGTTTATA GGAGGTCTTA TTAATGACAT TATTTTATT AGAAGCTAAC AATCTTGATT	10500
	TTGCATCAAC GAAAGAAGAA CTAGAAGCAA AGGCAGCATC ACTATCTACG AAGACAATTC	10560
	CAACATTAAT TGAAGTACAA GCTACTGAAA ATTTAACTCA TGGTTATTTT ATTGTGGAAG	10620
50	CAAATGACGA aGCAGAAGCT AAACAATTTT TAACAGAAGC AGATATTAGT ATTCAATTAG	10680

55

	TTGATTACCT TGTAACCTGG AACATTCCGG AAGGCATTAC GATGGATCAA TATTTAGCAC	10800
	GTAAAAAGAA AAATTCTGTT CATTATGAAG AAGTGCCAGA AGTTGAATTT AAACGCACAT	10860
5	ATGTATGTGA AGATATGTCT AAATGTATTT GTTTATACAA CGCACCTGAT GAAGAAGCGG	10920
	TACGTCGCGC GCGCAAAGCA GTTGATACAC CGATTGATGG CATCGAAAAA CTTTAATAAG	10980
10	ACAACAAGTT GATGAGATAT ATGTATATAG GTTTGGCATG GATTTTCGATT GCAGTTAATT	11040
	AGAATAGCTC AATGCTATAA ATGTAAGTAG TTGATATGAA GAACTAATG AACTAAATGC	11100
	AAGTATTGTC TAAAACAATC ATTTTATTGA AATTTAGTAG AGCTGAAATT AATATAACGT	11160
15	CGTTAATTGA ATAACGCTTA TGTATAAGA GCACTCATAC CAAACCATAA TCATCTATAG	11220
	ATATAACAAT TCACGATATA AGGGCTGTGT TTGGCATAGC CCTTTAGATA TACACTTAAT	11280
	TCCTATTAAA ATAGTAGGGA TTTAAAGGGG GCTTGTCTAG ATTAAAATTC AACAAATTACA	11340
20	ACATCACTTT GGATCACATA AAGTAATTCA TAACCTTAAT TTGGACATTA GCAAGGGAGA	11400
	AATAGTCACT TTCATAGGGA AAAGTGGTTG CGGAAAGTCT ACTTTACTCA ATATTATCGG	11460
	TGGATTTATT CATCCATCGT CTGGTCGTGT CATTATTGAT AACGAAATTA AACAAACAGCC	11520
25	ATCTCCAGAT TGTTTAATGC TATTTCAACA TCATAATTTG CTGCCATGGA AAACGATTAA	11580
	TGACAACATT AGGATTGGAT TACAACAGAA AATTAGTGAT GAAGAGATTA ACGCACAGCT	11640
	TAAATTAGTT GATTTAGAAG ACAGGGGAAA GCATTTTCCC GAGCAACTGT CCGGGGGTAT	11700
30	GAAACAACGT GTGGCACTAT GTCGAGCGCA TGTGCATAAG CCTAACGTTA TATTGATGGA	11760
	TGAGCCATTA GGTGCATTAG ATGCATTTAC ACGTTATAAA CTTCAGGATC AACTAGTGCA	11820
	AACTAAAACAT AAAACGCAAT CAACTATTAT TTTAGTGACC CATGACATTG ATGAAGCTAT	11880
35	TTATCTTTCC GACCGCATTG TTCTGTTAGG TGAAGGGTGC AATATTATTT CTCAATATGA	11940
	AATFACAGCA TCACATCCAC GCAGTCGTAA TGATAGCCAC CTACTTAAGA TTCGTAATGA	12000
	AATTATGGAA ACATTTCAT TGAATCATCA TCAAGTTGAA CCTGAATATT ATTTATAAGG	12060
40	AGTGAGTGAC GATGAAAAGG TTAAGCATAA TCGTCATCAT TGGAATCTTT ATAATTACAG	12120
	GATGTGATTG GCAAAGGACG TCTAAAGAAC GGTCTAAAAA TGCCCAAAT CAGCAAGTGA	12180
45	TTAAAATTGG ATATTGCGG ATTACACATT CAGCTAATTT GATGATGACT AAAAAATTAT	12240
	TATCACAATA CAATCATCCG AAATATAAAC TAGAATTAGT TAAATTCAAT AATTGGCCAG	12300
	ATTTAATGGA CGCATTAAAC AGTGGTCGTA TTGATGGTGC ATCAACTTTA ATAGAGCTAG	12360
50	CGATGAAATC AAAACAGAAG GGCTCAAATA TAAAGGCTGT GGCATTGGGC CATCATGAAG	12420
	GCAATGTCAT TATGGGACAA AAAGGTATGC ACTTAAATGA ATTTAATAAT AATGGCGATG	12480

55

	GTAAACAATT AAAGATTAAA CCGGGGCATT TTAGCTATCA TGAAATGTCG CCAGCAGAAA	12600
	TGCCAGCCGC ATTGAGTGAA CACAGAATTA CAGGGTATTC TGTAGCCGAA CCATTCGGTG	12660
5	CACTGGGTGA AAAGTTAGGC AAAGGTAAGA CTTTGAAACA TGGTGATGAC GTTATACCTG	12720
	ATGCGTATTG CTGTGTGCTA GTACTGAGAG GGGAAATTGCT TGATCAACAC AAGGATGTAG	12780
	CGCAAGCATT TGTACAAGAT TATAAAAAGT CTGGCTTTAA AATGAATGAT CGCAAGCAAA	12840
10	GTGTAGACAT TATGACGCAT CATTTTAAAC AAAGTCGTGA CGTTTTAACA CAGTCAGCGG	12900
	CATGGACATC CTATGGTGAT TTAACAATTA AGCCATCCGG CTATCAAGAA ATTACGACAT	12960
	TGGTAAACA ACATCATTG TTTAATCCAC CTGCATATGA TGACTTTGTT GAACCGTCAT	13020
15	TGTATAAGGA GGCATCGCGT TCATGACACG TCCCACAAAT AACAAATTTA TATTACCTAT	13080
	TATCACATTT ATTATTTTCT TAGGCATTTG GGAATGGTC ATTATTATTG GGCATTACCA	13140
20	ACCTGTATTG TTACCGGGTC CTGCTCTTGT AGGAAAAAGT ATATGGTCTT TCATTGTTAC	13200
	TGGAGAAATT TTCCAACATT TAGCAATTAG TTTATGGAGA TTTGTAGCGG GCTTTGTTGT	13260
	CGCATTGTTG GTTGCTATTC CATTGGGCTT CTGCTTGGA AGGAATCGTT GGCTATACAA	13320
25	CGCTATCGAA CCGCTATTTT AATGATTAG GCCGATATCT CCGATAGCAT GGGCACCATT	13380
	TGTTGTTCTA TGGTTTGGTA TTGGTAGTTT GCCAGCGATT GCGATTATTT TTATCGCTGC	13440
	TTTTTTCCCA ATTGTGTTCA ATACTATTAA AGGCGTTAGA GACATTGAAC CTCAATATTT	13500
30	AAAAATAGCA GCAAATTTAA ATTTAACTGG GTGGTCATTG TATCGCAATA TATTATTTCC	13560
	CGGGGCATTT AAACAAATCA TGGCTGGGAT ACATATGGCG GTAGGAACAA GTTGGATATT	13620
	TTTAGTTTCT GGTGAAATGA TTGGTGACA ATCGGGATTA GGTTTTTTAA TCGTTGATGC	13680
35	ACGAAATATG TTGAACCTAG AAGATGTTTT AGCAGCAATA TTCTTTATCG GATTATTTGG	13740
	TTTTATTATT GATCGATTCA TTAGTTATAT TGAGCAGTTT ATACTTAGAA GATTGGTGA	13800
	ATAAGGAGAG ATGATGATGA CTTTAGAAAC GCTTATCAAA GAACAATTAG ATCCTCATTT	13860
40	AGTAGAAGTT GATGAAGGGA CGTATTATCC GAGAACATTT ATTCAGCAAT TATTTGTAGA	13920
	TGGTTATTTT GGTGAGGCGG CATTGAGAAA AAATGCTGAA GTAATCGAAG CTGTATCGCA	13980
	GTCTTGTTTG ACAACAGGAT TTTGTTTATG GTGCCAATTA GCTTTTTCAA CGTATTTAGA	14040
45	AAATGCCACG CAGCCACATT TAAATAATGA CTTACAACAG CAATTGTTAT CTGGAGAAAT	14100
	ATTAGGTGCT ACCGGATTGT CTAATCCGAT GAAGTCATTT AATGATTTAG AAAAGTTGAA	14160
50	CCTTGAACAC ACTTATGTTG ATGGACAATT GGTGTGTCAGT GGACGTATGC CAGCTGTAAG	14220
	TAATATTCAA GAAGACCATT ATTTTGGTGC GATTTTCGAA CATGAATCAT CAGATGAATT	14280

	TTTAGGAGTC AACGGGTCAG CAACGTATCA AATCACATTG AATCAAGTCG TAGTGCCACA	14400
	ATCACAAATT ATCACGCATG ATGCGAAGCA GTTTGCGGCA ACTATTCGCC CGCAATTTAT	14460
5	TGCTTACCAA ATTCCAATAG GATTAGGCTC AATTAAAAGT TCTTTAGAGT TAATTGATGC	14520
	ATTTTCAAAT GTGCAAAACG GAATAAATCA ATATTTAGAG TATGATGTTG AAGCTTTTAA	14580
	AAAACGTTAT CGTCAACTTA GAGAGGAATA TTATGCAATA TTAGATGACG GTAACCTAAC	14640
10	TTCACATTTA AATGAATTAA TATCATTGAA GAAGGACATC GGCTATTTAT TGTTAGATGT	14700
	AAATCAAGCT TCTGTTGTCA ATGGTGGTTC TAGAGCGTAC ACACCATATT CGCCACAAGT	14760
	TCGCAAGTTA AAAGAAGGAT TCTTCTTCGC AGCATTGACA CCGACATTAA GACATTTAGG	14820
15	TAAACTTGAA GCAGAGTTGA AGGGGTAAGT GTGATAAGCT GATTTTTTGT TTAGATGCGT	14880
	TTGTTGAAAC ATTTTTTAAA ATAATATAAA TCTTAGTTTA TAAACATTTT CTGTTAATTT	14940
20	GTTATATCCT TTAACTAGG AAAATATACA TTTCGTAATA ATAATAATCG TTATCATTGA	15000
	AAAAGTGTTA ATAAGGTGTA TAATGAAAT GTGAACAATT AATGAACTTC TTATTTTAAA	15060
	GAAGGTGAAT ACTATAGATA CGCATACTAA AGAACAACAA TTCTCGAATC TAGTAAGATC	15120
25	TTATCGTAAA GAATACGTGG GTAAAGGACC CAATAGTATT CGAGTGTCCG TTAAAGATAA	15180
	TTGGGCGATT GCACATATGA CAGGTGTTTT GAGTAAAGTT GAGAGTTTTT ACCTAAACGA	15240
	CAAACGCAAT GAATCGATGC TCCATTATAC ACGCACAGAG AAGATTAAAC AGATGTATAA	15300
30	AGAAATAGAT GTAAATGAGA TGGAAAGTCT TGTAGGCGCT AAGTTTGTAA AATTATTTAC	15360
	AGATATTGAT TTGAATGATG ATGAAGTCAT TTCAATATTT GTTTTCGATA AGTCAATAGA	15420
	ATAAGTGTTG CTGGTGTAAG GTACACGGTG CTGTTTGCTA ACTTCGCTTT GAATTTAACA	15480
35	ATAATTCAAG GGGGTGGTAT GTCAAACGGT GCCGTTTTTT TGTCATATTT TTAAACAAG	15540
	CAACATGCAA CACGTACTTT AAGGAAGTCA AAATTTATCA TTTAGGAGAG ATGGATATGA	15600
	AAATCGTAGC ATTATTTCCA GAAGCAGTAG AAGGTCAAGA AAATCAATTA CTTAATACTA	15660
40	AAAAAGCATT AGGATTAAAA ACATTTT TAG AGGAAAGAGG ACATGAGTTC ATTATATTAG	15720
	CAGATAATGG TGAAGACTTA GATAAACATT TACCAGATAT GGATGTGATT ATTAGTGCGC	15780
45	CATTTTATCC TGCATATATG ACTCGTGAAC GTATTGAAAA AGCACC GAAC TTGAAATTAG	15840
	CAATTACAGC AGGTGTAGGA TCTGACCATG TAGATTTAGC GGCAGCAAGT GAACACAATA	15900
	TTGGTGTGCGT TGAAGTTACA GGAAGTAATA CAGTTAGTGT GGCAGAACAT GCGGTTATGG	15960
50	ATTTATTAAT ACTTCTTAGA AACTATGAAG AAGGTCATCG TCAATCAGTA GAAGGTGAAT	16020
	GGAACTTGTC TCAAGTAGGT AATCATGCGC ATGAATTACA ACACAAAACA ATTGGTATTT	16080

55

TACAACACTA TGATCCAATC AATCAACAAG ACCATAAATT GTCTAAATTT GTAAGCTTTG 16200
 ATGAACCTTGT TTCAACAAGT GATGCGATTA CAATTCATGC ACCATTAACA CCAGAAACTG 16260
 5 ATAACTTATT TGATAAAGAT GTTTTAAGTC GTATGAAAAA ACACAGTTAT TTAGTGAATA 16320
 CTGCACGTGG TAAAATTGTA AATCGCGATG CGTTAGTTGA AGCGTTAgCA TCCGAGCATT 16380
 TACAAGGATA TGCTGGTGAT GTTTGGTATC CaCAACctGC ACCTGCTGAT CATCCATGGA 16440
 10 GAACAATGCC TAGAAATGCT ATGACGGTTC ACTATTCAGG TATGACTTTA GAAGCACAAA 16500
 AACGTATTGA AGATGGAGTT AAAGATATTT TAGAGCGTTT CTTCAATCAT GAACCTTTCC 16560
 AAGATAAAGA TATTATTGTT GCAAGTGGTC GTATTGCTAG TAAAAGTTAT ACAGCTAAAT 16620
 15 AGAATAAGGA TGCTGGGCTA GCGATTAAAG CTTTCAATTT TATATAAATG AATCATATAA 16680
 GCACTACTGC TGTGTAAAG ATGGCAGTAG TTTTTTTATG ATTACATCTA AGTATAGTCA 16740
 CGGCTATGTT AGGACAATGA TTTAACATTT ACGCACATAT GTGTTCACTT ACGCAATTAT 16800
 20 TGaAAATht CATTCAATGTG GnaATC 16826

(2) INFORMATION FOR SEQ ID NO: 47:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4012 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

TTCAATGAGA GTAGTGGGCT GATGTTTAGC GATATCGCGT AAGATTAAAC ATTGGCCATA 60
 35 ATATATATTG TGTCTTTCTA AAATCGGCTC GGCTAATTTT AAATAGGGGC GATATATTGT 120
 TATAAACTA TTGAAAAATT CTTGTGATAG CATAGTGACA TCTCCTAAGA CAAAATAGTT 180
 AGCTTAGCTA mCCTTTTtAC AACAAATAGTA ATTATAAAC GGGAGCAATT AGAAATCAAT 240
 40 ATATAATTAT TAAGAGCAAA AATAATTATA CTTTGTTAAA ATAAGCGTAA TTACATGTAA 300
 ATAGGGGGAT ACTAATGATA TTGAAATTTG aTCACATCAT TCATTATATA GATCAGTTAG 360
 ATCGGTTTAG TTTTCCAGGA GATGTTATAA AATTACATTC AGGTGGGTAT CATCATAAAT 420
 45 ATGGAACATT CAATAAATTA GGTATATCA ATGAAATTA TATTGAGCTA CTAGATGTAG 480
 AAAATAATGA AAAGTTGAAA AAGATGGCAA AAACGATAGA mGGCGGAGTC GCTTTTGCTA 540
 CTCAAATTGT TCAAGAGAAG TATGAGCAAG GCTTTAAAAA TATTTGTTTG CGTACAAATG 600
 50 ATATAGAGGC AGTTAAAAAT AACTACAAA GTGAGCAGGT TGAAGTAGTA GGGCCGATTC 660

55

	ATCAGGATGA TGATGAAATT AAGCCACCAT TTTTATTCA ATGGAAGAA AGTGATTCCA	780
	TGCGTACTAA AAAATTGCAA AAATATTTTC AAAACAATT TTCAATTGAA ACTGTTATTG	840
5	TGAAAAGTAA AAACCGATCA CAAACAGTAT CGAATTGGTT GAAATGGTTT GATATGGACA	900
	TTGTAGAAGA GAATGACCAT TACACAGATT TGATTTTAAA AAATGATGAT ATTTATTTTA	960
	GAATTGAAGA TGGTAAAGTT TCAAAATATC ATTCGGTTAT CATAAAGAC GCACAAGCAA	1020
10	CTTCACCATA TTCAATTTTT ATCAGAGGTG CTATTTATCG CTTTGAACCA TTAGTATAAA	1080
	TATACGTAAG TGCTATGAGC GAGAATGCCC ATATGAATAA TGACAAGCAC AATGGAAAGA	1140
	ATCGTTAATA TATTATTTAA TCGTGATGAC TTAATTAAAA TGAAAAGAT TGATAATATA	1200
15	AATGTGAAAA AGATAAGTAT AACCCGTAAA CTAAGTAAT TCACGGTGAG AGGTTGACTC	1260
	AATGTCATAA TGATTGCAAC GATGTTTATA ATTATAAATA GACTTAAAAT AATTGTTCTC	1320
	ATATCAAACA CCTCATTGTT AGATTATTGA CATTATAACA GGGTAATTG TATATGAACA	1380
20	TTAATGTGGT TGCTTGAGGA AAAATTTATT CATTGAAGTC AAGTTGGTTC ATTTAGAAA	1440
	TGAATATCGT GTTAGATGAT GAAAGTATAT TGAAGTATAG GTAAGTAGTT GAAAAGTATT	1500
25	AATTGTACGA TAACATTAAA TTTAACACGA AACATAGATA TAAATGATT CACAATTAAA	1560
	ATGGGTAAAT TTGAACCTGC TAAACTATTA ATTGGAGCAT GGACATTTCA AAAATAAGAG	1620
	TTCAATCTT ACACAAGCTC TGAATCGACA CTATAAGATA CAACTGTAT AATTAAAGGT	1680
30	ATTGTTAAAT AGAAGGAGAT ATCATAAATC ATGGAAGA TGCATATCAC TAATCAGGAA	1740
	CATGACGCAT TTGTTAAATC CCACCCAAAT GGAGATTAT TACAATTAAC GAAATGGGCA	1800
	GAAACAAAGA AATTAAGTGG ATGGTACGCG CGAAGAATCG CTGTAGGTCG TGACGGTGAA	1860
35	GTTCAGGGTG TTGCGCAGTT ACTTTTTAAA AAAGTACCTA AATTACCTTA TACGCTATGT	1920
	TATATTTCGC GTGGTTTTGT TGTGATTAT AGTAATAAAG AAGCGTTAAA TGCATTGTTA	1980
	GACAGTGCAA AAGAAATTGC TAAAGCTGAG AAAGCGTATG CAATTAAAAT CGATCCTGAT	2040
40	GTTGAAGTTG ATAAAGGTAC AGATGCTTTG CAAATTTGA AAGCGCTTGG TTTTAAACAT	2100
	AAAGGATTTA AAGAAGGTTT ATCAAAAGAC TACATCCAAC CACGTATGAC TATGATTACA	2160
	CCAATTGATA AAAATGATGA TGAGTTATTA AATAGTTTGG AACGCCGAAA TCGTTCAAAA	2220
45	GTGCGCTTGG CTTTAAAGCG AGGTACGACA GTAGAAGAT CTGATAGAGA AGGTTTAAAA	2280
	ACATTTGCTG AGTTAATGAA AATCACTGGG GAACGCGATG GCTTCTTAAC GCGTGATATT	2340
	AGTTACTTTG AAAATATTTA TGATGCGTTG CATGAAGATG GAGATGCTGA ACTATTTTTA	2400
50	GTAAAGTTGG ATCCAAAAGA AAATATAGCG AAAGTAAATC AAGAATTGAA TGAACCTCAT	2460

55

CAAAATATGA TTAATGATGC GCAAAATAAA ATTGCTAAAA ATGAAGATTT AAAACGAGAC 2580
 CTAGAAGCTT TAGAAAAGGA ACATCCTGAA GGTATTTATC TTTCTGGTGC ACTATTAATG 2640
 5 TTTGCTGGCT CAAAATCATA TTACTTATAT GGTGCGTCTT CTAATGAATT TAGAGATTTT 2700
 TTACCAAATC ATCATATGCA GTATACGATG ATGAAGTATG CACGTGAACA TGGTGCAACA 2760
 ACTTACGATT TCGGTGGTAC AGATAATGAT CCAGATAAAG ACTCAGAACA TTATGGATTA 2820
 10 TGGGCATTTA AAAAAGTGTG GGGAACATAC TTAAGTGAAA AGATTGGTGA ATTTGATTAT 2880
 GTATTGAATC AGCCATTGTA CCAATTAATT GAGCAAGTTA AACC GCGTTT AACAAAAGCT 2940
 AAAATTAAAA TATCTCGTAA ATTAAAACGA AAATAGATTA ACGACTGAAA TCTGAACGCT 3000
 15 CATAAGACTG TCATTTGCGT TCAGATTTTT TTACACAATA TAGAATGGTT GAGTAAAATA 3060
 TTTTTGAATA TAGTGAAAGA GGGGGAAGTA CTGTGATAAA AAAGCTATTA CAATTTTCTT 3120
 TAGGGAATAA GTTTGCTATC TTTTAAATGG TTGTTTTAGT TGTCTTGGGC GGTGTATATG 3180
 20 CGAGTGCTAA ATTGAAATTA GAATTACTAC CAAATGTACA AAATCCAGTT ATTTCACTTA 3240
 CAACAACAT GCCGGGTGCA ACGCCACAAA GTACCCAAGA TGAAATAAGT AGTAAAATTG 3300
 ACAATCAAGT AAGATCATTG GCATATGTGA AAAATGTTAA AACGCAATCC ATACAAAATG 3360
 25 CTTCAATTGT AACAGTTGAA TATGAAAATA ATACAGATAT GGATAAAGCA GAAGAACAGC 3420
 TTAAAAAAGA AATCGATAAA ATTAAATTTA AAGATGAAGT TGGTCAACCA GAATTAAGAC 3480
 GTAATTCGAT GGATGCTTTT CCGGTTTTAG CATATTCATT TTCAAATAAA GAGAATGACT 3540
 TGAAAAAAGT AACGAAAGTA CTGAATGAAC AATTAATACC AAAATTGCAA ACGGTAGATG 3600
 GTGTGCAAAA TGC GCAATTA AATGGGCAGA CGAACCGTGA AATCACCCCTT AAATTTAAGC 3660
 35 AAAATGAACT TGAAAAATAT GGGTTGACTG CTGATGATGT AGAAAACTAT CTAAAAACGG 3720
 CAACAGAAC AACGCCACTT GGATTGTTCC AATTTGGTGA TAAAGATAAT CAATTGTTGT 3780
 TGATGGTCAA TATCAATCTG TTGATGCTTT TAAAAACATA AATATTCCAT TAACGTGGCA 3840
 40 GGAGGACCAA GGCATCTCA TCCCAAAGTG ACCATAAACC AAATTCAGCC ATGTCAGACG 3900
 TTATCAGGCA TCACCACAGC AAATTCAAAG CGTCAGCnCC AATATATAGT GGATGCCGCA 3960
 45 nGAACTAGGG GTTTAGCGnT ATCAGTGGTG TGGCGACTCT ATTCTAAACG AT 4012

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7778 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

	CAATATAGGT CGCCGAGTTT CAACTaCATC AACTGGTTCA GTTACATTAG ATAATGCGCT	60
5	AGGTGTAGGT GGCTATCCTA AAGGACGAAT TATTGAAATT TATGGTCCTG AAAGTTCTGG	120
	TAAGACAACA GTAGCGCTTC ACGCTATTGC TGAAGTACAA AGTAATGGCG GGGTGGCAGC	180
	ATTTATCGAT GCTGAACATG CTTTAGATCC AGAATATGCT CAAGCATTAG GCGTAGATAT	240
10	CGATAATTTA TATTTATCGC AACCGGATCA TGGTGAACAA GGTCTTGAAA TCGCCGAAGC	300
	ATTTGTTAGA AGTGGTGCAG TTGATATTGT AGTTGTAGAC TCAGTTGCTG CTTTAACACC	360
15	TAAAGCTGAA ATTGAAGGAG AAATGGGAGA CACTCACGTT GGTTTACAAG CTCGTTTAAT	420
	GTCACAAGCG TTACGTAAAC TTTCAGGTGC TATTTCTAAA TCAAATACAA CTGCTATTTT	480
	CATCAACCAA ATTCGTGAAA AAGTTGGTGT TATGTTCCGT AATCCAGAGA CTACACCAGG	540
20	TGGACGTGCA TTAATAATTCT ATAGTTCAGT AAGACTAGAA GTACGTCGTG CAGAACAGCT	600
	TAAACAAGGA CAAGAAATTG TAGGTAATAG AACTAAAATT AAAGTCGTTA AAAATAAAGT	660
	GGCACCACCA TTTAGAGTAG CTGAAGTTGA TATTATGTAT GGACAAGGTA TTTCTAAAGA	720
25	GGGTGAACTT ATTGATTAG GTGTTGAAAA CGACATCGTT GaTAAATCAG GAGCATGGTA	780
	TTCTTACAAT GGCGAACGAA TGGGTCAAGG TAAGGAAAAT GTTAAATGT ACTTGAAAGA	840
	AAATCCACAA ATTAAGAAG AAATTGATCG TAAATTGAGA GAAAAATTAG GTATATCTGA	900
30	TGGTGATGTT GAAGAAACAG AAGATGCACC AAAGTCATTA TTTGACGAAG AATAGTACAC	960
	AAATTTATAT CTATAGTTAA ACTTAGCAAA TATCCTTATA GGATTGATTG AAAGTGATAT	1020
	TCATCTCATA AAGCTAGAAT AATATCTAAC TTTATGGGAT ACACTACAAA TCGAGACTAT	1080
35	AAGGTTTTTT ATTTTATTTA TTATTACATT ATCAATAGTT TTATAATCGA GCTTCAAAAC	1140
	TTTAGAAAAT AGTAGAAATA GCATTCAATA TAGTGCAAAA GTGCAAATTG ATAACCTGAC	1200
	ACTTATCTCC TATAAACCGT ACAATTAATT TGTATGATT ATATATAATT TCATAAAGTC	1260
40	ATATTGAATT TCATATAAAG AGCAAACCCT AGAAAAGGAG GTGTTTGTGT GAATTTATTA	1320
	AGCCTCTAC TCATTTTGCT GGGGATCATT CTAGGAGTTG TTGGAGGGTA TGTGTTGCC	1380
45	CGAAATTTGT TGCTTCAAAA GCAATCACAA GCTAGACAAA CTGCCGAAGA TATTGTAAAT	1440
	CAAGCACATA AAGAAGCTGA CAATATCAAA AAAGAGAAAT TACTTGAGGC AAAAGAAGAA	1500
	AACCAAATCC TAAGAGAACA AACTGAAGCA GAACTACGAG AAAGACGTAG CGAACTTCAA	1560
50	AGACAAGAAA CCCGACTTCT TCAAAAAGAA GAAAACCTAG AGCGCAAATC TGATCTATTA	1620
	GATAAAAAAG ATGAGATTTT AGAGCAAAAA GAATCAAAAA TTGAAGAAAA ACAACAACAA	1680

55

	CGCATCTCCG GTCTCACTCA AGAAGAAGCT ATTAATGAGC AACTTCAAAG AGTAGAGGAA	1800
	GAAGTGTAC AAGATATTGC AGTACTTGTT AAAGAAAAAG AAAAAGAAGC TAAAGAAAAA	1860
5	GTTGATAAAA CAGCAAAAGA ATTATTAGCT ACAGCAGTAC AAAGATTAGC AGCAGATCAC	1920
	ACAAGTGAAT CAACGGTATC AGTAGTTAAC TTACCTAATG ATGAGATGAA AGGTCTGAATC	1980
10	ATTGGACGAG AAGGACGAAA CATCCGCACA CTTGAACTT TAACTGGCAT TGATTTAATT	2040
	ATTGATGACA CACCAGAAGC GGTATATTA TCTGGTTTTG ATCCAATAAG AAGAGAAATT	2100
	GCTAGAACAG CACTTGTTAA CTTAGTATCT GATGGACGTA TTCATCCAGG TAGAATTGAA	2160
15	GATATGGTCG AAAAAGCTAG AAAAGAAGTA GACGATATTA TTAGAGAAGC AGGTGAACAA	2220
	GCTACATTTG AAGTGAACGC ACATAATATG CATCCTGACT TAGTAAAAAT TGTAGGGCGT	2280
	TTAAACTATC GTACGAGTTA CCGTCAAAT GTACTTAAAC ATTCAATTGA AGTTGCGCAT	2340
20	CTTGCTAGTA TGTTAGCTGC TGAGCTAGGC GAAGATGAGA CATTAGCGAA ACGAGCTGGA	2400
	CTTTTACATG ATGTTGGTAA AGCAATTGAT CATGAAGTAG AAGGTAGTCA TGTTGAAATC	2460
	GGTGTAGAAT TAGCGAAAAA ATATGGTGAA AATGAAACAG TTATTAATGC AATCCATTCT	2520
25	CATCATGGTG ATGTTGAACC TACATCTATT ATATCTATCC TTGTTGCTGC TGCAGATGCA	2580
	TTGTCTGCGG CTCGTCCAGG TGCAAGAAAA GAAACATTAG AGAATTATAT TCGTCGATTA	2640
	GAACGTTTATG AAACGTTATC AGAAAGTTAT GATGGTGTAG AAAAAGCATT TGCGATTGAG	2700
30	GCAGGTAGAG AAATCCGAGT GATTGTATCT CCTGAAGAAA TTGATGATTT AAAATCTTAT	2760
	CGATTGGCTA GAGATATTAA AAATCAGATT GAAGATGAAT TACAATATCC TGGTCATATC	2820
	AAGGTGACAG TTGTTGAGA GACTAGAGCA GTAGAATATG CGAAATAATT TTTGTCTCCC	2880
35	TCACAAATTA GTGAGGGAGC TTTTAAAGT TGTAGTCTTA AtCTAGTTAG ACAGCACTTT	2940
	ATCGGTAATA ACTATATTAA ACAGTAGTTA TTTGAAAGTA AGACGGACCT TATATTAAAT	3000
	AAGAAGTTAT TGCTTTTAAT AAAAATGTTT TAGGCTTCGT AATTACTATA TTTATATTAT	3060
40	GTAAACCTAT AAAGATGATT GGTTTTCTAT CCAATAAAAA AGAAGAGAAG ATGTAACACA	3120
	TCTTCTCTTC YGCAATATTA ATTAGGATTT ATTTCTAAGT TGAGTTATTT TAATTGTAAA	3180
45	TCTGTTTTCT TTAATCTTT TATAACTTCT GCAGTATCAT AACAATTGTG TGCAATTGTT	3240
	GAATATCTCT CTGCTAAACG ATATGCATTA ATGTAAAGCT TTAACTTTC TTTAGCTATA	3300
	TCCTCTGCAT CTTGGAATTT TGATGGGTTA GACATAACCA CTAATCTGTC AAATTTTCT	3360
50	GGATCAATAT TAATAGACAT GTATTTATTT ACAACTCCTA TTTATTTTGA TGTCTTAATA	3420
	CTAACATATT GAAGTTTTCA GACAAAGTAA TGTCTCTCTA TAATTGAAGA AAAATAATTC	3480

	GGATGAACAA AACATGAGAA TAATGTTTAT AGGGGATATC GTAGGTAAAA TTGGACGAGA	3600
	CGCAATTGAA ACGTACATAC CTCAACTGAA GCAAAAAGTAT AAACCAACAG TTACAATTGT	3660
5	AAATGCTGAA AATGCAGCAC ATGGTAAAGG TTTGACTGAA AAAATATATA AACCAATTACT	3720
	AAGAAATGGT GTAGATTTCA TGAATATGGG TAATCACACA TATGGTCAAC GTGAAATTTA	3780
	TGATTTTATA GATGAAGCAA AACGACTAGT AAGACCAGCG AATTTTCCGG ATGAAGCGCC	3840
10	GGGAATTGGT ATGAGATTTA TACAAATTAA TGATATTAAA CTTGCAGTTA TTAATCTGCA	3900
	AGGAAGAGCG TTTATGCCAG ATATTGATGA TCCTTTTAAA AAGGCAGATC AATTAGTCAA	3960
	GGAAGCAGAA GAACAACTC CGTTTATATT TGTTGATTTT CATGCAGAAA CAACTTCTGA	4020
15	AAAGTATGCA ATGGGATGGC ATTAGATGG TAGASTAGCG CTGTTGTGG AACGCATACA	4080
	CACATTCAAA CAGCAGATGA ACGTATTTTA CCAAAGGGGA CAGGGTATAT AACGGATGTT	4140
	GGTATGACAG GTTTTTATGA TGGCATTTTA GGAATAAATA AAACAGAGGT AATTGAGCGT	4200
20	TTTATCACTA GTTTGCCACA AAGACATGTT GTTCCAAATG AAGGTAGAAG TGTATTATCT	4260
	GGTGTGTGTA TTGATTTAGA CAAAGAAGGT AAAACAAAGC ACATCGAAGC TATATTGATA	4320
25	AATGATGACC ATCCATTTTC AACATTTTAA AATTACGTAA GTAAACATTC GAATTGGACC	4380
	CTATCGTCCA TTAGTATGAA TTTAATATAG TACCACTGTT TACATAGTAA ATCGGTGGTT	4440
	CTTTTGTGTA TCATTTAATA TGAAATATAT CCATAGGAGG CATATACTA TGAAACCACA	4500
30	ATTATCGTGG AAAGTTGGCG GTCAACAAGG CGAAGGTATT GAATCAACTG GGGAAATCTT	4560
	CGCTACGGCT ATGAATAGAA AAGGATATTA TTTATATGGA TATAGACATT TTTCAAGTCG	4620
	TATCAAAGGT GGACATACGA ATAATAAAAT TAGAGTTTCT ACGACGCCTG TTCATGCAAT	4680
35	TAGTGATGAT TTAGATATTT TGATTGCATT TGACCAAGAA ACAATTGATG TTAACCATCA	4740
	TGAAATGAGA GAAGACAGTA TTATTTTATC TGATGCCAAG GCTAAACCTG TGAAACCAGA	4800
	AGGATGTCAT GCACAGCTTA TTGAATTACC TTTTACAGCA ACCGCTAAAG AATTAGGTAC	4860
40	AGCATTAAATG AAAACATGG TTGCAATAGG TGCTACTAGC GCATTGATGA ATTTGAATAC	4920
	AAATACATTT GAAGAACTTA TTACTAATAT GTTTTCTAAA AAAGGTGACA AGGTAGTTGA	4980
	AGTCAATATC CAAGCATTAA ACGAAGGTTA TCAATTAATG CAATCTCGCT TACCTGAAAT	5040
45	CTACGGGGAC TTTGAATTAG AGTCAACAGA TGCACTACCA CATCTATATA TGATTGGTAA	5100
	CGATGCCATT GGATTAGGTG CAATTGCTGC AGGTTCCAAA TTTATGGCGG CATATCCTAT	5160
	TACACCTGCG TCTGAAGTTA TCGAATATAT GATTGCCAAT ATATCTAAAG TAAACGGAGC	5220
50	GGTTATTCAA ACAGAAGATG AAATTGCTGC TGTAAGTATG GCTATTGGTG CAAATTATGG	5280

55

	TGGATTATCT	GGTATGACTG	AAACGCCATT	AGTCATTATT	AATACCCAAC	GAGGTGGACC	5400
5	TTCTACTGGA	TTACCTACGA	AACAAGAAAC	GTCAGATTTA	ATGCAAATGA	TTTATGGTAC	5460
	ACATGGTGAT	ATTCCAAAAA	TTGTTGTAGC	ACCAACAGAT	GCAGAAGATG	CATTTTATTT	5520
	AACTATGGAA	GCATTTAATT	TAGCAGAACA	ATATCAATGC	CCTGTTATAG	TTCTAAGTGA	5580
10	TTTGCAATTA	TCTTTAGGTA	AACAACTGT	TGAAAAATTA	GATTATAATC	GTATTGAAAT	5640
	TAAACGTGGT	GAAATCATTG	AATCTGATAT	TGAACGTGAA	GAAGATGATA	AAGGTTATTT	5700
	CAAGCGTTAT	GCGTtAACAT	CCGATGGTGT	TTCTCCTAGA	CCTATCCCCG	GTGTTAAAGG	5760
15	AGGTATTCAT	CATATAACTG	GTGTGGaaCa	CAATGAAGAA	GGTAAACCTA	GTGAATCTGC	5820
	GTCAAATAGA	CAACAACAAA	TGAAAAAAG	AATGCGTAAA	ATTGAGCAGT	TACTAATTGA	5880
20	ATCGCCAGTA	GAAGCTAACT	TACAACATGA	GGATGCAGAT	ATTCTTTTATA	TCGGTTTTAT	5940
	TTCTACAAAA	GGTGCAATTC	AAGAAGGTAG	TAACCGTTTG	AATCAACAAG	GCATAAAAGT	6000
	TAACACTATA	CAAATTAGAC	AATTGCATCC	ATTCCCAACA	AGCGTTATTC	AAGATGCAGT	6060
25	TAATAAAGCG	AAGAAAGTCG	TTGTAGTGGA	GCACAATTAT	CAAGGACAAT	TGGCTAGTAT	6120
	TATAAAAATG	AATGTCAATA	TTCATGATAA	GATTGAAAAT	TATACAAAGT	ATGATGGGAC	6180
	ACCTTTCTTA	CCACATGAAA	TCGAAGAAAA	AGGCAAAATA	ATTGCTACTG	AAATAAAGGA	6240
30	GATGGTATAG	ATGGCGACAT	TTAAAGATTT	TAGAAATAAT	GTTAAGCCTA	ACTGGTGCCC	6300
	CGGATGTGGC	GATTTCTCAG	TACAAGCTGC	AATTCAAAAA	GCAGCCGCAA	ATATAGGGTT	6360
	AGAACCTGAA	GAAGTAGCTA	TCATCACCGG	TATAGGATGT	TCTGGCCGTC	TTTCAGGATA	6420
35	TATTAATTCT	TATGGCGTTC	ATTCTATTCA	CGGACGTGCA	TTACCTTTAG	CTCAAGGTGT	6480
	AAAAATGGCG	AATAAAGATT	TAAGTGTAT	TGCATCGGGA	GGAGATGGTG	ATGGTTATGC	6540
40	TATAGGTATG	GGGCATACAA	TCCATGCTTT	AAGAAGAAAT	ATGAACATGA	CGTATATAGT	6600
	CATGGATAAT	CAAATTTATG	GTTTGACAAA	GGGACAAACA	TCGCCGTCAT	CAGCAGTAGG	6660
	ATTTGTTACT	AAAACAACGC	CAAAAGGTAA	TATAGAAAAA	AATGTTGCGC	CTTTAGAATT	6720
45	AGTATTATCA	TCTGGTGCCA	CATTTGTAGC	CCAAGGTTTT	TCAAGCGATA	TTAAAGGATT	6780
	AACAAAATA	ATTGAAGATG	CAATTAATCA	TGATGGATTT	TCATTCGTTA	ATGTCTTTTC	6840
50	ACCATGTGTG	ACTTATAATA	AAATTAACAC	ATACGATTGG	TTTaAAGAAC	ATTTAACAAG	6900
	TGTTGATGAC	ATTGAAAATT	ATGATTCTAC	AGATAAACAA	TTAGCGACTA	AAACTGTTAT	6960
	TGAACATGAA	TCTTTAGTAA	CTGGTATTGT	TTATCaAGAT	AAAGAAACAC	CATCATATGA	7020
55	ATCtCAAATT	AAAGAGTTAG	ATGATmCACC	ACTTGCTAAA	AGAGATATCa	AAATTaCTGA	7080

TGTATTTATA ACAGATCCAT TTATGCTACT CAGTTTTTTA CTATTACAAA AAATAAAGGA 7200
 GTTTTTAAAA ATGAAAGACA CATTAATGAG TATACAAATA ATTCCTAAAA CACCAAACAA 7260
 5 TGACAATGTT ATACCTTACG TAGACGAGGC GATTAAAATA ATTGACGAAT CTGGTTTGCA 7320
 TTTTAGAGTA GGTCCGTTAG AAACGACAGT ACAAGGAAAT ATGAATGAAT GTTTAATTTT 7380
 10 AATACAATCA TTAAATGAAC GAATGGTGGA ACTTGAATGT CCAAGTATTA TTAGCCAAGT 7440
 TAAGTTTTAT CATGTGCCAG ATGGCATCAC TATTGAACT TTAAGTAAA AATATGATGA 7500
 ATAACATTAA AAGTGAAGTA AACTGGATTT GAATTGGCTT GTTAGAGATG ACGTATAACT 7560
 15 TTAAGTGT TTGCACTTTA TAGTTAAATT TAATATAATT ATTAAATGAT ACGGGCAAAT 7620
 AGAAAGGATT TTGTAAAGTG AACGAAGAAC AAAGAAAAGC AAGTTCTGTA GATGTTTTAG 7680
 CTGAGAGAGA TAAGAAAGCA GAAAAAGATT ATAGTAAATA TTTTGAACAT GTTTATCAGC 7740
 20 CGCCTAATTT AAAAGCAAGC GCAAAAAAAG AGGTnAAA 7778

(2) INFORMATION FOR SEQ ID NO: 49:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1128 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

AGATGAAGTT GTTACgAAAA TTGCGTACGC TGTTTCAGAA CATGTCAAAA TAGAAACAGG 60
 35 TAATCCATTC TTTCAAACAT CACATAGTGG TTGTGCGACG GGCGGATCCT GTAATTGTTT 120
 ATTATAAAAA ACATCGAGTC AGAAAAAGGT GGTTATTGAA cCACTAACTA GCATCTGACT 180
 40 CGATGTTTTT ATTTATTCGG GATTGTTTGT TTGAATTGTT GTGCTAAATC TGGTCGATCT 240
 GTCACAATCG TGTGTGCACC TTTTGGTAT AAATCATTCA TCAGATTAT ACTATTTACG 300
 CCATAATAGC CTGGAATGAT ATTCATATCA TTTAACCATT TGATAAACG AGATGAAGTC 360
 45 AAATCAATGC CTTTAAATG AGTAGGCATT TGGAACGTTT GTGCTAATGG TTGGTAGTAC 420
 CTACCACCTA ATAAATGATA TTTTAAAAAT GCTTCTGTAA CTTCTGTG GCTAGCACCA 480
 ATTGCGACGG ATCCTTGTGC AATTTTATTA AAACGAACGA TTTGTTCTTT ATAAAACTT 540
 50 GTCACAAGAA CGCGGTCAAA TGCTTGATTT TCTGCAATTG TATCAAACAT AATTTGTGGT 600
 GCGATTGAGC CTTCATAGGA TTCAGGAGCA TCTTTAAGT CTACGTTTAT ATACATATCA 660
 55 GGATATTGCT TCAGCAACTc ATCGAAGGTT AGTATAGCTG TGTGTGCATG ACCACGATAT 720

AATGTATGGG CACTAACTTT TCCAGAGCCG TTCGTCGTTT TATCAACAGT TGGGTCATGA 840
 AAAACGATAA GCTGTTGATC TTTTGTGAGT CTCACATCTG TTTCAAAGCC ATCAACGCCT 900
 5 AATTGTTTTAG CATAGTCAAA TGCAAGTTGC GTTTGCTCTG GTCTTAAAGC CATACCACCG 960
 CGATGCGCAA ATATATATGG TGCATTGCCT TTGAAAAAAG CAGGGATGGT TTGCTTTTTA 1020
 10 GTAATCACTT TATTTTTATT GATCATTAAAT AGACTACTTA AAAATCCAGC ACCGACTAGT 1080
 ACCGCATTTA AAATGTTTCT GTTTACnTTT TTCATAAAAA ATTCCTCC 1128

(2) INFORMATION FOR SEQ ID NO: 50:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6252 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

25 CAAGCAAACA ATCGTCGATA AAATTGCTAA AATAATAAAA GTAATTCGAA CTTTCATCAT 60
 GATCATCCTT TGTTTATAGA GTCAATATAA GTATGGAATA TGTTAGGTAT ATAGTCAAAT 120
 GCGTCAACTA ATGGGAATTT TGGCATAGAT AGAGAATTTA AGGCAATTAA AAAGGCATCA 180
 30 AACAGTAATA TGCTGCTTGA TGCCCAAATG ATGACTTTAG CTAAATTGAT TAGTCACTTT 240
 TAAAGATAAA GAATTGTCAT GAATTAATAAC TCATGTAATG ATGTGTTACA TTTCGCAATG 300
 ATGGCTTTCA GTTATTTATC GATAACATCA CTCTTGATAC CTTTAGATTT TAAGAAATCT 360
 35 TTAATTTTAT CTGTGTTGCTT TTTATTAACA TCACCGGCAT ATTTTGTTGG CACGTCGACA 420
 ACATTGATTT TATTTTGCGG TTGATAGCTA AGCTTTTCAA TATCTTCATC AACATTGGCG 480
 40 ATTGTACTAT TTAAAGCTTT GAAGTAATTC ATCATTAAAT CAACGGGTTT CTTATATTCT 540
 TTAGGAATAT TGTTTTAGT GACAAATTC TTGAAATGCA AATCGTTTTT AACAGCTAAG 600
 TTAGATAAGT GGCTAAGTGT TTCTGCTTGT TTTTCAGTCA CTTTGTTTG ACTGTCAATT 660
 45 TGTTTATCTA GTTTATGTTG CATAATATAT TTGTTATCAA GTATATCGCT ATTTACAGAC 720
 AAATACTTTT CTATAGCTTG CTTCACTCTC GCATCACTAA TATCACTATT TTTCTTATCT 780
 GAGTTAAAGA TATCTTTTGT tTCTAATTTT TTAGCGCTTT TAGGTGCATG GATGCCAGTA 840
 50 CTTGTATGAT GATCTTCGTT ATCAGATTGA TCGGACGCGC AACCTGTAAG AATTAATGTC 900
 GATGCTAAAA ATGTACTTAG TAGTAATCTC TTTTTCATAA TGTAATATAA CTCCTTAGTT 960
 55 TATCTTTAAT TGAAAAAATA TGTATTCATG TTTAATAGAG TAACATTGAA TTAGTTTGGA 1020

	TCTATCAATA ATGCATCATT TTGGACGTTG TTAAGGATAG CTTTATCTAT AAATAACTGC	1140
	ATAATTGGTT GACTAATTT AGACGTAGGT ATCGTACGTA AAAGCATAAT AATTCGTTC	1200
5	ACATACTTTT CTTTCTCAAT ATCATTTTTC ATATTGATTT GTTTGCGAGA GGTACATACT	1260
	TTAAGCATT TCGCACATCT CGTTGTATAT ATTAAGTTTA TCATAACATG ATTTTATGTC	1320
10	GGGATAAAAA AATAACAGCA TCTTAACAAA TGTAAGATAC TGTCAGTGAA ATGAATGAAA	1380
	CTTTAGTTTC TGaTAATATA GTCAAAGGCA TTTAATGCTG CATTTGCACC AGCGCCCAT	1440
	GAAATGATAA TTTGTTTGGT CTTCTGATCT GTGACATCGC CAGCAGCAAA TATTCCAGGA	1500
15	ACATTCGTAT TATTGTTACG ATCAATCACA ATTTACCAC GTTCGTTTAA TTCAACAGCA	1560
	TCGTTTAACC ATGATGTGTT TGGAAGTAAA CCAATTTGAA CAAAGATACC ATCTAAGTTA	1620
	AGTAGATGTT CTTGCGCCGT GTTCATGTCT TCGTAACGTA TACCTGTAAC ATGGTCTTCT	1680
20	CCGACAACCT CAGTAGTTTT GGCATTTGTT TTGATATCAA CATTTGATAA AGAACGTAAA	1740
	CGATCTTGTA ACACGTTGTC TGCTTTTAAT TCGCTAGCGA ATTCGAATAA TGTAACATGA	1800
25	TTAACGATAC CAGCAAGGTC AATTGCTGCT TCAACCCAG AGTTACCGCC ACCGATAACT	1860
	GCTACGTCTT TATTTTCAA TAGAGGTCCG TCACAGTGAG GGCAGAATGC AACACCTTTA	1920
	TTAATCAATT GCTCTTCACC TGAATGTTT AGCTTACGCC AACCTGCACC AGTAGCAATA	1980
30	ATGACTGTTT TACTTTCTAA GACAGCACCG TTTTCTAACG TAACTTTAAT TGCTTCGTCA	2040
	GTCTTTTCGA TATCTGTAGC ACGTATACCT GTCATTGCAT CAATGTCATA TTGATCAATG	2100
	TGCGCTGCTA AGTTAGAAGA AAATTCAGAA CCAGTTGTTT CTTTAACAGT AATGAAGTTC	2160
35	TCAATACCAG CAGTATCATT AACTTGGCCA CCGATACGAT CAGCAACTAT ACCAGTACGT	2220
	AAACCTTTAC GTGCTGTGTA AATCGCTGCA CTACCACTAG CAGGACCACC ACCAACGATT	2280
	AAGACATCAT AAGGTTCTTT ATTTTCAAAC TCAGATGCAT CTGCCGTACT GCCTAGTTTC	2340
40	GAAAGAATAT CTTGGATTGT CATACGACCA TGCCAAATT CTTGCCATT TAAAAAGACA	2400
	GCAGGGACTG CCATGATGTT TTCAGATTCT TCACGGAACA CTGCACCATC AATCATAGAA	2460
45	TGCGTGATGT TAGGGTTGAT CACACTCATT AAGTTAAGTG CTTGAACGAC ATCAGGACAT	2520
	TTTTGACACG TTAAACTAAT GAATGTTTCA AAATGGAATG AACCTTCTAA TTTTCTAATT	2580
	TGGTCAATGA TTGACTGTTT TTCTTTAGGT GCACGACCAC TAACCTGTAA AATTGCTAAA	2640
50	ACAAGTGAGT TAAACTCGTG ACCTAATGGA ATACCTGCAA ATGTTACACC TGTTTCTTCG	2700
	CCAGGACGAT TGA CTGAGAA ACTTGGTGTA CGTTTTAAAG ATTTTTCAGA AAGAGATAGT	2760
55	CTAGGTGACA TATCAGTAAT TTCTGTCAAC AAATCTTTAA GTTCTTTGGA TTTATCATCT	2820

	TGTTGTTTTA AATCAGCATT AAGCATGGTT GTAATGCCTC CTTAGATTTT ACCTACTAAA	2940
5	TCTAAACCAG GTTGCAATGT TTTAGCGCCT TCTTCCCAT TAGCTGGGCA TACTTCGCCA	3000
	GGGTTTTTAC GAACATATTG AGCTGCTTTG ATTTTGTGAG CTAATGTACT AGCGTCACGG	3060
	CCAATTCGGT CAGCGTTAAT TTCAGATGCT TGTACAACAC CGTCTGGGTC GATAATGAAT	3120
10	GTACCACGTT GAGCTAAACC AGTAGCTTCA TCTAATACAT CAAAATTACG AGTGATTGTT	3180
	TGTGATGGGT CACCAATCAT AGTGTAAGTG ATTTTGCTAA TTGCATCTGA ATGGTCATGC	3240
	CATGCTTTGT GTACGAAGTG AGTATCAGTT GATACTGAGA ATACATTTTAC GCCTAATTTT	3300
15	TGTAATTCCT CATATTGGTT TTGTAAGTCT TCTAATTCAG TTGGACAAAC GAATGAGAAG	3360
	TCAGCAGGAT AGAAGCATAC TACGCTCCAA GAACCTTTTA AATCTTCTTG TGTAACCTCT	3420
	TTAAATTGAT CTTTTTTTGG ATCGAAATCT TCGCGCTGTA ATGGTAAGAT TTCTTTGTTA	3480
20	ATTAATGACA TAAATATCTT CCTCCTAAGA ATTTAAGTAT GAATTAGAAC TATCAATTGA	3540
	TTGCGCTTAA TTATAATAAT TCTAATCTCT TAGTTAGCAT TATTACATTT TGATCCAGAA	3600
25	TAGTCAACTG GATAACTTTG TAAAGTGAAT GATTACTTTT AAAATAAAGA AAGATAATAT	3660
	AAAGTGCTTT GATAATGGAT TTTGTAGTTG ATGATTTAAA AGGTTGTGTC TATATTTAAT	3720
	ATCTTGATTT TAATGTAAAA AATGTAAAAA AAGAAGATTT GTATTCTCAA CTAAGTCAAC	3780
30	CTTATTGATA ATGGTATGAG AATATTTGTT CGAGATGGAT GAAGGTAATG AGTGAGAAAC	3840
	TGGATTTTAA AAGTATGAGA CAATATTTTA AAAAGTTCAA TTATTAACCT ATAAGCAAAT	3900
	AATTGCTATA AAAAAGTTTG GACGTGTACA ATTGCAATAT GAAGATTTTA AATTAATTGT	3960
35	AAAGTATCGA GGAGTGGGTA ACGTGTGAGA ACATGTATAT AATCTTGTGA AAAAGCATCA	4020
	TTCTGTTAGA AAATTTAAGA ATAAACCTTT AAGTGAAGAC GTTGTTAAGA AATTGGTAGA	4080
40	AGCTGGACAA AGCGCTTCGA CGTCAAGTTT CCTGCAAGCA TACTCAATTA TTGGTATCGA	4140
	CGATGAGAAG ATTAAAGAAA ATTTACGAGA AGTTTCTGGA CAACCTTATG TTGTAGAAAA	4200
	TGGCTATTTA TTGTCCTTTG TTATTGATTA TTATCGTCAT CATTTAGTTG ATCAACATGC	4260
45	TGAAACTGAT ATGGAAAATG CATATGGTTC AACGGAAGGT TTGCTAGTAG GTGCAATCGA	4320
	TGCAGCATTG GTTGCCGAAA ATATTGCGGT AACTGCTGAA GATATGGGGT ATGGCATTGT	4380
	CTTTTATAGG TCATTAAGAA ATGATGTTGA ACGCGTTTGA GAAATTTTAG ACTTACCTGA	4440
50	CTATGTCTTC CCGGTATTTG GTATGGCAGT AGGGGAACCC GCAGATGACG AAAATGGTGC	4500
	AGCCAAGCCA CGCTTACCAT TTGACCATGT CTTCCATCAT AATAAGTATC ATGCTGATAA	4560
55	GGAAACACAG TATGCACAAA TGGCAGATTA CGACCAGACA ATCAGCGAGT ACTATGATCA	4620

	CAAAGCAAGA TTAGATATGT TAGAACAATT GCAAAAATCA GGCTTAATAC AGCGATAgCA	4740
	AGATACCAAA ATAACCCGCC CCCCTCTAGC TTAAATGAT AAGTATAGCT AGAGGGGGCG	4800
5	GGTATTTCTT GCAATGAATT AGTGTGAAGT TAATGCAGCA TTATCATTTG AATCGAAAGT	4860
	ATCTTTATCC CAATGTTTAG TTAACCTGGC GGTACCTGTA CCAGCTAGCA TTGAATCGTT	4920
10	CACGTTTAAT GCTGTTCTAC CCATGTCAAT CAATGGTTCA ACGGAGATGA GCACGCCGGc	4980
	TAAAGCGACT GGCAAGTTTA ACGTTGACAA CACCAATATG GATGCAAATG TAGCCCCGCC	5040
	ACCGACGCCA GCAACGCCGA ATGAACTAAT AATCAGGACA GCGATTAACG TTACAATAAA	5100
15	TTGTAAATCA ATTTCTACAT TAGCGACGGG TGCGACCATA ATTGCAAGCA TGGCAGGGTA	5160
	AATGCCTGCA CAACCATTTT GTCCAATCGA CAATCCAAAT GTCGCAGCGA AATTGGCAAT	5220
	ACCTTCTGGC ACGCCTAGAC GTCTTGTTTG TGTTTGTACA TTCAATGGTA AGGCACCCGC	5280
20	GCTTGAGCGT GATGTGAATG CAAAGATTAA TACTTCCAAA GTCTTTTAA CATAGCGAAT	5340
	TGGGCTAATA CCTAACAGGC TTAAATAAT TAAGTGAATG ATATACATCG TAATTAATGC	5400
	AGCGTACGAT GCGATTAAGA ATTTTCCTAA AGTCCAAATG GCGCCAAAGT CACTTGTCGA	5460
25	TAATGTGTTG GCCATAATTG CTAATACACC GTATGGCGTT AAACGTAAGA CGAACGTCAC	5520
	AATCGCCATT ACTAGTGAAT AGATAGCGTC AATCGCACGC TTAAGCAATT CACCATGATC	5580
30	AGGTTGTTTG CGTnTACGCG TAAATAAGCA AATCCTATAA ACGAAGCAAA TATCAGGACA	5640
	GCAATCGTGG aAGTTGCACG TTGTCCaGTG AAATCTAAGA ATGGATTTTT AGGCAATAAT	5700
	TCCAAAATTT GTTGTGGTAA CGTATGTGCT GTTAAATCTT TCGCTTGTTT AGCAATTTTCG	5760
35	CTTCCACGTG CTTGTTcAGC GTTACCAAGG TTAATTGTTG ATGCATCTAA ACCAAACACC	5820
	AAGGCATACA CAACACCAAC AATCGCAGCA ATGGTGACAG TGCCAATTAA AAAGATAAAA	5880
	ATGASACTAC CAATTTTAGC AAACTTTTCT CCGATTTGAA TTTTAGTGAA TGCAGCTACA	5940
40	ATAGAAATGA AAATTAAAGG CATAACAATC ATTTGCAACA ATGCAACGTA ACCTTGTCGG	6000
	ACAATGTTGA ACCAGTCACT TGTTGATGTA ATAACATTcG AATGTGTGCC ATAAATAAGA	6060
45	TGCAATAACA CACCGAATAC TATACCAATC CCTAAAGCTG TAAACACACG TTTCGCAAAA	6120
	GATATATGTT TGCGAGCCAT CATGTGCAAT ATTACGATGA AAATCACCAA TACAATAATA	6180
	TTAATCAGTG TAAGAAAAGC ATTCATGAAC GTCACTCCTT AAATTTTTGA ATATAATTCC	6240
50	GACTAGTATG CT	6252

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6730 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

5	ATCAAATCnC AAAATATTTA TTAATnAnAA GGGGATTATC CaTGtGAGAA ACAAAGTAAT	60
10	GCTCTTTTTT TACCTCTTGT GGGTTGAAAA aTGGATCATC AGAGATAGAC TTCTTCTTTT	120
	TCGAAGATGA CATTTGATAC TTTAATCTTC TAAAACCATA ACTTGTCGCA TCAAAAATGC	180
	CTTCTTGTA C AAGTAAAATC AAAAATATGC TAATAAAAAT AATTAATGAA ACATAAAACA	240
15	ATATATTTAA ATATGTAATG ATAGTATGGC TATTAAAAAG CCATATAATA AACGTTAATA	300
	TTGGCGTTAT TAGTGCCATT CCAAGCCATT TTTTCAACAT TTGATCACTC CCACTTATAG	360
	AAAACCTCTTA GGCATAGTTT ACATTAAAAT CAGACATTGA GGAATGATTT TTTAATTTCT	420
20	TCAGCTTTAT TGAAATTCTA AAATCAATCA TTCTTCATTA GTTTAAAGCA AAAAAATATT	480
	GATATATAGT AAATATTGTA TATATAATAT TAGTTAAGAT TTCaGAAAAT TTTGAAGGGA	540
25	ATGGAAATTT AGAAATCGGA ATTTGTTAGA GGAGGGGATT AGATGGGGAA ATATATTTTC	600
	AAACGATTTA TTTATATGCT TATTTCTTTA TTTATTATTA TTACAATTAC ATTTTCTTA	660
	ATGAAATTAA TGCCAGGTC GCCATTTAAC GATGCTAAAT TAAATGCTGA ACAAAAAGAA	720
30	ATTTTAAATG AAAAAATATG ATTAAATGAT CCTGtAGCTA CGCagTATTT ACATTATTTA	780
	AAAAATGTTG TTACAGGCGA TTTTGGAAT TCATTCCAGT ATCATAATCA ACCTGTGTGG	840
	GATTTGATTA AACCGAGACT ACTACCTTCT TTTGAAATGG GTCTTACAGC AATGTTCaTC	900
35	GGTGTGATAC TGGGACTTAT TTTAGGTGTT GCAGCAGCTA CTAAACAAAA TTCTTGGGTT	960
	GACTATACAA CTACAGTTAT TTCAGTTATT GCAGTATCTG TACCATCTTT TGTACTTGCT	1020
40	GTACTTTTAC AATATGTATT TGCAGTTAAA TTAAGATGGT TCCCAGTAGC TGGATGGGAA	1080
	GGTTTTTCGA CCGCGGTATT ACCGTCACCT GCATTATCTG CAGCTGTTTT AGCAACTGTC	1140
	GCCAGATACA TAAGAGCAGA GATGATAGAG GTATTAAGTT CAGACTATAT TTTATTAGCG	1200
45	AGAGCTAAAG GTAATTCGAC AATGCGTGTA CTTTTTGGAC ATGCACTTAG AAATGCTTTA	1260
	ATTCCAATTA TTACAATTAT CGTTCCCATG TTAGCAAGTA TTTAACAGG CACTTTAACA	1320
	ATTGAAAATA TTTTGGAGT TCCTGGATTA GGGGATCAAT TCGTACGTTT AATTACAACA	1380
50	AATGATTTCT CAGTAATCAT GGCAATCACA CTATTATTTA GCACACTGTT TATCGTTTCT	1440
	ATTTTATTG TAGATATTTT GTACGGTGTG ATAGATCCAC GAATTCGTGT TCcAAGgAGG	1500
55	TAAAAAATAA TGGCTGAAAA TAAAAACAAT TTGTCGATTA ACGACGATCA TTCTAATGCA	1560

	TGAATCAGGA ACCTGAAATG CAACGAGAAA GCAAAAACCTT TTGGCAAGAT GCTTGGGCTC	1680
5	AGTTAAAACG AAATAAGTTA GCTGTTGTCG GTATGATAGG TTTAATTATC ATTGTAATAT	1740
	TTGCTTTTAT CGGTCCAGTT ATAAATAAAC ATGATTATGC TGAACAAAAT GTAGAACATA	1800
	GAAATCTTCC GGCAAAAATA CCTGTATTAG ACAAAGTTCC ATTTTTACCT TTTGATGGTA	1860
10	AAGATGCAGA TGGCAAGGAT GCTTATAAAG CAGCAAATGC TAAAGAAAAT TATTGGTTTG	1920
	GTACTGATCA GTTGGGTCGA GATTTATGGA CAAGAACATG GAAAGGTGCT CAAATTTTCAT	1980
	TGTTTATCGG TGTTGTGCA GCGATGTTAG ATATTTTTAT TGGTGTGTA TATGGTGCGA	2040
15	TTTCTGGATT CTTCCGTGGA CGTGTGATA CGATTATGCA ACGTATACTT GAAGTCATAG	2100
	CATCTATTCC GAATTTAATT GTCGTAATTT TATTTGTATT AATTTTTGAA CCATCCATTT	2160
	GGACAATTAT ATTGGCTATG TCTATCACAG GCTGGTTAGG CATGAGCAGA GTTGTACGTG	2220
20	GAGAATTTTT AAAATTAAAA AATCAAGAGT TTGTCATGGC TTCGAAAACA TTGGGGGCTT	2280
	CAAAATTCAA ATTGATATTT AAGCATATTT TACCTAATAC ATTAGGTGCT ATCGTGGTTA	2340
25	CATCAATGTT TACAGTACCT AGTGCTATTT TCTTCGAAGC ATTTTAAAGT TTCATTGGTA	2400
	TAGGTGTACC CGCACCTCAA ACATCGTTAG GGTCAATAGT AAATGATGGG CGCGCAATGT	2460
	TATTAATTTA TCCACATGAA TTATTTATAC CAGCAATGAT TTTAAGTTTA TTAATTCTAT	2520
30	TCTTTTACTT ATTTAGTGAT GGATTACGTG ATGCATTGA TCCGAAAATG CGTAAATAAA	2580
	AAGGGGGCAT AGCATATGAC TGAAAGAATA TTAGAAGTAA ATGATTTGCA TGTTTCCTTT	2640
	GATATTACAG CAGGGGAAGT GCAGGCAGTG AGAGGCGTAG ATTTTTATTT GAACAAAGGG	2700
35	GAAACATTGG CAATTGTTGG TGAATCAGGT TCAGGTAAAT CTGTAACAAC AAAAGCAATT	2760
	ACAAAATTAT TCCAAGGGGA CACAGGAAGA ATTAAAAAGG GAGAAATTTT ATTTTTAGGG	2820
	GAAGATTAG CAAAAAACC TGAAATGAG TTGATTAAAT TACGTGGCAA AGATATTTCA	2880
40	ATGATCTTTC AAGATCCAAT GACATCTTTA AACCCAACGA TGCAAATTGG TAAACAAGTC	2940
	ATGGAACCAT TAATTAAGCA CAAAAATTAT AGTAAAGCAC AAGCTAAAAA GCGCGCATTG	3000
45	GAAATACTAA ATCTGTAGG TTTACCAAAT GCAGAAAAAA GATTTAAAGC ATATCCTCAT	3060
	CAATTTTCAG GTGGACAAAG GCAAAGAATT GTTATTGCAA CCGCATTAGC TTGTGAACCT	3120
	AAAGTGCTCA TTGCTGATGA ACCAACGACT GCATTAGACG TAACGATGCA GGCACAAATT	3180
50	TTAGATTTAA TGAAAGAACT ACAACAAAAA ATCGATACAG CAATTATTTT TATAACGCAT	3240
	GATTTAGGGG TTGTTGCGAA TATTGCTGAT AGAGTGGCAG TTATGTATGG TGGTCAAATG	3300
55	GTTGAAACAG GAGATGTTAA CGAAATATTT TATGATCCAA AGCATCCATA TACATGGGGA	3360

	GGAGCGCCAC CTGATTTATT ACACCCACCT AAAGGTGATG CATTTGCGAG ACGTAGCAAT	3480
5	ATGCATTAGA TATTGATTTT AAAGTAGAAC CACCGTGGTT TAAAGTTTCA CCGACACATT	3540
	TTGTGAAATC TTGGTTATTA GACGCACGTG CACCAAAAGT TGAACCTACC GAGCTGGTAA	3600
	AACAACGTAT GAAACCGATG CCTAATAATT ATGAAAAACC ACTCAAGGTA GAAAGGGTGT	3660
10	CGTTCAATGA AAAATGATGA AGTGCTATTA TCTATTAAAA ATTTAAAGCA ATATTTTAAC	3720
	GCAGGAAAGA AAAACGAAGT GgaGCGATTG AAAATATTTT GTTTGATATA TACAAAGGGG	3780
	AAACATTAGG TTTAGTAGGA GAATCGGGGT GTGGTAAATC TACAACCTGGT AAATCAATTA	3840
15	TTAAACTTAA TGATATTACA AGTGGAGAAA TTTTGTATGA GGGTATTGAT ATACAAAAGA	3900
	TTCGTAAACG TAAAGATTTG CTTAAATTTA ATAAAAAGAT ACAGATGATT TTTCAAGACC	3960
	CATATGCGTC TTTAAATCCT AGGTTAAAAG TAATGGATAT AGTAGCTGAA GGTATTGATA	4020
20	TCCATCATTT AGCAACTGaT AAGCGTGACC GAAAAAACG TGTCTATGaT TTACTTGaAA	4080
	CTGTTGGATT AAGTAAAGAA CATGCCAATC GCTATCCTCA TGAATTTTCA GGTGGaCAAC	4140
25	GCCAACGTAT TGGaATTGCC CGTGcATTAG CCGTTGaACC AGAATTCATT ATCGCGGACC	4200
	AACCAATATC GGCATTGGAT GTTTCAATCC AAGCTCAAGT AGTTAATTTA TTATTAAAT	4260
	TACAACGTGA AAGAGGGATT ACGTTCCTAT TTATAGCTCA TGATCTATCA ATGGTGAAGT	4320
30	ATATTTTCAGA TCGTATTGCA GTCATGCATT TTGGGAAAAT AGTTGAAATT GGACCGGCAG	4380
	AAGAAATTTA TCAAAATCCA TTACACGATT ATACTAAGTC TTTATTATCA GCCATTCCAC	4440
	AACCTGATCC TGAATCAGAA CGCAGTCGCA AACGATTTAG TTATATTGAT GATGAAGCAA	4500
35	ATAATCATTT AAGACAATTA CATGAAATTA GACCGAATCA CTTTGTCTTT AGTACTGAAG	4560
	AAGAAGCGGC ACAACTACGA GAAAATAAAT TGGTGACACA AAATTAAGGG GAAGGGGGAA	4620
40	ATGCAATGAC GAGAAAATTT AGAACACTTA TTTTAATTTT GATTGCTACA ATTGCATTAA	4680
	GTGTTGTGTC TAATGACGAT GGTATTTATT CAGATAAAGG TCAAGTATTC AGAAAAATTT	4740
	TGTCATCAGA CTTAACATCC CTTGATACAT CATTATAAAC GGATGAAATA TCTTCTGAAG	4800
45	TGactGCGCA AACATTGAA GGTTTATACA CATTAGGAAA AGGTGACAAA CCGGTGTTAG	4860
	GTGTTGCGAA AGCTTTTCCT GAAAAGAGTA AAGATGGTAA AACTTTAAAG GTTAAATTAA	4920
	GAAGCGATGC TAAATGGAGC AATGGTGACA AAGTGACTGC ACAAGACTTT GTTTATGCTT	4980
50	GGAGAAAAAC AGTTGACCCT AAAACAGGTT CTGAATTTGC ATACATTATG GGGGACATTA	5040
	AAAAATGCGAG TGATATTAGT ACTGTAAGA AACCTGTAGA GCAATTAGGT ATCAAAGCAT	5100
55	TAAATGATGA AACATTACAA ATTGAATTAG AAAAGCCGGT TCCATATATT AATCAATTAT	5160

ACGGTACGGC AGCTGATAGA GCGGTATACA ATGGTCCaTT TAAAGTTGAT GATTGGAAAC 5280
 AAGAAGATAA AACCTTACTA TCTAAAAATC AGTATTATTG GGATAAAAAG AATGTAAAAT 5340
 5 TAGATAAAGT GAATTATAAA GTTATTAAAG ACTTACAAGC CGGTGCATCA TTGTATGATA 5400
 CTGAATCAGT AGATGACGCA TTTATTACTG CAGATCAAGT AAATAAATAT AAAGACAACA 5460
 10 AAGGATTAAA CTTTGTGTTA ACGACTGGGA CATTTTTTTGT AAAAATGAAT GAAAAACAAT 5520
 ATCCTGATTT TAAAAACAAA AATTTAAGAT TGsTATCGCA CAAGCAATAG ATAAAAAAGG 5580
 ATACGTTGAT TCAGTGAAAA ACAATGGCTC AATTCCCTTC GATACACTAA CAGCCAAAGG 5640
 15 AATTGCGAAA GCGCCTAATG GCAAAGATTA TGCGAGTACC ATGAATTTCGC CTTTAAAATA 5700
 TAATCCTAAA GAAGCAAGAG CACACTGGGA CAAAGCTAAA AAAGAGTTAG GTAAAAATGA 5760
 AGTGACATTT TCAATGAACA CAGAAGATAC ACCAGATGCA AAAATATCTG CTGAATATAT 5820
 20 CAAATCGCAA GTTGAGAAAA ATTTACCAGG AGTTACTTTG AAAATTAAGC AATTACCGTT 5880
 TAAACAAAGA GTATCACTAG AACTGAGTAA CAATTTTGAA GCATCACTTA GTGGTTGGTC 5940
 25 TGCAGATTAC CCTGATCCTA TGGCTTATTT AGAAACAATG ACCACAGGTA GCGCACAAAA 6000
 TAATACAGAC TGGGGTAATA AAGAATATGA TCAATTACTT AAAGTAGCAA GAACCAAATT 6060
 GGCACCTTCAA CCGAACGAAC GATATGAAAA CTTGAAAAAA GCAGAAGAAA TGTTCCTAGG 6120
 30 AGATGCACCG GTAGCACCAA TTTATCAAAA AGGTGTTGCA CATTTaACAA aTCCTCAAGT 6180
 AAAAGGATTA ATTtACCATA AATTTGGTCC AAATAACTCA CTTAAACATG TATATATTGA 6240
 TAAATCGATA GATAAGAAAA CAGGTAAGAA GAAAAAATAA TATGCTTTGT AAATTAGGCT 6300
 35 GGAGACATAT CTCCAGTCTT TTTGTGTTGG ATAAAAaCTT TGGGAATAAA AATTTAAAAT 6360
 AAGTCGTTTT TTAAATTACT GAAATTGATT AAATGCATAA ATAAGTGAAT ATTCTAAAAA 6420
 TAAaCTTGTA ATAATTTTTT CTATGAGTAA ACTAAAAAGA AAAAATTAGA TTGAAAGTAG 6480
 40 GAGGCATATG TATGGGGAAG CTAATTAAAT ATATTTCAAT ACTTCTTATT GTCGTTTTAG 6540
 TGTTGAGTGC TTGCGGAAAA AGCAGTAATA AAGATGAAGG AGTAAAAGAT GCTACTAAAA 6600
 45 CGGAAACCTC AAAACATAAA GGTGGTACCT TAAATGTAGC ATTAACAGCA CCGCCAAGTG 6660
 GTGTTTTATTC TTCGTTATTA AATAGTACAC ATGCAGATTC TGTAGTTGAG GGATATTTTA 6720
 ACGAAAGCTT 6730

50 (2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6482 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

5	AATTTTGTGTC ATTATTAAAA ACCTCGCTTT TAAAAGATTG AAAAGTAAAT GAGTGAAATT	60
	AAAGATTATG CACATTAAAA TCACGCCACA ATTTAATTGT GAAAAATATC ACAAATATAT	120
	TATAACACTA AATTTCCCAA AATTCAAAAG TGTGTTTTAT TGCAGAAAAC TTATAACaYg	180
10	TGCACAAGTT ATAGTGAATT GCAAACGGAT TACTTTAGTC TTTTAAAAAC ATGAAGTATA	240
	ATTTGTATAG CAATAAATAT AAAAATGGGA GGCTATGTTT AATGAGCAAT ATGAATCAAA	300
	CAATTATGGA TGCATTTTCAT TTCAGACATG CGACTAAGCA ATTCGATCCA CAAAAGAAAG	360
15	TTTCGAAAGA AGATTTTGAA ACAATATTAG AGTCAGGTAG ATTGTCTCCA AGTTCTCTTG	420
	GGTTAGAACC TTGGAAGTTT GTCGTGATTC AAGATCAAGC GTTACGTGAT GAATTAAAAG	480
	CGCACAGTTG GGGCGCAGCA AAACAATTAG ATACAGCGAG CCATTTTGTG CTAATTTTGTG	540
20	CGCGTAAAAA TGTAACGTCA AGATCACCGT ATGTACAACA TATGTTAAGA GATATTAAAA	600
	AATATGAGGC ACAAACGATT CCAGCTGTTG AACAAAAATT CGATGCATTC CAAGCAGATT	660
25	TCCATATTTT TGATAATGAT CAAGCCTTGT ATGACTGGTC AAGTAAACAA ACGTATATTG	720
	CATTAGGCAA TATGATGACG ACAGCCGCAT TGTTAGGTAT TGATTCATGT CCGATGGAAG	780
	GTTTTAGTCT GGATACAGTG ACAGACATTT TAGCAAATAA AGGGATCTTA GATACTGAGC	840
30	AATTTGGTTT ATCAGTGATG GTCGCATTTG GCTACAGACA ACAAGAGCCA CCGAAAAATA	900
	AAACACGCCA AGCTTATGAA GATGTTATTG AATGGGTTGG ACCAAAAGAA TAAATAGAAT	960
	ACCGTATGTC TAAATATATA AAATTAAAAA GTTAGCAATA AAAAAGCCTG CGATTACATA	1020
35	AATGAATCGC AGGCTTTTGC GTGAAAAAAT TGTATTAATA AAGTATGGAT GATTATTTTT	1080
	CTGGACAAG GTCAGTATTT GAATGAACTG TGATGTCAA CCCTTCTGGT GCCGTAAATG	1140
40	TATGTGTTGA GGCCTCGGGT TGATAAATAT CAACATGTGT TAATCCATAA CTTTGTGAAT	1200
	TGTTTTGTCT TGCTTGATTG GATTGCCAAG TATTAGCAGC AATATGATGG TGATAATGAT	1260
	TCGTTGACAT AAATAGCGCA CGTGAAAAT CAGACACATG TTGGAATCCT AATTGTTCAA	1320
45	TGTAACATTG ATATGCTGCG TCTAAATCAT GTGTTTTTAA ATGTAAGTGT CCAATCATGC	1380
	CTTTTGCTGG CATTCTTGC CAACCTTCAT CAGTACGATG TGTTAATAAG GTTTGGCTAT	1440
	CAACTTCTAA AGTATCCATT TTAACCTTGC CATTTTGCCA TTCCCATGAA GATGAAGGTC	1500
50	TATCGCGATA GACTTCAATA CCATTACCTT CGGGGTCGTT GAAATATAAA GCTTCACTTA	1560
	CTAAATGATC ACCAGCGCCG ATGCCCATAT TTTTTTGTGC CACGAAATAT AAGAAGTTAG	1620

55

	aAGTCTGACG GcCGTCTTCT AATAAATGTA ACGTTAGAGT ATGGcCACCA GTCCCAACAG	1740
	ATAATACGGT TGTATTATCG TCAGAACTTT TAACGGATAG TCCTAAAATG TTTTGTAAA	1800
5	ATGTTGTcAT TAAGTCTAAG TCTCTTACGT TCAGTACAAT GTTTGTCACT TGTGTTGCTG	1860
	TTTTATCGTG AAATGCCATT ATGCATCGCC TCTTTTCTA TTTTCTATA AGTTAGTATA	1920
10	AAAAGTATAC CAGAAAAGAA AATGAATTGA TAGCATAAAG TTTGAAATGC AAAATAACTA	1980
	GTCGTTTTGC AATTTTATcAT TGATGCGAAC AAAAAAGCGA TGGTACAGTT GCACCATCGC	2040
	AAAATTTATT TAACCAAGAT ATACATCTTG ATATGAATCT TCTTTTTCTA ACATATGTTT	2100
15	GGCAAAATGAA CATGAGGCAA TAATTTTCAA ATTATTTTCT CGAGCGTGTT CAACAACGTGc	2160
	TTTAAGTAGT TTTTGGCCAA CACCTTGACC ACCAAGTTCA TCAGATACGC CTGTATGATC	2220
	AATGTTAATT TCATTATTAT CCACAAAACG GTATGTGATT TCAGCTAAAG CATTATTTTC	2280
20	ATCATCACCA ATATAGAATT TGTCTCGCC TTGTTTGATT TCAAGGTTAC TCATACATAT	2340
	CAACTCCTAT CATGATTGAT TATAGTATTT CCCTATTCTA TTTTAACTTA AACGAAGTCA	2400
	AAGGTGCATG ACAGTCATGT GACGACATTG CCACATCTAT GTAGTCGTTT TTATTAAGCA	2460
25	CAGTTTGAAA TGAAGATGAA AACACGTATC TTGACATTAA ATCTATTCAG CTATATAATT	2520
	TATCTCGAAA TCGAAATAAA ATAAAAAGT TGGTGATCAT ATGGATCGAA CGAAACAATC	2580
30	TCTCAATGTT TTTGTGGAA TGAATAGGGC GTTAGACACA TTAGAGCAA TTACAAAAGA	2640
	AGACGTAAAG CGATATGGCT TAAATATTAC TGAATTTGCA GTGCTCGAGT TGCTTTATAA	2700
	TAAAGGTCCG CAACCAATTC AACGTATTAG AGACCGCGTA TTAATTGCAA GTAGCAGCAT	2760
35	TTCATATGTT GTAAGTCAAT TAGAGGACAA AGGTTGGATT ACACGTGAAA AGGATAAAGA	2820
	TGATAAACGT GTATATATGG CTGTTTAAAC TGAAAAAGGT CAAAGTCAAA TGGCAGATAT	2880
	TTTcCCTAAG CATGCTGAGA CATTAAACAA AGCGTTTGAT GTGTTAACAA AGGATGAATT	2940
40	AACAATCTTA CAACAAGCGT TTAAGAACT AAGTGCACAA TCTACAGAAG TGTAAGGCGT	3000
	GCACTAAAAA TTTACATTAA AGTATCTCGA TTTCGAGATA AATGCACTAA AAATATAAAG	3060
45	AGGGTATATA AAATGATAAA TAATCATGAA TTACTAGGTA TTCACCATGT TACTGCAATG	3120
	ACAGATGATG CAGAACGTAA TTATAAATTT TTTACAGAAG TACTAGGCAT GCGTTTAGTT	3180
	AAAAAGACAG TCAATCAAGA TGATATTTAT ACGTATCATA CTTTTTTTGC AGATGATGTA	3240
50	GGTTCGGCAG GTACAGACAT GACGTTCTTT GATTTTCCAA ATATTACAAA AGGGCAGGCA	3300
	GGAACAAATT CCATTACAAG ACCGTCTTTT AGAGTGCCTA ACGATGACGC ATTAACATAT	3360
55	TATGAACAGC GCTTTGATGA GTTTGGTGTT AAACACGAAG GTATTCAAGA ATTATTTGGT	3420

	TTAAATGAAG GGGTAGCACC TGGTGTACCT TGGAAGAATG GACCGGTTCC AGTAGATAAA	3540
	GCGATTTATG GATTAGGCCC CATTGAAATT AAAGTAAGTT ATTTTGACGA CTTTAAAAAT	3600
5	ATTTTAGAGA CTGTTTACGG TATGACAACT ATTGCGCATG AAGATAATGT CGCATTACTT	3660
	GAAGTTGGCG AAGGAGGCAA TGGTGGCCAG GTAATCTTAA TAAAAGATGA TAAAGGGCCa	3720
10	GCaGCACGTC AAGGTTATGG tGAGGTACAT CATGTGTCAT TTCGTGTGAA AGATCATGAT	3780
	GCAATAGAAG CGTGGGCAAC GAAATATAAA GAGGTAGGTA TTAATAACTC AGGCATCGTT	3840
	AATCGTTTCT ATTTTGAAGC ATTATATGCA CGTGTGGGGC ATATTTTAAT AGAAATTTCA	3900
15	ACAGATGGAC CAGGATTTAT GGAAGATGAA CCTTATGAAA CATTAGGCGA AGGGTTATCC	3960
	TTACCACCAT TTTTAGAAAA TAAAAGAGAA TATATTGAAT CGGAAGTTAG ACCTTTTAAT	4020
	ACGAAGCGTC AACATGGTTA ATTGGAATGA GGAGGATTTG TGATGGAACA TATTTTTAGA	4080
20	GAAGGACAAA ATGGTGCGCC AACACTAATA TTATTGCATG GTACAGGTGG TGATGAGTTC	4140
	GATTTATTAC CGTTAGGCGA AgcATTGAAT GAAAATTATC ACTTGTTAAG TATTAGAGGA	4200
	CAAGTTTCAG AAAATGGGAT GAACCGTTAT TTCAAACGTC TTGGTGAAGG TGTTTATGAT	4260
25	GAAGAAGATT TGGCATTTCG TGGACAAGAA TTGTTGACGT TCATTAAAGA AGCTGCTGaA	4320
	CGTTATGATT TTGaTATTGA AAAAGCAGTA CTTGTTGGAT TTTCAAATGG ATCAAATATA	4380
30	GCGATTAACT TAATGTTGCG TTCAGAAGCA CCATTTAAAA AAGCATTGTT ATATGCACCG	4440
	TTATACCCAG TTGAAGTAAC GTCAACAAAG GATTTATCAG ATGTCAGTGT GTTGCTTTCT	4500
	ATGGGGAAAC ATGATCCAAT TGTGCCATTA GCTGCAAGTG AACAGTCAT TAACTTGTTT	4560
35	AATACACGTG GGGCACAAGT CGAAGAAGTT TGGGTGAAGG GCCATGAAAT TACAGAACT	4620
	GGATTAAACG CTGGTCAACA AATACTTGGG AAATAACAGT TCTATTAAGA AGCGGACAGA	4680
	TGGAaAAGAT TTTTACTTTT CATCTGCCCG CTTTTTTGAT TTTGAAGTGC TGTAATAAT	4740
40	TTTACAATAG TATAGATATT TTAATCGATA TGAGATTGTC CGGTAATACG CTTAATTAAA	4800
	CCTTTATAGA GTACAGGTAT GAGTAAGATG AAACCGAACA ATCCCATAAT AGGGAATACT	4860
	TTTCCAATTA ATGAAATGaA ACCGATAAAT GACTAATAT AAGTGATGAC AGCCATTGTA	4920
45	ATAATAATGA TGAAGTAACG TCTGCTGAAT GGAACGCTGA AACGTGACGC AAATGCATAC	4980
	ATTAATCCAA CAACAGTATT GTAGATGACA AGTATCATAA TGACAGACAT AATAATACCA	5040
50	ATTGACGGAG ACATTTGTGT CGCTAATTTT AATGTAGGTA GATCTACGTG TTTAATTTTA	5100
	TCGAATTGAG AAATTAAACC TAGATTAATC ATCATGAGTA AAAATGTAAT GATTAAACCG	5160
	CCAATCAAGC CCCCCTATAA CGTTGAGTCA CGATATTTAA CTTTACTACC CATCACTGAT	5220

55

CCAGGTGATA ATGATTTCTG CTTATGAATC TGAGCATCAT TATTAGCGGC AGTAAAATCA 5340
 AGATGACTTG TTGTGAAATA GTAGACCGCA ATCATAATGA CAATCGCAAT TAAAAATGGG 5400
 5 GTAACACCGC CAAGCACAGC AATTAAACGA TCGAATTTTA GAAACAGTGT TGCTAAAATA 5460
 AAGGCGACTA ATATGAGTGC GCTCAGCCAA TACGGTAAGT TGAAACTTTG ATGAATGGTT 5520
 GACGCACCAC CTGCAGTCAT AATAATAGCT AAAGACAACA TAAACATTGT TAAAAATA 5580
 10 TCAAAACCTC TTGCAATAGA GGGGTATAAG AAATAGTTAA TTGAATCAGA ATGATTTCTG 5640
 GACTTTAGAT GATGACCTGT ATGCATGACA ACCATTCCAC CTAAAGTAAT CAATAGTCCT 5700
 15 GTTACAATAA TGCCTGAAAT GCTATATGCG CCATGACTTG TGAAAAACTG GAAAATTTCT 5760
 TGACCAAGTAG CAAAGCCGGC ACCAACGACA ACACCAACAA AGGCAAATGC CACAATAATG 5820
 GACTCTTTTA AGATACGCAT GATTTAAAAA TGTCCTTCG TAATTTTAAG TAATATAGAA 5880
 20 AATGTAACAT ACATGTTAAT GAAAAATATA GTACTAATAT AGTATTTTGT TAAATTGGAG 5940
 TAGAAGCGAG GGTGTCGGTC ATTCATTAA TTTATTAGTT GATTTTGCAT TTTTGTCTG 6000
 TAAAGTTGTT ATAATACAGT TAACAGGAAT TAGCATAGAT ACACCAATCC CCTCACTACT 6060
 25 CGCAATAGTG AGGGGATTTT TTTCGGTGTA GCTAGGTCGC CTATTTATCA TCGTGTTCG 6120
 GTAGCaATGC GTAAACACAG TACCACTAAA TAAGTGCACG ATACATGCAT CAAATGTCGT 6180
 CTTTAGTcTA AGTAACGATC ATGCATTAACT ATTTTCAAAA TATCTATTTG AGCTTGAAGA 6240
 30 TCTTTACCAA TATTGGTATC ACGAATCTTC TTACGTTGTA ATTCTTTATC TACGACGCGC 6300
 TTTATAGAAA GTTCATCGAT ACCTTCGGAA AGTATTTTTn CTTTAGCGTT AAATTGTTGG 6360
 35 TGTGCAACGA GTTGCAATACC GAATGAATTA TACAATAGTG TATAGCCTGC AATGCCAGTh 6420
 GTTGACTGAT AAGCTTTTGA AAAGCCACCA TCAATGACAA GCATCTTTCC ATCAGCCTTG 6480
 AT - 6482

40 (2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16592 base pairs
 (B) TYPE: nucleic acid
 45 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

ATTTAAGGCG ATTGCTTG TG TATTTCTCTC TTTGTAGGC AAACCTGCAC TCGTTCCAAA 60
 AAATGTAAC TCCATATATG CCCCTCCTTT TCTTCAATTC ATTTTATCAT AAAATTTGTA 120

55

	AATTTTCTA	ACTTTAACGT	AGACATAACT	ATATAAAATT	TGATAATTAC	GTTATACTTA	240
	TCATTAATAA	GTATCACATT	AAACATGATA	CATGAATCGA	TATTTCAITT	AAGACACTGC	300
5	ATACAGTCGA	GCATATTGTA	TGACCTACTG	AATGGATTAT	CTTATAATAA	TAAATCATAT	360
	ATCTAATTAA	GAATTGAGGT	TTTAATCTTG	AGTACTAAAA	ACAAACACAT	CCCATGTTTA	420
10	ATCACAATCT	TTGGTGCACT	GCGTGACTTA	AGCCATCGTA	AGTnGTTTCC	ATCAATATTC	480
	CATCTCTACC	AACAAGACAA	TTTAGATGAA	CATATTGCCA	TcATCgGTAT	TGGACGTCGT	540
	GACATkwnTA	ATGATGATTT	CCGTAATCAA	GTAAAATCAT	CAATTCAAAA	GCACGTAAAA	600
15	GATACAAACA	AAATTGACGC	GTTTATGGAA	CATGTCCTCT	ATCATAGACA	TGATGTTAGT	660
	AATGAAGAAA	GCTATCAAGA	ATTACTAGAT	TTTAGTAATG	AATTAGATAG	CCAATTTGAA	720
	TTAAAAGGTA	ATCGACTATT	CTATTTAGCA	ATGGCACCAC	AATTCTTTGG	CGTTATTTCT	780
20	GATTATCTAA	AATCTTCTGG	TCTTACTGAT	ACAAAAGGAT	TTAAACGCCT	TGTTATCGAA	840
	AAACCATTTCG	GTAGTGATTT	AAAATCAGCC	GAAGCATTAA	ACAATCAAAT	TCGTAAATCA	900
	TTTAAAGAAG	AAGAAATTTA	TCGTATTGAC	CACTATTTAG	GAAAAGACAT	GGTTCAAAAT	960
25	ATCGAGGTAT	TACGTTTTGC	GAATGCGATG	TTTGAACCAT	TATGGAATAA	CAAATATATT	1020
	TCAAACATCC	AAGTTACATC	TTCTGAAATA	CTAGGTGTG	AAGATCGTGG	TGGTTATTAT	1080
30	GAATCAAGTG	GCGCGCTAAA	AGATATGGTG	CAAAACCACA	TGTTACAAAT	GGTTGcATTA	1140
	TTAGCTATGG	AAGCACCTAT	TAGTTTAAAT	AGTGAAGATA	TCCGTGCTGA	GAAAGTAAAA	1200
	GTACTIONAAAT	CACTGCGTCA	TTTCCAATCT	GAAGATGTTA	AAAAGAACTT	TGTTCGTGGT	1260
35	CAATATGGCG	AAGGCTATAT	CGATGGTAAA	CAAGTTAAAG	CATACCGTGA	TGAAGATCGC	1320
	GTTGCAGATG	ACTCTAACAC	ACCTACCTTT	GTTTCAGGTA	AATTAACAAT	TGATAACTTT	1380
	AGATGGGCTG	GTGTACCATT	CTATATTCGT	ACTGGTAAAC	GTATGAAATC	TAAAACAATT	1440
40	CAAGTTGTCG	TTGAATTTAA	AGAAGTACCA	ATGAACCTAT	ACTATGgAAA	CTGaTAAACT	1500
	GTTAGATTCA	AACCTATTAG	TAATCAATAT	CCAACCTAAT	GAAGGTGgTA	TCTTTtACAT	1560
	CtAAATGcTA	AGaAAAATAC	ACAAGGTATC	gAAACAGrAC	CTGtCCmATT	GtCTTACTCm	1620
45	ATGaGCGcTC	aAGaTAAAAAT	GaATACTGTA	GATGCATATG	AAAATCTATT	ATTTGATTGT	1680
	CTTAAAGGTG	ATGCCACTAA	CTTCACGCAC	TGGGAAGAAT	TAAaATCAAC	ATGGAAATTT	1740
50	GTTGATGCAA	TTCAAGATGA	ATGGAATATG	GTTGaTCCAG	AATTCCCTAA	CTATGAATCA	1800
	GGTACTAATG	GTCCATTAGA	AAGTGATTTA	CTACTTGCTC	GTGATGGTAA	CCATTGGTGG	1860
	GGACGATATT	CAATAATTGA	ATTAAAACGC	ACATGTTTAA	CAAAAATAAA	TGAGCGAATG	1920

55

	TATATTATGA AATTATATTT TACAATGCCC AAAACTATTT TAATAATCAT TGAACAAATG	2040
	GGTGTATAAT TTATAGAAAT AATGTAGAAT AAAAATAAAT GATTGAATTA ATTGGAGTGA	2100
5	AAGTTTTGGA CGTTATCAAG CAAATACAAC AGGCAATTGT TTATATTGAA GATCGTTTAT	2160
	TAGAGCCTTT CAATTTGCAA GAATTAAGTG ATTACGTTGG TCTTTCGCCA TACCATCTTG	2220
	ATCAATCATT TAAAATGATT GTCGGCTTAT CTCCAGAAGC TTATGCACGC GCGCGTAAAA	2280
10	TGACACTCGC TGCAAATGAT GTGATTAATG GTGCTACACG ACTTGTAGAT ATCGCTAAAA	2340
	AATATCACTA TGCAAATTC AATGATTTTG CAAATGATTT TAGTGATTTT CACGGCGTAT	2400
	CACCTATTCA AGCCTCTACT AAAAAAGATG AATTACAAAT TCAAGAGCGA TTATATATCA	2460
15	AATTATCAAC TACTGAGAGA GCACCTTATC CATAAGATT AGAAGAGACA GATGATATTT	2520
	CATTGGTTGG ATATGCACGA TTTATAGACA CTAAGTATTT GTCACATCCT TTTAATGTTT	2580
20	CGGATTTTTT AGAAGACTTG CTCATTGATG GTAAAATTAA AGAGTTACGA CGATATAATG	2640
	ACGTTAGTCC ATTTGAACTA TTTGTTATTA GTTGTCCTCT TGAAAATGGT TTAGAAATAT	2700
	TTGTAGGTGT ACCAAGTGAA CGTTATCCTG CACACTTAGA AAGTCGATTT TTACCTGGCA	2760
25	AACATTGTGC GAAATTCAAT TTACAAGGTG AAATTGATTA TGCAACTAAT GAAGCTTGGT	2820
	ACTATATTGA ATCAAGTTTG CAGTTAACAT TGCCATATGA ACGAAATGAT TTATATGTTG	2880
	AAGTGATCCC TCTCGATATT TCATTTAATG ACCCATTAC TAAAATTCAG CTTTGGATTG	2940
30	CTGTTAAACA GAGTCCTTAT GACGAAGATT AAATAATAAA AAACAAAGAA GCCCCCTAAT	3000
	ATATCTATAG GTCTACAAAT GGCCTTAGAT TCTATTAGGG GGCATATTAA TATGTTAATT	3060
35	TAGTTCGATA ACACATGCTT CATATGGACG TAACTGTTTT AAATTAACCT TGGCATCATA	3120
	ATTAAATAGC TTTACTTCTC CATGGCTTAA ATCAAATGGT ACAGTTAATT CTGCTTCGTG	3180
	GTTAGTAAGA TTACCTACAA TAAGAACTTG CTTTTCATTT AATGTTCTCG TGTACGCAAA	3240
40	AACTTGTGAA TTTTCAGCAT CTACTAAATC AAATTGACCA TATACGTATA CATCATTAGA	3300
	CTTCTTAAT TGAATTAAAT CTTTATAAAA TTGTAATACT GAATGCTCAT CTTCTAATTG	3360
	TTGTGCAACA TTGATAGTTT TATAATTCGG ATTCACTGGG AACCACGGTT CACCATTGTG	3420
45	AAATCCTCCA TTAAACGTAT CATCCCATG CATTGGTGTG CGAGAATTAT CTCGGTTCTC	3480
	ATCTTTATAT TTCGCAAGTA AAGCGTCTAC ATCTCCACCT TGAGCTTTCA CTATTTGATA	3540
	GTCATTTTTA ACAGCAACAT CGTTAAACGT TTCAATACTT TCAAATGGAT AATTCGTCAT	3600
50	ACCAATTTCT TGACCTTGAT AAATGAATGG CGTACCTTGT TGCAAGAAAT AAACAGCTGC	3660
	ATGACTTGTT GCTGATTCAT ACCAATACTT GTCATCGTCA CCCCACGTCG ATACACGTCG	3720
55		

	CCATCTATTT AATACAGATT TATACGAATT TACATCAAAG TGAGAATCAC CACTATTCCA	3840
	CAGTCCCAAA TGTTCAAATT GGAATATCAT ATTAAATTTA CCATTTTCTT CCCCAGCCCA	3900
5	GTCATCAGCA TCATCAGGGC TTACACCATT CGCTTCACCA ACAGTCATAA TGTCACTTT	3960
	ACTTAATGAG CGATCTTTCA TCTCTTGTA CCAAGTTTGT ATACCTGGCT GATTCATATC	4020
	TACATCAAAT GCTGGGGCAT ATGTTTTACC CTCAGGTACA GGTAAGTCAC CCGCTTCAAA	4080
10	CGTCTTCTTA ATATGCGTAA TTGCATCTAC TCTAAATCCA TCAATGCCTT TATCAAACCA	4140
	CCAGTTCATC ATTTCAAATA CAGCATCTCT AACTTCCGGA TTACCCCAAT TCAAATCAGG	4200
	TTGTTTTTTA CTGAATAAAT GGAAATAATA TTGCTCAGTA TTAGCATCAT ATTCCCATGT	4260
15	AGATCCATTA AATATACTTT CCCAGTTGTT AGGTTTCAGAG CCATCTGGCT TTGGATCTTG	4320
	CCAAATGTAC CAATCACGTT TGGGATTGTC TTTACTAGAT TTGGATTCTA TAAACCAAGG	4380
20	ATGTTTCATCA GATGTATGAT TTACAATAA ATCTAAATA AGCTTCATGC CTCTATCATG	4440
	AACACCTTTT AATAAACGAT CAAAGTCTTC CATCGTTCCA AATTCATCCA TAATCTCTTG	4500
	GTAGTCACTA ATATCATAAC CATTGTCATC ATTAGGTGAT TTAAACATTG GACTGAGCCA	4560
25	AATGACATCG ATACCGAAAT CTTTTAAGTA GTCCAATTTA TCAATCATTC CAGGTAAATC	4620
	CCCAATACCA TCGTGATTAC TATCATTAAA ACTTCTTGGA TATACTTGAT ATGCTACTGC	4680
	TTCTTTCCAC CATTGCTTAT TCATTTTAAA ACTCCTTTCG TATCGCTGTG TTGATTTTCT	4740
30	TATTTTTAAT TCTGTATCTA TAATGACGAG TTCAATAACA TCCTGTGCTT TGTTTTTCAA	4800
	TATATTTAAA ATTGCTGCAC CAGCCTGTTG ACCTAACATT CGAGGCTTGA TGTCAATACA	4860
	GGTTTGTGGT GGTGACGCAA TTTCGGTAA ATAAGAATCA TTGAACGTTG CTGTCATTAC	4920
35	ATCTTTCCGA ATTTCAATAT TAAGTTCATA TAGGACACTT AAAATCGCTA AATGTAACAT	4980
	AGCATCTAAC GAAATGATTG CCTGTTTAAAT ATTTGGGTCC TTCAAACGCG TATGTAGATT	5040
40	TTGCATGTAA TTAAAAATAA CTTCTCTTTC ATTACTAGTC TCAATAATTT GATAATTAAT	5100
	TTTATTTTGA GAAGCTATCG TTTCAAATCC TTGAATTCTA TCTTTTGAAA CTTCAAAATT	5160
	TCCTTTTCT GTAATAAATA TTAATTCATC TACACCTTGT TCAATAACAT GTCGTGTCAA	5220
45	ATTTTCAGAA GCTAATATAT TATCATTATC TATATGTGTA AATTGATGAT CTATATCCGA	5280
	TGTAGGCTTA CCAATCACAA TAAATGGCAT GCTTTCATCA ATTAACATTT GTTTAATCGG	5340
	ATCATTTTCT TTTGAATAGA GCAGTATAAA CGCATCAACC ATTCGTTGTT TAATCATTTT	5400
50	ATAAACTTCA TCCATTAAAT CATTCAATAT ATTTGAGACT GTCGTTTGTG TACCATAGCC	5460
	ATGCTGGTTA CACGTTTCAG AAATTCCTAG CAATACATTG ATGTAGAATG GATTCAGTCG	5520

55

	AGTTCTAGCA GCGGTATTAG GAAAATAATT CAATTCTTCC ATAACCTTCT TCACCTTTGA	5640
	AATTGTGCGT TCGCTAATAC GTTGATTTCC TTTTATAACT CTTGAAACTG TCGAAGGAGA	5700
5	AACACCGGCT TTTAGTGCAA CATCTTTAAT CGTAACCATT TAATCACCTC CTGTTAATTT	5760
	CTGCATCGGA AAACGCTTCC AACCACTGTA TAATACCAGT TTAGTCACAC TTTCTAAAAA	5820
	AGTCAAAAGA TTTGTGCAAA CGATTGCATA AAACGATAAA AATAAACCT TCATACTGAA	5880
10	ATTCAATCCG AAAATCAATA TAAAGGTTTG TATAAATATT AAAATCGATT GTTTAGTCAC	5940
	TAACTGCAAA ATAGTTACCT TGGCCATCTT GAAAATTAAA TACACGTTGA CCATTCAATT	6000
15	CTACTATATC ATGCCCAGTT AAACCTAAAT CATTTAATTT TGAGTATAAT GCATCAAAGT	6060
	TTTTCTCTTT AAACATTAAA GATGGTGTTT CTAGGTTTAC TTCCGGGCTA TGCTTTTCAA	6120
	TAAATTCTTT TGCCATAATC GTCAATGACG TTTTACGATC TTTGGTAGGT GATACTTCAA	6180
20	CTGCAACATA GTCCTCAGCT AACGGTGTTT CACTTACAAC AACAAATTCT AAAGTTTCTG	6240
	TCCAAATGC TTTGCTTTT TCGACATCAT CAACATATAA CATAACTGA TTTAACTTTT	6300
	CCATAAAATA GTACCTCTAT TTCTCTATAG TACATGCTAT CATAACACAG TAAATATTTT	6360
25	ATTACTTCAC AAAATGCTTA AAAATATGGC GGGATGCTTT TAAGGTCAAG GATAATACTT	6420
	GTGTAATTTT TTATAGGTTG TAGCTACTCT ATCACACTCT CTTTTATATT TATCAAAAGA	6480
	TATAAAAAAG GATAGTATCT TTCAACTATC CTTTAATCAA TATTATTCTT CAATCCATTG	6540
30	TGTATGGAAT ACGCCTTCTT TATCTTTTCT TTCGTACGTA TGAGCACCGA AGTAGTCACG	6600
	TTGTGCTTGA ATTAAGTTTG CAGGTAAATC AGCAGCACGG TAACTATCAT AGTAATTAAT	6660
35	ACTTGATGAG AAACCAGGTG TTGGTACACC ATTTTGAACA CCAGTTGCGA CAACATCACG	6720
	TAACGCATCT TGATATTCAG TAACGATGTT TTTAAAGTAA GGATCTAGCA ATAAGTTTGT	6780
	TAATCCTGGA TTATTATCGT AAGCATCTTT GATCTTTTGT AAGAATTGTG CACGGATAAT	6840
40	GCAACCTTCT CTCCAAATCA TAGCTAAATC ACCAAGTTT AAATTCCATT CATTATCTTC	6900
	ACTTGCTTTA CGCATTGTGCG CGAAACCTTG TGCATAAGAA CAAATTTTAC TCATATATAA	6960
	TGCTTTACGA ATTTTTTCTA AAAAGTCTTT CTTGTCACCA TCAAATGATG CTTTGGACC	7020
45	ATTTAATTCT TTAGAAGCAT TTACGCGCTC TTCTTTGATT GAAGAGATAA AACGTGCAAA	7080
	TACAGATTCA GTAATGATTG TTAATGGAAT ACCTAATTCT AATGCGTTAA TTGAAGTCCA	7140
	TTTTCCTGTA CCTTTTTGAC CTGCAGTATC AAGAATTTT TCAACTAATG CTTCTTTATT	7200
50	TTCATCTAAT TTCATGAAAA TATCACCAGT GATTTCAAAT AAATAACTTT CTAATTCACC	7260
	AGCATTCCAG TCTTTGAACG TTTGAGCAAT GTCTTCATGA GACATGCCTA ATAATTCTTT	7320
55		

	CATTTTCACA TAGTGTCCAG CACCATTAGG TCCAATATAA GTAACACATG AAGCACCGTC	7440
	TTTTGCCTTT GCAGCAATTG CATCAAGAAT ATCTGCAACT TTGTTATAAG CTTCTTCTTG	7500
5	TCCACCCGGC ATTAATGACG GACCAGTTAA CGCTCCAATT TCACCACCAG AAACGCCCAT	7560
	ACCAATAAAG TTGATTGCAC TTTGTGyWAA TGCTTTATTA CGTCTGATAG TATCTTGATA	7620
	GTTTGTATTA CCACCATCAA TTAAATATC TCCATCATCT AATAAAGGTA ACAAACATATC	7680
10	AATCGTTGCG TCCGTAGCTT TACCTGCTTG AACCATTAAAT AAAATTTTAC GTGGTTTTTC	7740
	TAAAGAATTA ACAAATTCTT CCAATGAATA CGTTGGATGA ATATTTTTCC CTTTTGATTC	7800
	TTCAACCATT AAATCAGTTT TTTCACTTGA GCGGTTAAAT ACAGATACAC TATATCCGCG	7860
15	TGATTCAATA TTCCAAGCTA GGTTTTTTACC CATAACGGCT AAACCAATAA CTCCAATTTG	7920
	TTGTGTCATA TTACTTACCT CACTTGTTGA TTTTTCATTA GTATTGTATC ACAAATAGA	7980
20	CATACACTAC ACTAAATCAT TTCGAATGTC GCGCAACTAT TTTGATTATT TCTAACACTT	8040
	GACTTGCAAG CAAGTTCAAT GATTTAATCG GCATTCTCTC ATTTGTTGTA TGGATTTTTT	8100
	CATAACCCAC TCCTAAATG ACTGAAGGAA TACCAAATGT ATTAATAATA CTGCCGCTCG	8160
25	AACCGCCACC AGAAATAATT GTATTTGCAG ATAATCCTAA ATTACGAGCA CTTTCTTGTG	8220
	CAATTTTAAC AACCGCTTCA TTATCATTAA TTTTAAATCC TGGATAACTT TGCTCCACTG	8280
	TAACACTGCT TTTCCACCT AATTCTGATG CAGTAGTTT AAACACATCA GTCATATGTT	8340
30	TGACTTGTGT TTTTATTCTT TCTGGATCGT GAGAACGTGC CTCTGCTTCT AAAATGACTT	8400
	CATCTGCAAC AATATTCGTA GCTGAACCGC CATGAACTT ACCAATATTG GCAGTAGTTA	8460
	TTTCATCAAC TTGTCCTAAT TTCATTGAC TAATTGCTTT CGCCGCAATA TTAATAGCAC	8520
35	TAACACCCTC TTTTGGCGTA CTTGCATGAG CCGTTTTGCC AAAAATTTTA GCTGAAATTA	8580
	ACATTTGCGT CGGTGCACCT ACAACCGTAG TACCGACATC AGCACTTGCA TCAATAGCAT	8640
40	AACCAAAGTC CGCGTCCAAC AACTCTGAAT TTAATTCTTT AGCACCAATT AAACCTGATT	8700
	CTTCTCCAAC AGTAATCACA AATTGAATTT GTCCATGTGG GATTTGTTGT TCCTTTATCA	8760
	CTTGCAAAAC TTCAAGCATC GCTGATAATC CTGCTTTATC ATCTGCACCT AGAATAGTCG	8820
45	TACCATCAGA GTATATGTAG CCGTCATCTT TTACAATTGG CTTTACATTA ATTGCGGGTA	8880
	CAACAGTATC CATATGGCTC GTCAAATATA ATTTAGGTAC TTCGCCTTCT TCGATAGTAC	8940
	TATTCATTGT ACACACTAGA TTATTGGCAC CTAATTTAGG ATGTTTAGCC GCTTCATCTT	9000
50	CTTTAACATC TAACCCTAAT GCTATGAATT TTTCTTTTAA AATAGGTTGG ATTGTTGATT	9060
	CATTCCCTGT CTCAGAATCG ATTTGTACAA GTTCAAAAAA CGTATTAAAGT AATCTTTGCT	9120

55

	GATGAAATAA AATGTTACAG TAATTGACGT TACACAGATT TATCAGGTTT GTAAATTGTG	9240
	TCATATTATT TTCAATTTAT TATATATAAT TATTGTAAGT CAAACTAAGC TTTGTCAAAA	9300
5	ATATATTGAT TGATTTTTCA AAGATATCGT ATAATGAGGA AAATGACATA AGCAAACCTA	9360
	CTCATGTTTT TTATTATATT CCTTTATGAT GATTGCTAGT TATATCGTCT CAAGTTAAAA	9420
10	GTTTTATATC TTATGTCGTA ATTATTAATA CAAAGGTTAT TCATTTGGAG GCACACAAAA	9480
	TGCAAAATAA AGTTTAAAGA ATTATCATT TCGTTATGCT TGTATCAGTT GTATTAGCAT	9540
	TGTTATTAAC GAGTATCATT CCAATTTTAT AAACATATATC TCAACTACCT ATACAAAATC	9600
15	ATACAATTAA AAATCCATCC ATTATAAAGC CATGTATTAA TAAGTTATCG TATTGCAACG	9660
	ATTACTTTCA AACATGGGTC ATACGGATGG ATTATTTTTT AAGCTACTTC ACTATGCATT	9720
	TTCAATGAAC CAAATTGCGA TTTGATTTGT AAATATTCTT CTAATTCATT TAATATTTGA	9780
20	ATAATACTTG CTCTCGAGTT AAGCGCTTTG TGTGTTGTTG GCAATGGCAG TTCATCCAAT	9840
	TTCAAACGCG TCTCATACAA ATTGTGTAAG CGCATTGCTG TATAGTCATT ACTATTCACA	9900
	TTTAGACCAA TTTCTTTCAG CAGTGACGCA ACATCATTTA AAAGCGGATC TTTATGACAG	9960
25	ATACTTTCGA TGAGCGGTTT CATTCTCATT AACAATTCCA CTTGCTCTTC TCGCATATCA	10020
	AAATAATGAT AGTATGAATT TTCGTTTCTA ACAAATGAT TTTTAACATC TCGGAACGCG	10080
	ATAGACTTCG CCTTTTTAAT ATTTAAAAGT AACACTTCAA ATTCAATCGC AATGGTATCT	10140
30	TCATATTTTT CACAAATATA ACTATATTTA CTAAAAATAT CAGCAATTTG TTGCTCAATT	10200
	TTACATTTGT ATTGCTCTAG TTGTTTGTCT AAACCTGGCA TCATTAAATT CaTTGTAAAT	10260
35	GCAATGCTTA GTCCAATTAA CAGTAATAAT GTTTCATTAA CAATTAAATG TGCATCAATT	10320
	GATTTTGCAT TAAAAACATG AAGTAATATA ACGCAACTCG TAATGACACC TTCTTGTAAT	10380
	TTTAATACGA CAGTTAATGG TATAAATAAC AATACGATAA TACCGAGTAC AATTGGACTC	10440
40	TGACCTAATA AACTAAATAT TGCTGAACCT AAAACAATA CTAAAAACA TGATACTAAT	10500
	CTTGAAATAA TCGCTTGTAG CGAATGTACT TTTGTATGTT TAATACATAA TACGACTAAT	10560
	ATGGCGCTTG AAGCATAATT ATCTAAACCT AACAGCTTAC TAATAATTAC ACCTAAAGTC	10620
45	ATACCCACTG CTGTTTTTAT TGTTCTAAAT CCAATCTTGT AAGGATTTAA CTTTAACATG	10680
	GGTTAGCGCC TCTTATCTTT CTTTACAATA TTTATTGAAT AATGTTTGTA ATTGATTAAT	10740
	TACGTTTATC ACATCATGAC CTTGATTTG ATGCTTTTCA ATCATTCTTG TAATCTTTCC	10800
50	ATCTTTTACT AATGCAAATG ACGGACTTGA AGGCGCATAA CCTTCGAAGT ATTCACGCGC	10860
	TCTTTGTGTC GCTTCTTTAT CTTGTCCAGC AAATACTGTC ACTAGACGAT CAGGTAATAC	10920

55

	AGAATTGATC	ATAACTAGTG	TTGTACCATC	TTGTTTAAGA	ACTTTGTCAA	CATCTTCTGC	11040
	AGTAGTTAAT	TGCTCATATC	CCGCAGATTC	AATTTCAATC	CTTGCTTGTT	CTACAACACC	11100
5	GTTTCATGTAT	AAATCGAAAT	TCATGnCCAT	AAGTTCAATC	ACCTATCCCT	TTATATTTAA	11160
	ACTAtCCTCA	TTCTACTAAT	TAATAACATA	TTGTTCAATA	AACTAATCTG	AATCACACCT	11220
10	ATATTTAGAC	ACAATTTTAA	CAATATACCA	AACATTATTG	TGCTTAAAT	CATGGTAACT	11280
	AATTTGTTCA	CATGTTTTC	TTAATATGTT	TCAAGTATGA	TGTCTTATTT	TGACTTTACT	11340
	GCAAAAATGC	ATTCAACCAT	GTTGATTATT	GTTCTTTATC	TTTTTTGAAT	ATATTGCACA	11400
15	TATTTTAGTG	CCAAAAAATA	ATACATCCAT	CGACAAGAAC	AAGATAAAAC	AAGTTGTCGA	11460
	TAGATGCATC	TATGTTATCA	CTAATATATA	TTTGATTATTT	CTAAAGTATA	CTGTTTCGATA	11520
	CGCTGTTTAA	TATGATTCAT	ArATTACCT	GTTTGTAAC	CATCTAAAAT	ACGATGATCA	11580
20	ATTGAAATAC	ATAAATTAAC	CATGTTACGA	ATTGCAATCA	TATCATTAAAT	TACTACTGGC	11640
	TTTTTAACGA	TTGATTCTAC	TTGTAAAATC	GCTGCTTGTT	GATGATTTAT	AATACCCATT	11700
	GATGATACTG	AACCAAATGT	ACCAGTATTA	TTTACCGTAA	ATGTACCGCC	CTGCATATCT	11760
25	TCAGCTGTCA	ATTGCTTATT	ACGCGCTTTC	GTTGCTAAAG	TATTAATTTT	TCTAGCTATA	11820
	CCTTTGATTG	ACTTTTCGTC	TGCATGCTTA	ATCACAGGTA	CGTATAATTT	ATTTTCATCA	11880
30	GCAACAGCAA	TTGAAATATT	AATGTCTTTA	TGTAAGACAA	TTTCATTTCC	TTGCCAGCTA	11940
	CTATTTAATA	AAGGATATGC	TTTTAAAGCA	TCTGCTACAG	CTTTTACAAA	GAAAGCAAAG	12000
	AACGTTAGAT	TATATCCTTC	TTTATTTTTA	AAGCTGTTTT	TATAATGATT	TCTCGTATTC	12060
35	ACAAGATTTG	TAGCATCTAC	TTCAATCATC	ATCCATGCAT	GTGGAATCTC	TGTTACACTA	12120
	TTAACCATAT	TTTGCGCAAT	TGCTTTACGC	ACACCATTTA	CTGGTATTGT	GCTGTTTTCA	12180
	CTATGTCTT	CAGATGATTG	GTTACTTGAT	GTATCTACTG	ATGTTGATTT	TGTTTGAACT	12240
40	TGTTTGTCAG	ATTGAGCTGT	GGTACCACCA	TTTTCAATAA	CTGACATTAT	ATCCTTCTTA	12300
	GTTACACGAC	CTTCAAATCC	ACTACCTACA	ACTTGTGATA	AATCAATGTC	ATGCTCTGAA	12360
	GCGAGTTTAA	ATACAACAGG	TGAAAAGCGA	CCATTATTAC	GTGGTTGATT	TTGTTTAGCA	12420
45	GTAGATGTCT	GTTCCACTGT	TGCACTAGCT	TTTTTAGTAG	ATTTCTGAGT	ATGCTCATCC	12480
	ACTTTTGCTT	GTATCTCTTC	AGTTGTTTCA	TTTGTCTTTT	CATCAGCAGT	TTCAATTTTA	12540
	CAGATAATTG	TATCAATAGC	TACTGTCTGC	CCCGCTTCAA	CTAAAATTTT	TGTAATTGTT	12600
50	CCTGATATCG	TGGAAGGGAC	TTAGCTGTCT	ACTTTATCTG	TAATAACTTC	ACATAATGGT	12660
	TCATATTCAT	CAATATGATC	ACCAACAGAA	ACTAACCATT	GTTCAATGGT	GCCTTCATGA	12720

55

	AATTCACGCA TTTTATTTAA GATTTTTTCT GGATTCATCA TAATTTTCATT TTCTAATACA	12840
	GGAGAAAATG GCATAGATGG TACatCTGGA GCAGCTAAAC GCATGATTGG TGCATCTAAA	12900
5	TCGAACAAGC AATGCTCTGC AATAATCGCT GACACTTCTG ACATAATACT ACCTTCTAAA	12960
	TTATCTTCAG TTACAAGTAA AACTTTACCT GTATGTTTAG CACGATCAAT AATTGTTTCT	13020
10	TTATCTAATG GATAAACAGT TCGTAAATCA ACGACTTCAA CATTGATACC GTCTGCAGCT	13080
	AAAATATCCG CTGCTTGTA ACAATAATTG ACCATTAATC CATAACAAA TACTGTTAAA	13140
	TCTTCACCTT CACGTTTCAC ATCTGCTTTT CCTAAAGGTA CAGTGTAATA TTCTTCTGGC	13200
15	ACTTCTTCCT TTAAGAAACG ATAAGCTTTT TTATGCTCAA AGTACAATAC TGGATCATT	13260
	GATTTCGATAG ATGATAATAA AAGCCCTTTA GCATCATACG GTGTGGAAGG AATAACAATT	13320
	GTTAAACCTG GCGATGAAGC AAATATACTT TCAATACTTT GTGAATGATA TAGTCCTCCG	13380
20	TGAACACCGc CACCAAATGG TGCACGAATC GTTAATGGGC ATTGCCAATC ATTATTTGAA	13440
	CGATAACGCA TTTTCGCAGC TTCACTAATA ATTTGATTG TCGCAGGTAA AATAAAATCT	13500
	GCAAATTGAA TTTCTGCAAT TGGTCTTTTA CCTACCATAG CTGCACCAAT GGCAGTTCCA	13560
25	ACAATATTTG ACTCAGCTAA TGGCGTATCG ATAACTCTGT CTTCAACATA TTTTGTGTC	13620
	AGTCCTTGAG TAGTACCAA TACGCCACCT TTTCTACCAA CATCTTCACC AAGAATAAAC	13680
	ACATCTTTAT TTTGTTGTAA TGCTAAGTCT TGTGCCcGcG TATCGCCTCT AAATAAGATA	13740
30	ATTTAGCCAT TAGTTAAGAC TCCCTTCTTC GTACACAAAT GCATAGGCTT CTTGACACT	13800
	TGGATATGGC GCGTCTTCAG CAGCCTTTGT CGCTTTATTG ATGATGTCTT TrATgTCCGC	13860
35	TTCTATTTCT GCCAACCAAG CATCATCGAT AATGCCAGCT GAAAGCAACT CTTTTTTGAA	13920
	CTTTTCATTG CAGTCTGCTT TTTTAAGcGT TTCACGCTCT TCTTTCGTAC GATATTGGTC	13980
	GTCATCATCT GATGAATGAG CTGTCATACG ACTTGTTACT GCTTCAATCA AAGTTGAACC	14040
40	TTGACCAGAA ATAGCTCGAT CTCTTGCTTC TTTTCATCGCT TTATACATTG CTAATGGATC	14100
	ATTACCATCT ACTTGTTTAC CATGTATACC GTAACCAAGT GCTCTATCCG ATAATTTTTT	14160
	AGCTGCGTAT TGTAATGAAT CAGGTACTGA AATTGCATAT TTATTATTTA TAATGACACA	14220
45	TACAAAAGGA AGTTTGTGTA CACCCGCGAA GTTTAAACCT TCATGGAAGT CACCTTG GTT	14280
	TGAGCTACCT TCACCAACAG TTGCTGTTGC AATTTTCTTC TTACCATCCA TTTTAAAGC	14340
	TAAAGCAGCA CCAACAGCAT GGGGTATTTG AGTTGCTACC GGTGAACCTT GAGACAAAAT	14400
50	ATTCTTAGCT CTACTACTAA AGTGTGATGG CATTTGTTTT CCACCAGAGT TAACATCGTC	14460
	TTTCTTTCCA AACGCTGATA AAAACGTATC ATACGCTGAG ATACCCATAT AAGTAACGAA	14520
55		

	AATCTGAGTT GCTTCTTGTC CTTGACCACT TACAACAAAT GGAATTTTAC CTGCACGGTT	14640
	CAATAACCAC AGTCTTTTCAT CTATTTTTCT ACCTAAATCC ATCCATTTAT ATATTACTTT	14700
5	TAGGTCTTCT TCGCTAAGGC CTAATGATT ATAATCAATC ATGTTAAATC CTCCTATTTA	14760
	TACGTGAATA GCTCTACTTT CTGCTTTCAA TCCTAATTC ATCAACACTT CAGAGATGGA	14820
10	AGGATGTGCG TGTGTTGTTA GTCCTAATTC TAATGCCGAG CCATTCATGA ACTGTAACAG	14880
	TGATGCCTCA TTAATCAATT CTGTTACATG TGGACCAATC ATATTAATAC CCACAATTTT	14940
	TTCAGTTGAT TGATCAATCA CCATTTTCGT ATACCCTTCG TTTGTGTCAT GGCTATCAAT	15000
15	CACTGCTTTA CCAATTGCTT TAAATGGTAC TTTAAACTT TTAACITTTCA TTCCCTCTGC	15060
	CTTTGCTTGT TCAATGTTTA AACCGATAGA AGCAATTTCA GGTTGTGAAT AAATACACTT	15120
	AGGCATCATG TTATAGTTTA CTGGGATTGG GTTCCCTCA AACATATGAT CAACAGCCAC	15180
20	AACACCTTCT TTTGATCCAA CATGTGCCAA TTGTAATTTT CCTATACAAT CACCAGCTGC	15240
	ATAAATATGT TTATCTTCAG TTTGTTGAAA TTCGTTTCGT AAAATATGTC CTGATGTTGa	15300
	AAGtTTTATT TTAGTGTTGT TTAACCAAT ATCTGATGTG TTAGGTTTTT TACCAATCGA	15360
25	TAGCAACACT TTATCTACTT TAATTATGTC TGAGGAAATT TCAAACGTAA CACCATCTTC	15420
	GTTAACATTT ATATCATTTT CAGAAAGTTT TATTCCCTCA TAGAATTTAA CACCACGTGC	15480
30	TGACAATGAT TTTTTTAATA GTTGTGAAGC TTGTTTACTT TCAGTTGGTA AAATTCTTTC	15540
	ACCTGCTTCT ATAAGTGTTA CGTCAACACC TAAATCTATC ATCAATGATG CAAATTCCAT	15600
	TCCGATAACA CCACCACCAA TAATACCAAT ACTTGATGGT AACGTCTTTA ATGATAATAT	15660
35	ATCATCGCTA GATAAAATTT TATCATGATC AAATGATAAG AATGGCAACT CTGCAGGCGA	15720
	AGAACCAGTT GCAATTAATA CAAATTGGTT GGGTAATAAG TCTGATTCAC CATCTTCATA	15780
	TTGACAGAA ATTGTGCCAC TTTGAGGTGA AAATATAGAT GTACCTAGAA TACGTCCCGT	15840
40	GCCATTATAA ATGTCAATGT GATTGTGTTG CATTAAATGC TTTACACCTT GATACATTTG	15900
	ATTAATAATG TCTTCTTTTC GTGCCAACAT ATTTTCAAAA TTAACATTAG CATCTTTGAC	15960
	ATCAACGCCA AACATTGCTG CCTGTTTTAC TGTTTGAAAT ACTTCAGCAG ATTTAAGCAG	16020
45	CGATTTAGTA GGAATACAAC CTTTATGGAG ACAAGTACCT CCTAATAGTT GTCGTTCTAC	16080
	TATTGCCACT TTTTACCTA ATTGAGACGC ACGTATCGCA GCAACATATC CTGCAGTACC	16140
	TCCACCGAGA ACGACTAAAT CATATTGTTT CTCTGACATG TTCTTACTCC TAACTAATGA	16200
50	TATATATCCA TTGAAAATTT ATTAATACAT AGTTTTCATG TCCATTAAAT ACCTATTTTA	16260
	CATGATTGTC TATTTAGTTT GAATGCACAT AAATAAATCC ATAAATGAGT ATTCAACACA	16320

55

TAAATCAGTA ACACTTGCAC CTGAAATCAT TCGTGCAATT TCATCTACTT TATCATCGCT 16440
 AATTAACCTCT TGAACCTGTG TTGTTGTACG ATCATCTTTT GATGATTTTCG AAATTAATAA 16500
 ATGATGGTCG CTCATCGATG CAACTTGTGG TAAGTGAGAG ATACAAATAA CTTGTATATA 16560
 TTCTGCTaTA TCTCGCATTT TCTCTGCCAT TT 16592

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13794 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

CCAATACAAC GTAAAAGAT TGCTTGTGTT ATTAATGAGT TAGATAAAAT AATTAAAGGA 60
 TTTAATAAGG AAAGAGACTA CATAAAATAT CAATGGGCTC CAAAATATAG CAAAGanTTT 120
 TTTATACTTT TTATGAACAT TATGTACTCA AAAGATTTTT TAAAATATCG ATTAAATTTA 180
 ACATTTCTTG ATTTATCTAT CTTATATGTA ATATCATCTC GAAAAAATGA GATACTAAAT 240
 TTAAAAGATT TGTTTGAAAG TATTAGATTT ATGTATCCTC AAATGTTAG GTCAGTTAAT 300
 AGATTAAATA ATAAAGGTAT GCTAATCAAA GAACGATCCC TTGCAGATGA AAGGATTGTG 360
 TTAATCAAAA TAAATAAAAT ACAATATAAC ACTATTAAAA GCATATTCAC AGATACTTCC 420
 AAGATTCTCA AACCAAGAAA ATTTTCTTTT TAAATTAAAT CAGATTTACC TCTTGATAAA 480
 ATAAATAAGC AATCATACTA CTTCTCAATT TAGTATAAAT AAAAATACAT AATTAACTTT 540
 CTTTTGTTTT TATATTATTT CAATACCCTA CTATATATCA CAACACATAA ATTAAGCATG 600
 ACACTCATT C AATTTAGTTC ACCATTTCTG GTTCCAATTT TACTGAGTAT CATGCTTTTA 660
 ATGTTATAAA CCTAATGCTT TAATAAATCG TGTTAATTCT TCTCGCATAC TGTCATCTTT 720
 CAATGCATAT TCTATGGTAG TTTTAACGAA GCCTAATTTT TCTCCAACGT CATAACGTTT 780
 GCCTTCGAAG TCATATGCAT ACACTTGGTT ATCATTATTC ATACGTTCAA TCGCATCTGT 840
 TAACTGAATT TCGTTACCTG CGCCTTCTTT TTGCGTTTTT AAATAATCGA AAATTTCAGG 900
 CGTTAATACA TAACGTCCCA TAATAGCTAG GTTTGATGGT GCCGTACCTT GTGCTGGCTT 960
 TTCAACAAAC TTTTTCACCT CATACTGACG TCCGTTTTTA GTTAATGGGT CAATAATTCC 1020
 ATAACGATGA GTATCTGCTT CCGGAACCTT TTGGACACCT ATAACTGAGT GCCCTGTTTC 1080
 TTCATAAAG TCAATCAACT GTTTCACCTG TGGCACTTCA GATTCAACAA TATCGTCACC 1140

EP 0 786 519 A2

	TAAACCTTTT TGTCTTTCT GCCTTACATA AAAAATATTC GCAAGTCCG TTGAATACTG	1260
	AACCTTCTCT AGTAATTCAG ATTTACCTTT TTCTTTTAAC ACCATTCTTA ATTCTTTTGT	1320
5	ACTATCAAAA TGATCTTCAA TCGCGCGTTT GTGGCGACCT GTCACATATAA TAATATCTTC	1380
	AATTCCAGCT CTTGCAGCTT CTTCAACGAT ATATTGTATT GTGGGTTTAT CTAAGATAGG	1440
	AAGCATTTCC TTGGCATCG CTTTAGTTGC TGGTAAAAAT CTAGTCCCTA AACCAGCAGC	1500
10	GGAATGATT GCCTTTTTTA TTTTTTCAA AGTTAATGTG CTCCTTTTCC TAAGTATTAA	1560
	ATCTATGTAT CAACGTCATT TTAACACTAA TTAGAACGCC TTCATAGTGT CATTGAGTAT	1620
	GTAATTATTT CTTGGGAAAT TTGTTTTAAT TTTAAAAAAC AGGCTTACTT CATATAATTT	1680
15	ATGAAATAAA CCTGTCAATT TTGGATTGAT TATGCTTTGT GATTCTTTTT ATTTCTGCGT	1740
	AATAACGCTA AACCTAAAAT GCTAAATAAT CCGCCGAACA ACATGCCGTT GTTTGTTGAT	1800
20	TCTTCTCCAC CTGTTTCAGG TAGTTCAGAT TTCTTAGATT GTGCTTTTTT AGTTGGTACC	1860
	ACTGCTTTAA CCTTTTCATT GATTTCAATA ACAGGTGTTA CTACTTTACC TTGTTCCACT	1920
	GGTTTAGAAG GTTTTTTAGG TTCTTCTTTA GCAGGTGGTA TTGGTTTACC AGGTTCAGTT	1980
25	GGTACCTCTG GCGTTGGCGG TGTGGTGT TCCGGCTCGC TTGGTACTTC TGGTGTCGGT	2040
	GGTGTGGTG TTTCCGGCTC GCTTGGTACT TCTGGTGTG GTGGCGTTGG TGGCAGGATT	2100
	GGAGGTGTTG TATCTTCTTC AATCGTTTGT TGACCTTCAT TATGACCACT TACTTGTTGA	2160
30	AGTGTATCTT CTTCAAAGTC AACACTATTG TGTCCACCGA ATTGATAATT TGGTTTATCT	2220
	TTATTTGTAT CTTCTTCAAT AATTTCAAGT TGCTTATTGA ATCCGTGAAT ATGTGGCACA	2280
	CTGTGGAAGT CGATATCAAT GATATTACCA CCTTGTTTCAT ACTTAGGTTT GTCTTCTCT	2340
35	GATCTTCTT CGAATGATTG GTTACCATT TTTTGACCAT GAAATTGAGG TACACTATCG	2400
	AAATCGATAT CTACGATATT GCCACCTTGT TCATATTTCTG GTTTATCTTC TTCTGTGTCT	2460
40	TCCTCAAATG ACTGATTACC GCTATTTTGG CCACCTTCGT AACCTAATTC ACTCTTAATA	2520
	TCCACGTGGC TATTTTCTTC GATTTCTTCA ATCAGCCCAT AATTACCGTG ACCATTTTCA	2580
	GTTCTTAAAC CAGAATGAGA AATATGATGA TTGTTTTTCTG TAATTCCTC GATTGGTCTT	2640
45	TGCGCTTGAC CATGTTCTTC AGGTAGTTCA TCTACTAGTT CAATCAGATT ACTTTCAGTC	2700
	GTATATTCTT TCGTATCTTC AATTGTTGTA TGATCGCTAA CAGCACCAGT TACAATACCT	2760
	TTTGTAGAAT CTTTCGTCAA TTCAACTAGG TTAGACTCAG TAGTAACCTG ACCACCACCT	2820
50	GGGTTTGTAT CTTCTTCATA TTCAACAACA TCAGCATGAT GTTTTGAATT TTCATGTGTC	2880
	GATTCTTCAA AGTCTACATG AATAGAATCT TCTTCAGTTT CAATGGTACC TTCTGCATGA	2940

55

	TCTTCGATTG TACCAGTCAA TTCATGCTTC TCCACTGGCG GCTCTGATTT AAATTCAAGT	3060
5	TCGATAGGAG TACTATGTTC TATAATAGGT TCCTTTAGTT TATCTTTGCC GTCGCCTTGA	3120
	GCGTTATTAG AGTAAAATGC AACGCCATTT TTCCaAGTTA AATTACTTGT ATAATAATAG	3180
	TTATAATATC CAAAAAGGTG TGTGTGAAAT TCTAAGTTGC TAGCATTTGA ATCATAATAC	3240
10	CCTTCATATT TTATTACATA ATTTTACTT TGGTCTAAAT TATTAAAGTT TAAAGAATAA	3300
	CCACCATTAG TATCAAAATC TAAACTCATA TTATCASTCA CATCTTCAAA TTTGCTGACA	3360
	TCATCAAGCT TTGCATAnTn AgctTTCAGC TAAATCGTCT GAACCAATGT GTTTATATAC	3420
15	CTTAACTGTT GGATTATTAA CCCCTGGTTT ATTTCCTTTA GTTACTTGAC CAGTTACTGT	3480
	CACAGAGCTT AACGACTGGT TGTTAGGTTT CATGTACGCA AAATGACTAA ATTTCCCATC	3540
	TACTTTATTT AAAGTATCAA TTCGACCATT AGCTGTTACT CCCCAATTAT CTCTAACTCC	3600
20	ACCTAAATAT TGAATATTAA ATATTTTGCT AACCGTAGTC TCACCCAATT TAACTTCAAC	3660
	ATTTTGGTTA CCTTTTTCG TCACTGTTGT AGGATCAATA AATAGATTTA AAGATAATTC	3720
	AGCAGTTAAA TCTTTCTTTT CTGTACATA TTCTTTAAAC GTATATCTAA CTTTTCTTTC	3780
25	TCCAATTATT TCTCCTGTCG CCATAACTTG ACCATCTGTA CTTTTTATCT CCGGAACTTT	3840
	ACGCAGTGTT GAGATACCAT GAGTTTCAAC ATTATCGCTT AATGTGAAAT CAAAATAATC	3900
30	TCCCGCCTTA ATTCCTTCTC CAAATTTCa TTTATATTTT AAGGTTACTC TTTCTGCGTT	3960
	ATGAGGATTT ACAACATTCTG TATCTTGTTT ATGTCTTACA ATTTCACTAC CTTCTTCTAC	4020
	TTCCACTTTA TTTGTTACAT CTGTACCTGT CGCTTTAGTT TCTTCCACTA CTTCTTTCTC	4080
35	TGCAACTGCT GTAACGTCAc TGatCTTTTC ATTCCTGGTT TAATTTCTGA GACGTTACTT	4140
	GGTTGAGCTA TGTCAACTTG AGTTCCTGTA GTTTCCTTAT CAGCAACTTT TTCCGATGGC	4200
	AAATCAACTC GCGAagTTTC TACTTTTGGT GCTTGCaCAG TTTTCGGTGC TTCTTCTGTT	4260
40	GTTACTTGTTG TTGATTGTGA TGGTTGCTCA GTTGATGTCG CGCTGTATGA TTGTGTTTCA	4320
	TCTATTGTAT TAACGTTATT TGTAGTTGTT TGTGTTTCGC TTGCTTTACT TTCAGTAGCT	4380
	GAACCTCCAC TTTCTCTAC TGTAGTATTG TTTTGTTCCG ATGCTGCAGC TTCTTTTCT	4440
45	TGTCCCATTC CAACAACGAT CATTGTTTCT AAGAATACTG AGGCCGCTCC CAATTGTGT	4500
	TTTCTTATGC CGTATCTAAG ATTGCTTTTC ACTATAATAT TCTCCCTTAA ATGCAAAATT	4560
50	CATTTATTTT TAAAACTCAA TAAATGCAAT TCTATATTGT TCGGTTTTTA AAAGCAATGA	4620
	AAAAAAGCGA GTTAATAAAA AGTTAAGATT GTTGTTAACT TTATGTATAA TGAGTTTTTT	4680
55	ATTATTTGAA ACTCACATAT ATATTGCATA CAAAGCTCTT GAACACCTTG ATATAACAGG	4740

	TACTAAACCA TACATAATAA TCGCCTGTAC AATGCATCAT TAACAAGTCA CTGAAACGCC	4860
	TTTCATTGTA TTAATAACGT CACTATAATT TTTATATCGT TCGGTTTTTG TTTGATTTTA	4920
5	ATGATTATTT ATACAAAAAC AGCCGTATTT CAAGCCGACA TTTTAAATTT AACTAAATTT	4980
	GCATCTAGTT AATAATTGCA TTTATCAAAT TTGTCTTATT GATCCAATCT AATTTGTACT	5040
10	CACAACTAG TTTAAAATTC TAACTTTATC TCTCAGTTCG TTATCAATCA TCAGACATAA	5100
	ACCAATGAAG CAATCAGAAA ACACTCTAAT TTTCTATTAG AAATTTGATT TAATATAAAA	5160
	AAACAGGCTT ACTTCATATA ATTTATGAAA TAAACCCGTC AATTTTTGTT TAATTATGCT	5220
15	TTGTGATTCT TTTTATTTCT GCGTAATAAT GCTAAACCTA GAATGCTGAA TAATCCGCCG	5280
	AACAACATAC CTTTGTGTTG TGATTCTTCT CCACCTGTTT CAGGTAGTTC AGATTTCTTA	5340
	GATTGTGGTT TTTTAGTTGG TGCCACTGCT TTAACCTTTT CATTGATTC AATAACAGGT	5400
20	GTTACTACTT TACCTTGTTT CACTGGTTTA GAAGGCTTTT TAGGTTCTTC TTTGGCAGGT	5460
	GGTACTGGTT TACCAGGTTT AGCTGGTACC TCTGGTGTG GCGGTGTTGG AGTTTCTGGC	5520
	TCACTCGGCA CTTCTGGTGT CGGTGGTGTG GGTGTTTCCG GCTCAGTTGG TACTTCTGGT	5580
25	GTTGGTGGCG TTGGTGTTC CGGCTCACTT GGTACTTCTG GTGTCGGTGG CGTTGGTGGC	5640
	ACGATTGGAG GTGTTGTATC TTCTTCAATC GTTTGTTGAC CTTCAATTTG GCCGCTTACT	5700
	TTTGAAGTG TATCTTCTTC AAAGTCAACA CTATTGTGTC CACCGAATTG ATAACCTGGT	5760
30	TTATCTTTAT TTGTATCTTC TTCAATAATT TCAGTGTGCT TATTGAATCC GTGAATATGT	5820
	GGCAGACTGT CGAAGTCGAT ATCAATGATG TTACCGCCAT GTTCATACTT AGGTTTGTCT	5880
35	TTTTCTGTAT CTTCTCTGAA TGACTGATTA CCTTTATTTT GACCATGAAT TTGAGGTACA	5940
	CTATCAAAAT CGATATCTAC GATATTGCCA CCTTGTTCAAT ATTTAGGTTT GTCTTCTTCT	6000
	GTGTCTTCCT CGAATGACTG GTTACCGCTA TTTTGCCAC CTTCATAACC TAATTCACCTC	6060
40	TTAATATCAA CGTGGCTATT TTCTTCGATT TCTTCAATCA CGTCATAATT CCCGTGACCA	6120
	TTTTCAGTTC CTAAACCAGA ATGAGAAATA TGATGATTGT TTTTAGTAAT TTCCTCGACT	6180
	GGTCCTTGTT CTTGACCATG CTCTTCAGGT AATTCATCCA CTAATTCAAT CAGATTACTT	6240
45	TCAGTTGTAT ATTCTTTTCGT ATCTTCAACT GTTGTATGAT CGCTCAGTGC GCCAGTTACA	6300
	ATACCTTTTG TAGACTCTTC GTCAAATTCA ACTAAGTTAG ACTCAGTAGT AACCTGACCA	6360
	CCACCTGGGT TTGTATCTTC TTCATATTCA ACAACATCAG CGTGATGTTT TGAATTTTCA	6420
50	TGTGTAGATT CTTCAAAGTC AATTGGATTT GATTCCTCAG AGGACTCAGT GTATCCTCCA	6480
	ACGTGACCTG CTTGCTATC CACAGCAGTA TGGTAATCGA TATCAATAGC TGATGAATCC	6540

55

	TGGAATCAA TGTCAAGAGT TGATGAATCA TATTCCTCTT CAACAGTAGT TACTAAATTC	6660
	TTATCATATT GACCTGTAAG AGTTTCTTTA ATTGTATCTT CTTTATATTC AAATTTATTA	6720
5	TTTTGAATAA TCGGACCATT TTTCTCATTT CCGTTCGCTT TATTACTGTA TAAAACTAAA	6780
	CCATTATCCC AAGTTAAGGT ATATCCTCTA TCATAATAAT ACTTATAAAG TTGCTCTGGA	6840
10	TGTCCTACCA TTTGTGTTCT AAAATCAACT TCATCAGTAC CATTTAAATA CTCTCCATCA	6900
	TAGTGAACAA CATAAGTTTT ATCTAGATTT TCTATATTCA ATGAATAGCT TCCATTATTT	6960
	TGTAAATTCA AATTCCTACT CATATTACTT GTGACTTCTT TAAATTTAGA AGTATCTGTC	7020
15	GTATTTGTCAT ATACACTCTT CGCTATGTCT TCATTATTAC CCAAGTATTC AAATATCCTA	7080
	ACTTTTGTTT GATTTCCATT CTGATTACTA CCTTTCATTA AAGTTCAGT AACAGTCACA	7140
	CTTGTCGTTT TACCATTATT AGGTTTAATA AATGCAACAT GCGAAAATCT ATTATTGCT	7200
20	TTATTAAATG TCTCAATCGA TCCATTTAAA TTGGCATAAT AATTCCTAAT ACCATCTTTA	7260
	TATTTAACAT CTAATTCCTT TGAAGTTTGT TCTTCATTTA GTGTTGAAGT TATAGTTTGA	7320
	TTTCCATTAG TTTGTACAGT TTTAGGATCA ATAAATAAAT TAATTTCTAG TTCAGCCGTT	7380
25	ACATCAACCT TATCTTCAAT ATCATTTGTA AATGTATATC TAATCTTTCC ACCTTCTAAA	7440
	ACTTCACCTG TCGCCATTAC GACTGAACCA TTTTAAATTT CTGGTACTTT TCTAGCAGTT	7500
	GATACGCCAT GCGTATTTAC ATTATTTGAT AAAGTAAAGT CAAAGTAGTC ACCTTGATGT	7560
30	AAACCATTCT CAAATTTCAA CTTATATTTT AGTACCGCTC GTTGTCTGTC ATGAGGTTCT	7620
	ACTTTATTTG TATTGTTATG CCCCTCAATA GAACCAATTT CTAAGTAAAC TTTACTTGTT	7680
35	ACATCTGTAC CCGTTTCCAC TTTGCGGTTA CTAGCTTCCT TAGCTTCCGC TACATCTGCT	7740
	GATCTTGTC CACGTGGCTT ACTTTCTGAT GCCGTCTTG GCTGTGCCAC TTCAACTTGT	7800
	GTTTCTGCGA CTTGATTTTG TGTAGCCTTT TTAGGTGTTA AATCTACTTG TCTTTGATCT	7860
40	CCGCTATTGT CTTGAGATTG TGTGTTTCC TTAAGTTGAG GTTTCGCTTC TTCCTTAACT	7920
	ACCTCTTCTT TAACTGTTTC TATATTTGCT GGTGTGTCAG TTTGTGGTGC TTGACTGCT	7980
	TTGGTGCTT CTTCAAGTTG TACTGTGTT GCGTTTGACG GTTGTCTGT TACTGTTGCG	8040
45	TTATATGATT GAGTTTCTTC TATATGATTA ACGTTAGTTG CAGTTGTTTG TGTTTCACTT	8100
	GTTTTATAT CAGTAGCTGA ATTCCCATTT TCTTCTACTG TAGTTGTCTT TTGTTCTGAT	8160
	GCTGCAGCTT CTTTGTCTTG TCCCATCCCA ACAACGATCA TTGTTCCCTAA GAATACTGAT	8220
50	GCTGCTCCCA ATTTATGTTT TCTAATGCCG TACCTAAGAT TGTTTTTCAC TATAATATCT	8280
	CCCTTTAAAT GCAAAATTC TTAATTTTTT AAAGTTAATA AATGCAAGTC TATATTGTTT	8340

55

	ATGTTAATTG ATAATTTTAT TATTTGAAAT ATACCTATAA ATTGTATTCA AGTCATCAGA	8460
	AACCCCTTGT ACACAAGGCT TGTATTTTTT ATACTTATTT TTAAATTAA ATTCATCATT	8520
5	ATCTAATTTA AAACAATATA CTAAACGTTT CATAATTATC GCCTGTACAA TACGCACAAA	8580
	AACATGTCTT GAAACGCCTT TCATTACTCT AAAATACCCA ATATACTTTT TATATCGTTC	8640
10	GGATTCTGAG TATTTGAGAC GATTTTCTGC ATAAAAATAA ACGTGTTC AAGCAATATA	8700
	TTGCAATTAC CTAAAAACAC GTTTACTTAA TATTTAGTTA AACAAATAAG CTAATGAATA	8760
	AAATGAAGAT GATACCTGAA ACGGAAATAA TCGTTTCTAA TAATGACCAT GTTAAGAATG	8820
15	TTTCTTTTAC AGTTAAACCA AAATATTCTT TAAACATCCA AAATCCTGCG TCATTTACAT	8880
	GAGACAAAAT CACACTACCT GCACCTATCG CAAGTACAAC TAATGCAACA TTTACATCTG	8940
	ATGATTGTAA TAATGGTAAG ACAATACCTG TAGTTGAAAT CGCAGCTACT GTAGCCGAAC	9000
20	CTAATGCGAT ACGTAGCACA GCTGCAACAA TCCATGCTAG TAAAATCGGA GACATCTCTG	9060
	TACCTTCAAA CATTTTAGCA ATTGTATTTT CGACACCGCC GTCAATTAAT ACTTGTTTAA	9120
	ATGTACCGCC ACCGCCAATA ATCAATAACA TCATTCCGAT TGGATAAATC GCATTCTGCA	9180
25	CTGATTCCAT AATATGATTG ATCTTACGCT TTCTCATTAA TCCCATCGTA ACGATTGCAA	9240
	ATAATACTGC TATTAGCATG GCTGTCCCTG CTGTTCTTAT CATATAAATG ATAGATTCAA	9300
	ATAGATTTGT AGGTTTGTCA TGCCAGTTA CAAGTTGCGT TATCGTAGAC ACTAACATTA	9360
30	ATATGACTGG TAATGTTGCT GTTAATAAAC TCATACCAA TCCTGGCATC TCTTGATCCG	9420
	TAAATTCTTT TTGTGCACCT AACGCTGAAA TATCGCCTTC TCGTGATAC GCAGACGGAA	9480
35	TCATTTTTTG TGCACTTTGT TAAATATAGG CCCTGCAATG AGTGTAATG GAAATGGCAAT	9540
	AATCATACCA TACAGTAATA CATCTCCAAC ATTTGCCTTT AATTCTTTTG CGATGACTAC	9600
	CGGTCCTGGA TGTGGTGGTA AAAAGCCATG TGTCATGAT AAAGCTGTTA CCATAGGTAG	9660
40	TCCTAGTTTT AACACTGAAA CATTGCGCG TTTTGCTACT GTAAATACTA ATGGAATCAG	9720
	TAAGACTAAA CCTACTTCAA AGAACAATGC AATACCGACG ATAAATGCTG CAACAAGCAT	9780
	TGCCCATGTG ACATGTTTTT GACCAAATTT TTGAATCAAC GTGTCTGCGA TTCGAGTTGC	9840
45	ACCACCACCA TCAGCAAGCA ATTTCCCAAG TATGGCACCT AAACCGAATA TCAGTGCAAT	9900
	GTGGCCGAGC GTACTGCCCA TTCCTTTCTC AATCGTCTCC ATAATTTTAG TCAATGGTAT	9960
	ACCTAGCATT AACGCTGTAA TCATCGATGT GATAATTAAT GAAATAAATG TATTTAATTT	10020
50	AAACCCAATA ATTAATACTA ATAAAATAAC GATACCTAAA ACAACACTGA TTAACGGCCA	10080
	TATTTGTTA AACATGACAT TCCCCTCTTT CTCTTTTCAA TAGAATGTAA CACCGTCGTC	10140

55

	GAGTGACGTA	TTTATTGTGT	TTTATTTTCA	GCGATATGTT	GGCGTTGAAA	ATCTGCAATT	10260
	TGTTTCATAAT	TCTCTGTAA	AGAACGACTT	AAATTGATAA	AAATGGATAC	GATCTCTTGG	10320
5	TAAACAGTGA	CATTTTCTTC	AATCGGCGTA	TGATTGTTTG	TGGCACCGAC	CATCGATGAA	10380
	ACGATTGAAA	AATCTTCAAT	GTCACCTACA	GCTTTAAGTC	CGAGCACGCA	GGCACCTAAG	10440
10	CATGAACCTT	CATAACTTTC	AGGAACCACT	AATCTGTGT	CAAATATATC	TGACATCATT	10500
	TGACGCCATA	CTTCACTTTT	CGCAAAACCA	CCTGTTGCTT	TTATCATCTT	AGGTGTTTCA	10560
	TTCATTACTT	CAATAAGCGC	AAGATAGACG	GTATACAAAT	TGTAAAGAAC	ACCTTCTAAT	10620
15	GCAGCGCGAA	TCATATGTTT	TTTTTTATGA	GATAAAGTTA	AACCGAAGAA	TGAACCTCTT	10680
	GCATTGCGT	TCCAAAGCGG	CGCACGTTCT	CCTGCTAAAT	AGGGATGGAA	TATTAAACCA	10740
	TCTGCACCTG	GTTTAACACG	CTTTGCAATT	TGAGTTAAGA	CATCATAAGG	ATCAACACCG	10800
20	AGACGTTTCG	CAGTTTCGAC	TTCACCTCGT	AGCAACTCGT	CGCGCAACCA	TCTCAATACG	10860
	ACACCACCAT	TATTTACAGG	ACCTCCGATG	ACGTAGTGGT	CCTCTGTAA	GACATAACAA	10920
	AATATTCTAC	CTTTGTAATC	AGTACGCGGT	TTATCTATCA	CAGTACGAAT	CGCCCCAGAT	10980
25	GTACCGATTG	TGACAGCAAC	TTCTCCTTTA	CCAACACTAT	TGACACCTAA	ATTAGAAAGG	11040
	ACCCCATCAC	TCGCACCAAT	AACAAACGGT	GTATCTTTAT	TAAGCCCCAT	TAATGTTGCA	11100
	TAACGTTCTT	TCATACCTTT	CATCACATAC	GTTGTTGGAA	CTAATCCGG	CAACATTTC	11160
30	TTGGAAATAC	CCAGCAGTTC	TAATGCCTCA	ACATCCCAAT	CTAATGTTTC	TAAATTAAAC	11220
	ATCCCTGTTG	CGGAAGCCAT	TGAATAATCA	ATGATATATG	TATCAAATAA	ATGATAGAAA	11280
35	ATGTATGTTT	TAATATCTGC	AAACTTAGCA	GTACGTTGAA	ATACATCTTG	CCATTTCATG	11340
	TTCATCCAAA	AAATCTTCGC	TAATGGCGAC	ATAGGATGAA	TCGGTGTGCC	TGTTGCTGG	11400
	TAAATCGCAT	TGCCATCATG	CACCTTCATT	ATTACTGTTG	CATATTTTGC	AGCGCGGTTA	11460
40	TCTGCCCAAG	TAATATTATT	TGTTAATCTT	TGATGTTGCT	GATCCATCGC	AATCAAGCTA	11520
	TGCATTTGCG	CACTAAATGA	CACAACTTA	ATGTCGTCTT	TATTAACCTT	GGATTCTCTC	11580
	ATAACATATT	TAATAGTCAT	TAGTACTGCA	TCAAATAATT	CATCTGGGTT	TTCTTCTGAG	11640
45	ACATCAACGT	TTGGTGTGTG	TAAATCATAG	CCTATTTGAT	GTTTCATGAT	AAAAGTTCCA	11700
	TTTTCATCAT	ATAAGACTGA	CTTGGTACTC	GTCGTTCCAA	TGTCGACACC	AATCATATAT	11760
	TTCATGATAA	ATCCTTCTTT	CTTTCATTTT	AATTCAACCA	AAATCCTTCA	ATATCTTTAC	11820
50	CAACATCGTC	GAAATTTAAA	TGAAACGCTT	CTTTCAAAAT	TTGACTGTCG	TATTGTTCCA	11880
	CTGCATCAAT	AAACACTTGA	TGATTATGAT	GTATGCGTTC	AAAATCTTGC	GGGTTCTGTT	11940
55							

	AAAATGAGTT TAAATATTGA TGATTAGATG CTTTGATTAA TGTTTCATGA AATTCAAAGT	12060
	CATGCTTCGT AAATGATTCT GCATCCTCAA ATTTTACTGC CACTTTCATC ATTTCAAGTT	12120
5	GTTTCTTCAT TTCTTTTACG ATAGGTAGTC GCTCTTGATT TTAACTCTT GAAAATGCAA	12180
	ATGACTCTAA CATCAGTCGC AAATCATACA TTTCTTTCTT TTCTTGTTCC CCAAACGGCA	12240
10	ACACATGTGC ACCCATTTCTT TCTAATTGGA TGAGTTGATT TTGTTGCAAT AATTTAAATG	12300
	CATCTCGAAT TGGCGAACGA CTCACATTAA ATTGCTTTGC CATTTGATTT TCAGTGAGTA	12360
	ACGTACCTTC AGCTATGTGA CCATTCACAA TGCCTAAGCG TAATTCTGCC GCGATACCTT	12420
15	CTCCAGTTGT CATACTTCC AACCATTCTT CTGGATATCC ATACATCATC AAAGTCACTC	12480
	CTTCATTACA CGACATACTT GTATACAAGT ATGTTAATAT AGTTATTATG AGTTTGCAAG	12540
	CGCTTTCTTT ACGAGCACTA AAATAGTGAC CACCCCTTTT CGATTTAAAT TTAAAGGAAA	12600
20	TGGTCACTAT CACACGAATG ATTAAATTGT TATGTTGTAT GTGGGATATT TCTAATTGTT	12660
	CTGTACTCAT ATGCGCTTTA GGTACTTCAA TGCAATAATG CGTTTCATGA CAGTTTGGAC	12720
	ATTCGAATCG ACGTGTGTGTC GCTGTATGTT TCGCTTTGAT AACTGCCCCAC AAAGATGGTG	12780
25	AGAATATATG CTGGCAGTTA GGACATAAAT AGGCAACCTT TTGTTGGTAA TAAAAAGTAA	12840
	CACCAATGCC ATAACCAATC ATAAATGGTA AAGCAATTAA AAACGGCCAT TTATTTTTCA	12900
30	TCAAAATTGC ACTTATAATG CTAGAATATT GAATTATTCC TATAATACCA GCACTAATCC	12960
	AAATGTTACG ACGAATACTT TTCATTTTCAG CTGATTTACT CATGACATGC TCTATGTCTT	13020
	TTAAGTGTGT GATTGGAGAC GTCGACGCTT CATTTACGTA ATATTGAACA TTTTAAATTT	13080
35	TGTTTAAATAC CGCTTGTTGC TGTTTAACTT GTTGGTTAAT TTCTTGTTGT TTCATAGTTA	13140
	GTAAAGTATT GAGCGTCTTC AAAGTACCTT CACCTTTTAG CAACATATCT ATATCGCTTA	13200
	ACGCACAACC TAAATCTTTA AGCAATAAGA TTAACCTCTAA TGTTTGTCGC TGTTGTTCTG	13260
40	TATACACACG ACGCTTTCCT TCTGTAAATC CTGTGGTTT CAAAATACCT TTGCGATCAT	13320
	AATATTGAAT CGTTCGTGTT GTCACATTGC ATAATTTTGC GAGTTCTCCA GTCGAATAGT	13380
	TAGACATAGA TTCCACCTCC TATAATTACC ATAGTTGATG ACCCGACGTC ACGAGCAAGT	13440
45	ACAATTCCA CATTTTAAAG AAATTTATTA TACTAGGCGT CTTATTTTTA TGATTTGTA	13500
	CCATGTTGAT TTACAAACTC ACTCAAATA AGTAACACAC CTACTAAACA TCTACTCTGT	13560
	TATTTGAGAA TGAATTTGTT GTAATTTATC TTCAACTTCA GTAATCTCTG TCGCACATTC	13620
50	TTTCAGTAAA TCTCGATACT TTTCCGTCTC TGCATTGTTT TTATAACGTA TTTTATGTTT	13680
	TAACTTGcC CACATATCCA TACCTATCGT TCTAATTTGA ATTTCAACAG GCAATACCTC	13740

55

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1059 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

```

GGATAAGTTC AGGTAAATTC ATTTCTTTTT CAATTTTGAT TTTCATTGTT TCCGCCCTTT      60
TAAAATAAAG TTAGTTGCTT CTGTTCTCA TATTCCAAAT CACTTTGCTT TATATATGTT      120
TCAAGCTCTT CCGCTGTATC AAATGTCTTT TTCACACCTT GCCAACCTGG CACGATATGA      180
CCGTGAAAGT AATAAGTGCC ATTTACTACA TGGATATGTG CCACTCGTTC GTTATCCTGA      240
TACAGATATC TCTTAGATCC AAAGAATTGA TTTAGGTATT CTTTACGCGC GCTATCTGTC      300
ATGGTCATCA CTCCTTTTAA CAATTAGGCA GACCAAACGA CATGCATTCTG TCGTATAGCT      360
CTTCATTACT TATGCTTGCC TTATAGTTTT CAATCACATT GCTAACTTCT TTATGACTCA      420
TTGCTTTAAC TTGTTCTGCT GTATATTTTT CGCAGTCTTC TAATTCCAGT TGCTCCTGTA      480
ATGACATCAC ATATTCAACT TGTCTTTGGG TTGCCATCGT TAACCCTCCC ACAAGTCAAA      540
AGCTCTTTGG ACGTAAAACT TCGCCTTTGC TAAATCCTCA TGACCATTCT TTAACGGTGC      600
TCTAGACATG TATTTGATTG CATTACCTAT TGCGAATGCT AGTTGAGGTG GATACTGTGC      660
CGTAACCTGT TCGATAAAAT CTATAATTC AATGTCGCCG TATGTGTAGT GCGCTGGTTG      720
CTTAACATTG TCTTGCGCTT CGTTCATATC TACTTTTCTG TTACTGATTA CGCTCATTAT      780
GCTTCACTCC ATTTCTTGAA CATTTGGTTA TAAGTGACAT CGAACCAGTA CGGATCACGT      840
GAATGTTTTT GTGGCGTTCC ATCATAAAGC CATGGTCTTA ATCTTCTCTT TCTTCTCTGT      900
TCATATTCCG CTCTCACATT TCGTTGGTAT CGGTTCAAAA TCGCTTTTTT TCTGATTTTT      960
TCTCTCCCTT TTTCTTCATC TTTnATtTGA CTCTnCATAT ATTCAACTTC TTCTGTAGAT     1020
nTTGAGTCCT TTCTTCCACA CAATAATTCA nCGCCGCGC                               1059

```

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30246 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	GAAGTAAAAG AAGAATTAAA TTAAACATTA ACAATGGATG AAATTGAATA TGTCGGGACA	60
	ATTGTAGGTC CTGCATATCC ACAACAGGAT ATGTTAACTG AGTTAAATGG ATTTTCGCGCA	120
5	TTAACCAAAA TCGATTGGGA AAACGTAAC TCAATAATG AAATTACGGA TATACGCTGG	180
	ATTGATAAAG ATAATGATGC GTTGATTGCG CCTGCTGTCA AAGTTTGGAT TGAAACTTAT	240
10	GGTGGTAAAC ATGACAAATA ATGACACCAT CATGTTACGA CATTATGTCC CACAAGATTA	300
	TTCGATGTTA GAAGCTTTTC AATTAAAGTGA AAGTGATTG AAGTTTGTTA AAACGCCAGA	360
	GGAAAATATT ACAGCTGCAA TGTCTGATAA TGAAAGGTAT CCCATCGTTG TAATGGATGG	420
15	CAGGCAATGT GTGGCCTTTT TTACATTACA TCGTGAAAAA GGGGTCGCAC CATTTAGCGA	480
	TAACCAAGAT GCAGTATTTT TCAGGTCATT TAGTGTTGAT CAACGTTATC GTAATAGAGG	540
	AATAGGTAAA GTGGTAATGG AAAAAATGGC GTCATTTATC ACTTCAACAT TTCAGGATAT	600
20	TAATGAGATT GTGTTAACGG TTAATACTGA CAATCCACAT GCCATGGCAC TTTATCGCCA	660
	ACAAGGATAT CAATATATGG GAGATAGTAT GTTCGTCGGA AGACCTGTTC ATATTATGGC	720
	GTAACTATA AAATAAATTA AATTTAAAAG CATCTTTACT CATCGTCGAC CACAACAATT	780
25	AATGATGAAT AAAGGTGCTT TTTGTTATAG ATCATCGGAC AATTTACTAT AGTAAAAAGC	840
	GACCTAGTGA ACAATTGACA TATATCCACA GGTGCGTTAA CTTAAGTTAT ATTGCTAGTT	900
	GCGATTAATT GATAGACTCA TCATTTTTCG GCTGTCGAGA TGGTCTTTTT ATTAATAAATG	960
30	CCGTAATCCA AGCCGTAATC GGAATACTGA TTGCAACGGC AATACCGCCT AAAATAATAG	1020
	AAATAAATTC TTGGGCAAAT ATTTTCGAGT TTATAATATG ACCAAATGAA TATTTAAGTT	1080
	TGAAAAACCA AATAAATAAA GCAAGTTGGC CACCAAAAAA GGCAAGGTAA ATCGTGTTCC	1140
35	CAGATGTCGC TAAAATTTCT CTACCAACAC GCATGCCAGA TTGAATAAT TCGTATTGCC	1200
	TAACGTTGGA TTCACTTGAT GCAATTCATA AATGGGTGAA CTAATGGTAA TTGTTAAATC	1260
40	TATCACAGCT GCAATAACAG CAAGAATAAT AGTGAACACC ATAAATTGAA CCATATCAAT	1320
	GCCAATATTC ATTGAATACA CATATGTTTC ATCTTGTTGT TCGGTTGAAA AGCCTTGTAG	1380
	ATGACCGAAG TAGACCGATA AATAAATGAG TGTAATCAAC AATATTGTTG TAACGATAGT	1440
45	GCTGgATAAA TGCaGCTTGT GTTTTAAACAT TGTAAC TATT GAGTACGAAT AAATTACAAG	1500
	CGCCAATAAT AATGCAGAAA AAGAATGTGA CGACATAAAT CGGTACGCCA AAAATAATCA	1560
	ATACAATACT AATAATTAAA ATAGCGAAAT TTAAAAATAG GGTTAAATAA GAGATGAATC	1620
50	CCTTTTTACC TCCGAAAATT ATCATCAGAA AGAGGAGCAA TAACGCCAAT ATAAATACAG	1680
	CATTCATTGT TTCGCCCTCC TTAATGTTTC AAATATTTCC ATAAACAATA TTGTGATAGG	1740

55

EP 0 786 519 A2

	CATCGAAATA GTATAAGTCA CTGTATTGGC ATTTTTTAAA AAGATTAAAA ACATAGGTAG	1860
	TGCACCGGAT AAATATGAGA ATAATAAGAT GTTAGTCATT GTTCCCATAA TATCTTGGCC	1920
5	GATGTTTCGC CCAGCAAGCG CCCATCTCCT CATTGAAATG TGTGGCGTAC GCTGTAAAT	1980
	TTTCATGCATA CCACTAGCAA TTGTAATTGC AACATCCATA ATAGCGCCAA GTGAACCTAT	2040
	TAACACTGAG GCTAGGAAGA TATCTTTCGG TGGTAATGAT AAAAAGTTCA TCGTTTCATA	2100
10	TTTAATGCCT TTACCATCTG TCATATATAT GATTAAATCT GTTAAACCTA TACTCAAAAA	2160
	AGTTCCGATA ATTGTACTGG CTATGGTAAT GAGTGTACGC ATATGCCAGC CTGTAACGAG	2220
15	CAATAAAGTG AGTATTGTTG AACAGATCAT GGCAATGGTC ATGAGTAAGA ATAAATTAAT	2280
	ATTGCTATGT TGAATATGAA TGTAATTGC GATTAAATATG GCAATAGAAT TCAAGATTAA	2340
	CGATAAAATC GATTGCAGTC CGACTTTGCG ACCAACCAAT AATACAGTTA ATAAGAACAA	2400
20	ACCAGTGATG ATAACCGTTA AGGTATCAGC CTTCTTTTCT ATAATATAAG CATCACTCGG	2460
	CTTGTTAGAA ATATGTAATA ATACTTTTTC GTGTGTGCGA AATGCCTCAG AATCTGCTTG	2520
	CGATTTGACG TACTGATGAT TAATCGTCGT CGTTTCTCCA GCAAATTGAC CATTTAATAT	2580
25	TTTGACTTTT AATTGATTTT TATATTTAAT ATCACGATTA TTTTGTGCAT CTTTGTAGG	2640
	TGTCGAAGAA ACATGTTTGA CATCTATAAT TTGACCAATT GGTGTGTGT AAAAGTTCTC	2700
	ATTATTGAAT GTAAATAAAA TAGCACCAAT GAATGCGATG CAGAACAAAC CTAAAATTAT	2760
30	ATTAAATGGC TTTGTAAATA AATTTCTATA TTTCAAAAC AAAACCCCAA TTCTATGAAT	2820
	GAATTAATAT GGTGATTATA CGCCCTTAAT TTTTATTTT CAAAGATATT ACTGCTAAGT	2880
	GTAAACGAA AATCATCATT GATAGCATCG AATTACTTAA TGGAATGTAG ACGTTTTAGT	2940
35	CATTAATTGC TGAATAAGTG TTAATAATAT GCCAATATCA CTCCTTGTAT AAGGCTCCTT	3000
	TGTAATAGCA CATATCGTTC TTTTAAATTC AGTATGATCT AATTTTATAT CTATCCATGA	3060
	TTTAGATTCT GGTAAATGTA TATTTTGTGA TGAAATGATG TAACCTTCTT TTTGACGAAG	3120
40	GAGATACTGC GCAAGTGGTT GGCTACTGAT TGTGTATACA TCTGATTTAG TAATCTTGCG	3180
	CAATTGTTTT TTTACAGTTT CGGCAAATGG TGCCAAGCAA TAAATATGAC TATGCTCAAA	3240
45	CTGAATTAAT GGTGGGTGTG TCGCCATCGT AATTGGATCG TCTGAAGGCG CATATAAATG	3300
	ATAGTGCTCT TCGAATAAAG GTAGCATATG TAATTGTTTG TGTTTACGTA TTTCTGGTGT	3360
	AAGTTCCGTG AAACCAATGT CTATATCCC ATTTAATACG CTATTTATAA TTGTGTCATG	3420
50	TTCTAATAAG CTCGGTATGA CATGTGTATC ATTTTGTAAG TGAAACGTTT GGATAAGTGG	3480
	TAGTAACATG TGGGATACGT CACTCTCATC ATAGCCAATG TAGATACTTT TATTTTTAGT	3540

55

	TTCATTAAAT AATAATTTCC CTTGAGATGT GAGCGTAATA TTGCGTCCTT GCTTTTTTAAA	3660
	TAAAGACACA TTAAGTTCTT GTTCTAATAA TGTAATTTGA CGGCTTATCG CTGATTGAGC	3720
5	AATGTTTAGT TCAAGTGCTG TTTCGGAGAT ATGTTCTCTT TTAGCGACCT CGATAAAATA	3780
	TCTTAATTGT TTAATTTCCA TAGCGATATA GGCACCTCCA AAAATGAGTG TTTTGTAAC	3840
	ATTATAGCAA TATTATTGAT AAATGTTCTA TTTTITAGAT GAATATCTTC TATTTTATAT	3900
10	ATTGAACAGA TAAATTTTTT AGATTATAGT AATTATCATT AATAACTAAT ATCAGAATAT	3960
	TCTAAAAAAG GGGTGTGCAT CATGCACAAT GAGAAATTAA TTAAAGGCTT ATATGACTAT	4020
	CGTGAGGAAC ATGATGCGTG TGGTATTGGT TTTTATGCGA ATATGGATAA TAAAAGGTCT	4080
15	CACGACATCA TTGATAAATC GCTTGAAATG TTGCGACGCT TAGATCACAG GGGCGGGGTC	4140
	GGCGCAGATG GCATCACTGG TGATGGCGCA GGTATTATGA CTGAAATACC TTTTGCAATT	4200
20	TTCAAACAAC ATGTAACGGA CTTTGATATC CCAGGTGAAG GTGAATATGC CGTGGGGTTA	4260
	TTTTTTTCCA AAGAACGCAT TTTAGGTTCT GAACATGAAG TAGTTTTTAA AAAATATTTT	4320
	GAAGGCGAAG GGTATCAAT TCTTGGTTAT CGTAATGTAC CAGTTAATAA AGATGCCATT	4380
25	GCTAAACATG TAGCAGATAC GATGCCAGTC ATTCAACAAG TGTTTATTGA TATTAGGGAC	4440
	ATTGAAGATG TTGAAAAGCG TTTGTTTTTA GCGAGAAAAC AATTAGAGTT CTATTCGACT	4500
	CAGTGCGATT TAGAATTGTA TTTTACGAGC TTATCACGCA AAACAATTGT ATATAAAGGT	4560
30	TGGTTACGAT CAGACCAAAT TAAAAACTA TATACAGATT TATCGGATGA TTTATATCAA	4620
	TCAAAGCTAG GGTAGTGCA TTCGAGATTT AGTACGAATA CATTCCCGAG TTGAAAAGG	4680
	GCACATCCTA ACCGTATGTT AATGCATAAT GGTGAGATTA ACACGATTAA AGGTAATGTA	4740
35	AACTGGATGC GAGCAGCCA ACATAAATTA ATCGAAACAT TATTTGGCGA GGATCAACAT	4800
	AAAGTGTTC AAATTGTGCA TGAGGATGGT AGTGACTCTG CCATTGTAGA TAATGCGCTA	4860
	GAGTTCTTAT CGTTAGCCAT GGAGCCAGAA AAGGCAGCGA TGTTACTCAT ACCTGAACCT	4920
40	TGGTTATATA ATGAAGCGAA TGATGCAAT GTACGTGCGT TTTATGAATT TTATAGTTAT	4980
	TTAATGGAAC CGTGGGATGG TCCTACAATG ATTCGTTCT GTAACGGTGA CAAACTTGGC	5040
	GCGCTTACAG ATAGAAATGG ATTACGTCCA GGTGTTATA CGATTACTAA AGATAACTTT	5100
45	ATTGTCTTTT CATCTGAAGT GGGTGTGTG GACGTACCTG AAAGTAATGT TGCTTTTTAA	5160
	GGTCAATTGA ATCCTGGAAA GTTATTGCTT GTTGATTTTA AACAGAATAA AGTCATTGAA	5220
50	AATAATGATT TAAAAGGTGC GATTGCTGGA GAATTACCAT ATAAAGCGTG GATTGATAAC	5280
	CATAAAGTTG ACTTTGATTT TGAAAATATA CAATATCAAG ATTCGCAATG GAAAGATGAG	5340

	CAGGAACCTTG TAGAAGGTAA GAAGGATCCT ATCGGTGCAA TGGGATATGA TGCGCCAATT	5460
5	GCAGTGTTGA ACGAGCGACC AGAATCACTA TTAAATTACT TTAAACAGCT GTTTGACAAA	5520
	GTTACGAATC CACCAATTGA TGCATATCGT GAAAAAATCG TAACGAGTGA ACTTTCCTAT	5580
	TTAGGTGGCG AAGGTAACTT ACTAGCACCT GACGAAACGG TTTTAGATCG TATTCAATTG	5640
10	AAAAGGCCCG TATTGAATGA ATCACACTTA GCAGCGATTG ATCAGGAACA TTTTAAATTA	5700
	ACTTATTTAT CAACGGTATA TGAAGGGGAT TTGGAAGATG CGTTAGAAGC ATTAGGCCGA	5760
	GAAGCAGTGA ATGCTGTAAA GCAAGGCGCT CAAATTCTAG TGTTAGATGA TAGTGGATTA	5820
15	GTTGATAGCA ATGGCTTTGC AATGCCGATG TTAATCGCAA TAAGTCATGT GCATCAATTA	5880
	CTTATTAAAG CAGATTTACG TATGTCTACA AGTTTAGTCG CTAAATCTGG TGAGACACGA	5940
	GAAGTGCATC ATGTTGCTTG TTTACTCGCA TATGGCGCGA ATGCAATTGT GCCATACCTA	6000
20	GCGCAACGTA CAGTTGAACA ACTGACATTG ACAGAAGGGT TACAAGGCAC CGTTGTCGAT	6060
	AATGTTAAGA CATATACGGA TGTATTGTCA GAAGGTGTCA TTAAAGTAAT GGCTAAGATG	6120
	GGAATTTTGA CAGTGCAAAG TTATCAAGGG GCACAAATAT TTGAAGCGAT TGGCTTGTCT	6180
25	CATGATGTGA TTGATCGTTA TTTACTGGG ACACAGTCTA AGTTATCTGG TATTCGATT	6240
	GATCAAATTG ATGCTGAAAA TAAAGCACGT CAACAAAGTG ATGATAATTA TCTTGCATCA	6300
	GGTAGTACAT TCCAATGGAG ACAACAAGGT CAACATCATG CTTTAAATCC GGAATCTATT	6360
30	TTCTTATTGC AGCACGCATG TAAAGAAAAT GACTATGCGC AATTTAAAGC ATACTCTGAA	6420
	GCGGTGAACA AAAATAGAAC AGATCACATT AGACATTTAC TTGAATTTAA AGCATGTACA	6480
35	CCGATTGACA TCGACCAAGT TGAACCGGTA AGTGACATTG TCAAACGCTT TAATACAGGG	6540
	GCGATGAGTT ATGGATCGAT TTCAGCGGAA GCACATGAAA CGTTAGCACA AGCCATGAAC	6600
	CAATTAGGTG GAAAGAGTAA TAGTGGTGAA GGTGGCGAAG ATGCAAAACG TTATGAAGTA	6660
40	CAAGTTGATG GAAGCAACAA AGTAAGTGCG ATTAAACAAG TTGCTTCTGG GCGTTTTGGT	6720
	GTAAGTAGTG ATTATTTACA ACATGCCAAA GAAATTCAAA TTAAAGTTGC GCAAGGTGCA	6780
	AAGCCTGGTG AAGGTGGTCA ATTACCTGGT ACTAAGGTAT ATCCGTGGAT TGCGAAGACA	6840
45	AGAGGGTCAA CGCCAGGTAT CGGTCTGATT TCACCACCGC CACATCATGA TATTTATTCA	6900
	ATAGAAGATT TAGCGCAACT GATACATGAT TTGAAAAATG CGAATAAAGA TGCAGATATC	6960
	GCGGTAAAAT TAGTTTCGAA AACAGGTGTT GGTACCATTG CATCTGGGGT GGCAAAAGCA	7020
50	TTTGACGATA AAATTGTCAT CAGTGGTTAC GATGGTGGTA CAGGGGCTTC ACCCAAACG	7080
	AGTATTCAGC ATGCCGGTGT TCCTTGGGAG ATTGGTTTAG CAGAAACACA TCAAACATTA	7140

55

	AAAGATGTAG	CGTACGCATG	TGCGCTTGGA	GCGGAAGAAT	TTGGATTTGC	AACTGCACCA	7260
	TTAGTGGTGT	TGGGCTGTAT	TATGATGCGT	GTATGCCATA	AAGATACATG	TCCAGTAGGA	7320
5	GTTGCAACTC	AAAACAAAGA	TTTACGTGCT	TTATATAGAG	GTAAAGCACA	TCATGTTGTT	7380
	AATTTTATGC	ATTTTATTGC	ACAAGAATTA	AGAGAAATTT	TAGCATCTTT	AGGTTTGAAA	7440
	CGTGTAGAAG	ACTTAGTTGG	AAGAACTGAT	TTATTACAAC	GATCATCAAC	ATTAAAAGCG	7500
10	AATAGCAAAG	CGGCTAGTAT	TGATGTTGAA	AAACTGTTAT	GTCCTTTTGA	TGGGCCAAAC	7560
	ACAAAAGAAA	TTCAACAAAA	TCATAATCTT	GAGCATGGAT	TTGATTTAAC	AAATTTATAT	7620
	GAAGTAACGA	AGCCATATAT	TGCTGAAGGG	CGTCGCTATA	CAGGTAGCTT	TACAGTAAAT	7680
15	AATGAACAAC	GTGATGTAGG	GGTTATTACA	GGTAGTGAGA	TTTCGAAACA	ATATGGAGAA	7740
	GCAGGACTTC	CTGAAAATAC	AATTAATGTT	TATACGAATG	GTCATGCTGG	TCAAAGTCTT	7800
20	GCAGCATATG	CACCGAAAGG	CCTAATGATT	CATCATACTG	GAGATGCGAA	TGACTATGTT	7860
	GGTAAAGGAT	TATCTGGTGG	TACGGTCATT	GTCAAAGCAC	CTTTTGAAGA	ACGACAAAAT	7920
	GAAATTATTG	CTGGTAACGT	CTCATTCTAT	GGTGCACAA	GTGGTAAGGC	ATTTATTAAC	7980
25	GGTAGTGCAG	GAGAAAGATT	CTGTATTAGA	AATAGTGGTG	TAGATGTTGT	CGTTGAAGGT	8040
	ATCGGCGACC	ATGGATTAGA	GTATATGACT	GGTGGACATG	TCATTAATTT	AGGTGATGTA	8100
	GGTAAGAACT	TCGGTCAAGG	TATGAGTGGT	GGTATTGCTT	ACGTTATCCC	GTCTGATGTA	8160
30	GAAGCTTTTG	TTGAAAATAA	TCAACTAGAT	ACGCTTTCGT	TTACAAAGAT	TAAACACCAA	8220
	GAAGAAAAAG	CATTCATTAA	GCAAATGCTG	GAAGAACATG	TGTCACACAC	GAATAGTACG	8280
	AGAGCGATTG	ATGTGTTAAA	ACATTTTGAT	CGCATTGAAG	ATGTCGTCGT	TAAAGTTATT	8340
35	CCTAAAGATT	ATCAATTAAT	GATGCAAAAA	ATTCATTGTC	ACAAATCATT	ACATGACAAT	8400
	GAAGATGAAG	CGATGTTAGC	TGCATTTTAC	GATGACAGTA	AAACAATCGA	TGCTAAACAT	8460
	AAACCAGCCG	TTGTGTATTA	AGGAAAGGGG	GAGATACGAT	GGGTGAATTT	AAAGGATTTA	8520
40	TGAAGTATGA	CAAACAGTAC	TTAGGTGAAT	TATCACTGGT	AGACCGTTTG	AAGCATCATA	8580
	AAGCATATCA	ACAACGATTT	ACTAAAGAAG	ATGCCTCTAT	CCAAGGTGCA	CGATGTATGG	8640
	ATTGTGGAAC	GCCGTTTTGT	CAAACCGGAC	AACAGTATGG	TAGGGAAACA	ATAGGTTGTC	8700
45	CAATTGGAAA	CTACATTCCT	GAATGGAACG	ACTTAGTGTA	TCATCAAGAT	TTTAAACTG	8760
	CTTATGAACG	CTTAAGCGAA	ACAAATAACT	TTCCTGACTT	TACAGGGCGT	GTATGTCCTG	8820
50	CACCATGCGA	AAGTGCTTGT	GTGATGAAGA	TTAATAGAGA	ATCGATTGCG	ATTAAAGGTA	8880
	TTGAACGCAC	AATTATTGAT	GAAGCTTTTG	AAAATGGTTG	GGTAGCGCCG	AAAGTTCCGA	8940

55

	CTGAAGAACT TAATCTACTA GGATATCAAG TAACTATTTA TGAACGTGCT AGAGAATCAG	9060
	GCGGTTTATT AATGTATGGT ATTCCGAATA TGAACTTGA TAAAGATGTG GTTCGACGTC	9120
5	GTATTAAGTT AATGGAAGAA GCGGGCATT CTTTCATTAA TGGTGTGAA GTCGGTGTG	9180
	ATATTGATAA AGCAACGTTA GAATCTGAGT ATGATGCCAT TATATTATGT ACTGGTGCAC	9240
	AAAAAGGTAG AGATTTACCT TTAGAAGGAC GCATGGGTGA TGGTATACAT TTCGCTATGG	9300
10	ATTATTTAAC TGAACAAACG CAGTTGTTAA ATGGAGAAAT TGATGATATA ACAATAACTG	9360
	CAAAAGATAA GAATGTCATT ATCATTGGTG CTGGTGATAC AGGGGCAGAC TGTGTAGCGA	9420
15	CAGCATTAAG AGAAAATTGT AAATCGATTG TTCAATTTAA TAAATATACG AAATTGCCAG	9480
	AAGCAATTAC ATTTACAGAA AATGCATCAT GGCCTTTAGC AATGCCGGTG TTTAAAATGG	9540
	ACTATGCGCA CCAAGAGTAC GAAGCTAAGT TTGGTAAGGA ACCACGTGCA TATGGTGTTC	9600
20	AAACAATGCG TTACGATGTT GACGATAAAG GACACATACG TGGTTTGTAT ACTCAAATTT	9660
	TAGAGCAAGG CGAAAATGGT ATGGTCATGA AAGAAGGACC TGAAAGATTT TGGCCTGCTG	9720
	ACCTTGATTT ATTATCAATC GGCTTCGAAG GTACAGAACC AACAGTACCG AATGCTTTTA	9780
25	ACATTAAAAC GGATAGAAAT CGAATCGTGG CGGATGATAC AAATATCAA ACTAATAATG	9840
	AAAAGGTATT TGCTGCTGGA GATGCTAGAC GTGGTCAAAG TTTAGTTGTA TGGGCAATTA	9900
	AAGAAGGTAG AGGCGTAGCG AAAGCAGTAG ATCAGTATTT AGCTAGTAAA GTTTGTGTAT	9960
30	AATCTTTGTA TGGAAATGGT GGTACGTTG ACGTTGTGAC ATGCTGAATC GAGTTTGAAA	10020
	AAATCTAGTA TCTATCAACG TCACATGCCA TCTTTGTAAC CTAAAAACAA AGGTTTGTA	10080
	GACAACAAAT AGATTAATTA TAAGTAGTGA TTTTTTACAT TCGTTTATAG GTCAACTGTA	10140
35	GTGGAAGACA ATGATTTGTG GTAATCATGT AATGCTTAAA AACAAATATTG ACTTTTACAG	10200
	AACGTTTCATA TATGATAAAT ATTGTGTTTT GGAGGAATAC CCAAGTCCGG CTGAAGGGAT	10260
	CGGTCTTGAA AACCGACAGG GGCTTAACGG CTCGCGGGGG TTCGAATCCC TCTTCTCCG	10320
40	CCATCAATAT TTATATTAAA TTCTATATAT AATGAAGGTA AGTGCTCAA TTTTGAGTAT	10380
	TTACCTTTTT TATTTGTCTT TGAATGGCTC GTAATTTTGT ATAATAGAAA TGATAAGGCA	10440
45	TTGAGATTGG AAGGGCATTG GGCTTGTGCA ATATACATAG CTAAATGTCT TTTTGTGTTT	10500
	GTGAAATATG ATGGATGGCT TGTGTGGACA AGTTTGCTAT TTATAGATAT GCATTTTTC	10560
	ATTTAGGAGT TGGCCATGCA TCTACACTTT ATAATGGTGA GAGCGTGGTG AGGTATTGTT	10620
50	AATAACGCAA TTGTAGCGAG GAGTTATTGC TACATATGTC GTTATGGCTC ATTGATTTTC	10680
	TGAAATGGCT ACCCCAGATA ATTGTGACAA AATAAAAATA TTTTGTGTA AGCCTTTACA	10740

55

EP 0 786 519 A2

TAAAAAGAGA AGATGTAAAA GCCATCGTAA CCGCTATTGG GGGAAAAGAA AATCTTGAAG 10860
 CTGCAACGCA TTGTGTAACA CGATTACGTT TAGTGCTGAA GGATGAAAGT AAAGTTGATA 10920
 5 AAGACGCATT AAGTAATAAC GCGTTGGTCA AGGGGCAGTT TAAAGCAGAC CATCAATATC 10980
 AAATTGTCAT TGGTCCAGGA ACAGTCGATG AAGTGTATAA GCAGTTTATT GATGAAACAG 11040
 GTGCTCAAGA AGCTTCGAAA GATGAAGCGA AACAAGCAGC TGCACAAAAA GGAATCCAG 11100
 10 TACAACGTTT GATCAAATTG TGGGGGATA TTTTATACC AATATTACCT GCGATTGTGA 11160
 CAGCTGGTTT GTTAATGGGA ATCAATAATT TACTTACAAT GAAAGGTTTA TTGGTCCAA 11220
 AAGCACTTAT TGAGATGTAT CCACAAATTG CTGATATTC AACATCATT AATGTGATTG 11280
 15 CGAGTACGGC ATTATTTTC TTACCAGCAT TAATTGGTTG GAGTAGTATG CGTGTATTTG 11340
 GTGGTAGTCC GATTCTAGGC ATAGTCTTAG GTTTGATTTT AATGCATCCG CAATTAGTAT 11400
 CTCAGTATGA TTTGGCAAAA GGAATATTC CGACGTGGAA CTTATTTGGC TTAGAGATTA 11460
 20 AGCAGTTGAA TTACCAAGGT CAAGTGTTC CAGTTTAAT TGCAGCTTAC GTTCTAGCTA 11520
 AAATTGAAAA AGGATTAAAT AAAGTCGTTT ACGATTGAT AAAAATGTTG GTCGTTGGAC 11580
 CCGTAGCGCT TTTAGTTACT GGATTTTATG CATTTATTAT CATTGGACCA GTTGCCTTAT 11640
 25 TGATTGGTAC AGGTATTACA TCTGGTGTTA CATTTATATT CCAACATGCA GGATGGCTTG 11700
 GCGGAGCAAT ATATGGATTG TTATATGCAC CACTTGAAT TACAGGACTA CACCATATGT 11760
 30 TTTTAGCAGT AGATTTCCAA TTGATGGGTA GCAGCTTAGG CGGTACGTAT TTATGGCCAA 11820
 TTGTGCGAT TTCCAATATT TGTGAGGCT CTGCAGCATT TGGAGCATGG TTTGTCTATA 11880
 AACGTCGTAA AATGGTTAAA GAAGAAGGCT TGGCATTAA ATCTTGTATT TCTGGTATGT 11940
 35 TAGGTGTTAC TGAACCAGCC ATGTTGCGTG TGAACCTACC TCTGAAATAT CCATTTATCG 12000
 CTGGATATC AACGTCTTGT GTATTGGGGG CAATCGTTGG TATGAATAAC GTACTTGGAA 12060
 AAGTTGGTGT TGGTGGCGTG CCAGCATTC AATCAATTCA AAAAGAATTT TGGCCAGTAT 12120
 40 ATCTTATTGT GACAGCTATT GCTATTGTTG TACCATGTAT ACTAACAATT GTGATGTCTC 12180
 ATTTTAGTAA AAAAAAGCG AAAGAAATTG TTGAAGATTA ATAAAATAAA AAAGGGGCGT 12240
 TCGTTATTTG GACGTCCTTT ATTACGTTAT AAGGTGGTAA TTGTGTGTCG AAAGAAATAG 12300
 45 ATTGGAGAAA ATCCGTTGTA TATCAAATTT ATCCTAAGTC GTTTAATGAT ACGACGGGGA 12360
 ATGGTATAGG AGATATCAAT GGAATTATAG AAAAATTGGA TTATATCAAG TTATTGGGTG 12420
 TTGATTATAT TTGGTTAACA CCAGTGTATG AATCACCGAT GAATGATAAT GGCTATGATA 12480
 50 TCAGCAATTA TTTAGAAATC aATGAAGACT TTGGAACGAT GGATGATTTT GaAAAGTTAA 12540

55

	CGACGGAGCA	TGaaTGGTTT	AAAGAAGCCC	GTAAATCTAA	AGATAACCCy	TATAGAGATT	12660
	ATTACTTTTT	CAGATCATCT	GAAGACGGGC	CGCCAACAAA	TTGGCATTCT	AAATTCGGTG	12720
5	GTAATGCATG	GAAGTATGAT	TCTGAGACAG	ATGAATATTA	TTTACATTTA	TTTGATGTCA	12780
	GTCAAGCTGA	TTTAAATTGG	GATAATCCGG	AAGTACGTCA	ATCGTTATAT	CGCATAGTCA	12840
	ATCATTGGAT	AGACTTCGGC	GTTGATGGTT	TTCGATTTGA	TGTCATTAAAC	TTAATTTCTA	12900
10	AAGGTGAATT	TAAGGACTCT	GACAAAATAG	GTAAAGAATT	TTATACGGAT	GGTCCTAGAG	12960
	TGCATGAGTT	TCTGCATGAA	TTAAATCGTC	AAACGTTTGG	TAACACTGAC	ATGATGACTA	13020
	TAGGAGAAAT	GTCTTCGACG	ACGATTGAAA	ATTGTATTAA	GTATACACAA	CCAGAACGCC	13080
15	AAGAATTGAA	TAGTGTTTTT	AATTTTCATC	ATCTAAAGGT	TGATTATGTT	GATGGTGAAA	13140
	AGTGGACAAA	TGCGAgcTTG	nATTTTCATA	AGTTAAAGGA	AATTCTGATG	CAATGGCAAC	13200
	GAGGTATTIA	TGACGGTGGC	GGATGGAACG	CGATTTTCTG	GTGTAATCAT	GATCAGCCAC	13260
20	GGGTAGTGTC	TAGATTTGGT	GATGATACGT	CGGAAGAGAT	GAGGATACAA	AGTGCTAAAA	13320
	TGTTAGCTAT	CGCACTGCAT	ATGTTGCAAG	GGACGCCATA	TATTTACCAA	GGTGAAGAAA	13380
25	TTGGTATGAC	GGACCCACAT	TTTACATCAA	TAGCACAATA	TCGTGATGTT	GAATCGATTA	13440
	ATGCCTACCA	TCAGTTGTTA	AGTGAAGGGC	ATGCTGAAGC	GGATGTGTTA	GCGATTTTAG	13500
	GACAGAAGTC	ACGAGACAAT	TCGAGAACGC	CTATGCAATG	GAGTGATGAT	GTTAATGCTG	13560
30	GATTTACAGC	TGGTAAaCCT	TGGATTGATA	TTTCGGAAAA	TTATCATCAG	GTCAACGTTA	13620
	GACAAGCACT	TCAGAATAAA	GAGTCTATTT	TCTATACGTA	TCAAAAATTA	ATACAATTAA	13680
	GACATACGCA	TGATATTATT	ACGTATGGAG	ACATTGTGCC	ACGTTTATG	GATCATGATC	13740
35	ATTTATTGT	TTATGAACGT	CATTATAAGA	ATCAACAATG	GCTAGTAATT	GCGAATTTCT	13800
	CAGCAtCGGC	TGTTGATTTG	CCAGAAGGAT	TGGCTAGAGA	AGGTTGTGTT	GTGATTCAAA	13860
	CAGGCACAGT	GGAAAATAAT	ACGATAAGCG	GGTTTGGTGC	AATTGTAATC	GAAACAAACG	13920
40	CGTAAAATAA	ATTGAGTGGA	TGCGTTTATA	TGGCGAAACA	AAAAAAGTTT	ATGAAGATTT	13980
	ATGAGGCGTT	GAAAGAAGAT	ATATTAAACG	GGCAGATTCA	ATATGGTGAA	CAAATTCGGT	14040
	CTGAACATGA	TTTGGTGCAA	TTGTACCAGT	CATCTCGAGA	GACCGTGCGT	AAGGCATTAG	14100
45	ATTTGTTGGC	ATTAGACGGC	ATGATTCAAA	AGATTCAATG	TAAAGGGTCA	CTTGTCATTT	14160
	ATCAGGAGGT	TACAGAGTTT	CCATTTTCTG	AACTTGTTAG	TTTTAAAGAA	ATGCAAGAAG	14220
	AAATGGGCGT	CGCATATTTA	ACTGAAGTTG	TTGTGAATGA	GGTTGTTGAA	GCGCATGAAG	14280
50	TTCCAGAAGT	TCAACATGCT	TTAAACATCA	ATTCTAGTGA	ATCACTCATT	CATATTGTTA	14340

55

TTGTTTCAGA TATAGGTAAT GATGTTGCGA GTGATTCTAT TTATGATTAT TTGGAAAAGG 14460
 TATTAAATCT TAATATTAGT TATTCAAGTA AGTCTATTAC TTTTGAACCG TTTGATGAAC 14520
 5 AAGCATATCA ATTGTTTGGT GATGTATCGG TGGCTTATTC AGCAACAGTT CGAAGTATTG 14580
 TGTATTTAGA AAATACAATG CCGTTTCAAT ATAATATTTT AAAACATCTT GCAAATGAAT 14640
 TTAAATTTAA TGACTTCTCA AGACGTCGTA TAAAGTAAAC AATGATATAA ATGATTTATA 14700
 10 CTTGCAATTA ACTATTAAAA TATAGTAATA TATATCTTGC CGTGCTAGGT GGGGAGGTAG 14760
 CCGTTCCTTG TACTCGAAAT CCGCTTTATG CGAGGCTTAA TTCCTTTGTT GAGGCCGTAT 14820
 TTTTGCGAAG TCTGCCCAA GCACGTAGTG TTTGAAGATT TCGGTCCTAT GCAATATGAA 14880
 15 CCCATGAACC ATGTCAGGTC CTGACGGAAG CAGCATTAAG TGGATCATCA TATGTGCCGT 14940
 AGGgTAGCCG AGATTTAGCT AACGACTTTG GTTACGTTTC TGAATTACGT TCGATGCTTA 15000
 GGTGCACGGT TTTTTATTTT TTAAATATTA AACCgATTAT TAAGAGTTGA AAATATATAA 15060
 20 TTATAGAAGC TACTTTCTTG AAGACAATTC AGCGTATTAT ACGTGAACA TGTGTGTGGG 15120
 AAGTAGCTTT TTTATATGTG AAGTTTGATT CAAGTGAACG CGATGTGCAG TTTGAATGAT 15180
 TTTTGTGTCA ATGAAAAGTA AGAAGTTATA ATTTGATGAT AAAGAAATGA TGGTGAAATG 15240
 25 AGGGGGAGTA TCTTACAATA GAATTATTAA TGAGATACGT TATGATTATT GACAATCAAA 15300
 TGCCTACGGA GGACATATGC AAATATATTT AAGTACTTTA ACAGAGTTAG ATTATGATAA 15360
 ATCTTTAAAT AGTATTGAAG AAAGTTTGA TGATAATCCT GAAACGAGTT GGCAAGCAGG 15420
 30 TGCGAAAGTA AAACATTTAA GAAATCTCC TTGCTATAAT TTTGAATTAG AAGTAATAGC 15480
 GAAAAATGAA AATAACGATG TCGTTGGACA CGTTTTATTA ATTGAAGTAG AAATTAATAG 15540
 35 TGATGATAAG ACGTATTATG GTTTGGCGAT TGCTCTTTA TCAGTTCATC CTGAATTACG 15600
 TGGACAAAAA TTAGGTCGTG GCTTGGTTCA AGCAGTAGAA GAGCGTGCCA AAGCACAAGA 15660
 GTATAGTACG GTTGTGTAG ACCATTGTTT TGACTACTTT GAAAAGTTGG GTTATCAAAA 15720
 40 TGCTGCTGAG CATGACATTA AATTAGAATC TGGTGATGCA CCGTTACTTG TAAAATATTT 15780
 ATGGGATAAT TTGACGGATG CACCACACGG AATCGTAAAA TTTCCAGAAC ATTTTTATTA 15840
 ATTGTTCAAT TAAGAAGTAA AGGTATTATC ATGCTATAAT GAGAGGTAAT TGTTTATGGA 15900
 45 GGTGCTAACT TGAATTATCA AGCCTTATAT CGTATGTACA GACCCCAAAG TTTCGAGGAT 15960
 GTCGTCGGAC AAGAACATGT CACGAAGACA TTGCGCAATG CGATTTGCGA AGAAAAACAG 16020
 TCGCATGCTT ATATTTTTAG TGGTCCGAGA GGTACGGGGA AAACGAGTAT TGCCAAAGTG 16080
 50 TTTGcTAAAG CAATCAACTG TCTAAATAGC ACTGATGGAG AACCTTGTA TGAATGTCAT 16140

AATAATGGCG TTGATGAAAT AAGAAATATT AGAGACAAAG TTAAATATGC ACCAAGTGAA 16260
 TCGAAATATA AAGTTTATAT TATAGATGAG GTGCACATGC TAACAACAGG TGCTTTTAAAT 16320
 5 GCCCTTTTAA AGACGTTAGA AGAACCTCCA GCACACGCTA TTTTATATATT GGCAACGACA 16380
 GAACCACATA AAATCCCTCC AACAAATCATT TCTAGGGCAC AACGTTTTGA TTTTAAAGCA 16440
 ATTAGCCTAG ATCAAATTGT TGAACGTTTA AAATTTGTAG CAGATGCACA ACAAATTGAA 16500
 10 TGTGAAGATG AAGCCTTGGC ATTTAtcgCT AAAGCGTCTG AAGGGGGTAT GCGTGATGCA 16560
 TTAAGTATTA TGGATCAGGC TATTGCATTT GGTGATGGTA CGTTAACATT GCAAGATGCG 16620
 TTGAATGTCA CAGGTAGCGT ACATGATGAA GCGTTGGATC ACTTGTTTGA TGATATTGTA 16680
 15 CAAGGTGACG TACAAGCATC TTTTAAAAA TACCATCAGT TTATAACAGA AGGTAAAGAA 16740
 GTGAATCGCC TAATAAATGa TATGATTTAT TTTGTCaGAG ATACGATTAT GAATAAAACA 16800
 TCTGAGAAAG ATACTGAGTA TCGAGCACTG ATGAACTTAG AATTAGATAT GTTATATCAA 16860
 20 ATGATTGATC TTATTAATGA TACATTAGTG TCGATTCTGT TTAGTGTGAA TCAAAACGTT 16920
 CATTTTGAAG TGTGTTAGT AAAATTAGCT GAGCAGATTA AGGGTCAACC ACAAGTGATT 16980
 25 GCGAATGTAG CTGAACCAGC ACAAATTGCT TCATCGCCAA ACACAGATGT ATTGTTGCAA 17040
 CGTATGGAAC AGTTAGAGCA AGAACTAAAA ACACTAAAAG CACAAGGAGT GAGTGTCGCT 17100
 CCTGTTCAAA AATCTTCGAA AAAGCCTGCG AGAGGCATAC AAAAATCTAA AAATGCATTT 17160
 30 TCAATGCAAC AAATTGCAAA AGTGCTAGAT AAAGCGAATA AGGCAGATAT CAAATTGTTG 17220
 AAAGATCATT GGCAAGAAGT GATTGATCAT GCCAAAAATA ATGATAAAAA ATCACTCGTT 17280
 AGTTTATTGC AAAATTCGGA ACCTGTGGCG GCAAGTGAAG ATCACGTACT TGTGAAATTT 17340
 35 GAGGAAGAGA TCCATTGTGA AATCGTCAAT AAAGACGACG AGAAACGTAG TAGTATAGAA 17400
 AGTGTGTGAT GTAATATCGT TAATAAAAAAC GTTAAAGTTG TTGGTGTACC ATCAGATCAA 17460
 TGGCAAAGAG TTCGAACGGA ATATTTACAA AATCGTAAAA ACGAAGGCGA TGATATGCCA 17520
 40 AAGCAACAAG CACAACAAAC AGATATTGCT CAAAAAGCAA AAGATCTTTT CGGTGAAGAA 17580
 ACTGTACATG TGATAGATGA AGAGTGATAC ATGACAAGCG ATATAATCGT ATGTATAATG 17640
 AAAGAAACAT CATTTTATTG ATAAATATTT ATTGATTTTC AAGGAGGAAA TGAATATGC 17700
 45 GCGGTGGCGG AAACATGCAA CAAATGATGA AACAAATGCA AAAAATGCAA AAGAAAATGG 17760
 CTCAAGAACA AGAAAACTT AAAGAAGAGC GTATTGTAGG AACAGCTGGC GGTGGCATGG 17820
 TTGCAGTTAC TGTAACCTGGT CATAAAGAAG TTGTCGACGT TGAAATCAAA GAAGAAGCTG 17880
 50 TAGACCCAGA CGATATTGAA ATGCTACAAG ACTTAGTGTT AGCAGCTACT AATGAAGCGA 17940

55

EP 0 786 519 A2

	TCCCTGGaAT GTGATCATAG ATGCATTATC CAGAACCTAT ATCAAACTT ATTGATAGCT	18060
	TTATGAAATT GCCAGGCATT GGTCCAAAGA CAGCCCAACG TCTGGCTTTT CATACTTAG	18120
5	ATATGAAAGA AGACGATGTT GTTCAGTTTG CCAAAGCATT AGTAGATGTT AAGAGAGAAT	18180
	TAACATATTG TAGCGTATGT GGTACACATTA CTGAAAATGA TCCATGTTAT ATTTGTGAAG	18240
	ATAAGCAAAG AGATCGTTCA GTTATTTGTG TTGTGGAAGA TGACAAAGAT GTCATAGCTA	18300
10	TGGAAAAAAT GAGAGAATAC AAAGGTTTAT ATCACGTTTT ACATGGGTCT ATTTGCGCTA	18360
	TGGATGGCAT TGGACCAGAA GATATTAATA TTCCTTCATT GATTGAACGC TTGAAAAACG	18420
	ATGAAGTTAG CGAATTAATC TTAGCTATGA ACCCGAACTT AGAGGGGGAA TCTACAGCCA	18480
15	TGTATATTTT TAGATTAGTT AAGCCTATAG GTATCAAAGT GACGAGATTA GCACAAGGGT	18540
	TATCGGTAGG TGGCGATTTA GAGTATGCTG ACGAAGTAAC ATTATCTAAA GCAATCGCAG	18600
	GTAGAACAGA AATGTAATKT CTTCTATTAA ACATTTTGA TTTTAATACT ATAGTAAGAA	18660
20	AAGTCACAGT GTAATCATTG TGGCTTTTTT TATGGTGTGG TGTGATGTAC TACTTTATTT	18720
	GCGGTGTGGC GGTGGTATGG TTTACCTAGT TTTACTGAGG GATGGGTAAT CTTTAGGAAG	18780
	CAAGCCGTTG GTTGTGATTT GTTACTTCTA ATAGTAATGA TGTGAATTGG ATTATCGAAT	18840
25	TAGATCTATG GTTATGGTGT GTTGGTGCTA TTAATTTGAT AAATGCGGTT AATGACTATG	18900
	CAAATGAAAT TCTTTTGTA TTTGAAATGAT AGATGCTGGC TTAGTAAGTT GTACTTCTTT	18960
30	GGTCTAAAGC TTATTAAATC AGCCTGTATA GCGGTGTTTT GAGAGATTAT TTAAACTTG	19020
	TAAATTTATT TTTAATTTCT GGTAAAAAAA TAACGTTCTG TTTTGCCTTT TTTTGTATTG	19080
	ATATGGTTAG AGAAAAATCT GTTCTTGTG CTAAAAACG TACTATTTAT AAGTGGGGAT	19140
35	TTTTTAAGTT CGATTTTTAG GATAAGGGCG TTCAGTACAG ATGACAAAGG TGTAATTTTT	19200
	ACTGTGTGTA AGCAGTTTGA AAGCCTGTAT AGTATTTATT TGTGAGGCA AACAAAACAA	19260
	CTCAACTTAA GAAATAACTT GAATTACTAA CGAAAAATTA TTTTAAAAAG TTATTGACTT	19320
40	AAATGTTAAT AAAATGTATA ATTAATTCTT GTCGGTAAGA AAAATGAACA TTGAAACTG	19380
	AATGACAATA TGTCAACGTT AATTCCAAAA AACGTAACTA TAAGTTACAA ACATTATTTA	19440
	GTATTTATGA GCTAATCAAA CATCATAATT TTTATGGAGA GTTTGATCCT GGCTCAGGAT	19500
45	GAACGCTGGC GCGGTGCCTA ATACATGCAA GTCGAGCGAA CGGACGAGAA GCTTGCTTCT	19560
	CTGATGTTAG CGGCGGACGG GTGAGTAACA CGTGGATAAC CTACCTATAA GACTGGGATA	19620
	ACTTCGGGAA ACCGKAGCTA ATACCGGATA ATATTTTGAA CCGCATGGTT CAAAAGTGAA	19680
50	AGACGGTCTT GCTGTCACCT ATAGATGGAT CCGCGCTGCA TTAGCTAGTT GGTAAGGTAA	19740

55

	GAGACACGGT	CCAGACTCCT	ACGGGAGGCA	GCAGTAGGGA	ATCTTCCGCA	ATGGGCGAAA	19860
	gCtGaCGGAG	CAACGCCGCG	TGAGTGATGA	AGGTCTTCGG	ATCGTAAAAC	TCTGTTATTA	19920
5	GGGAAGAACA	TATGTGTAAG	TAAGTGTGCA	CATCTTGACG	GTACCTAATC	AGAAAGCCAC	19980
	GGCTAACTAC	GTGCCAGCAG	CCGCGGTAAT	ACGTAGGTGG	CAAGCGTTAT	CCGGAATTAT	20040
	TGGGCGTAAA	GCGCGCGTAG	GCGGTTTTTT	AAGTCTGATG	TGAAAGCCCA	CGGCTCAACC	20100
10	GTGGAGGGTC	ATTGGAAACT	GGAAAACCTG	AGTGCAGAAG	AGGAAAGTGG	AATTCCATGT	20160
	GTAGCGGTGA	AATGCGCAGA	GATATGGAGG	AACACCACTG	GCGAAGGCGA	CTTCTGGTC	20220
	TGTAAGTGAC	GCTGATGTGC	GAAAGCGTGG	GGATCAAACA	GGATTAGATA	CCCTGGTAGT	20280
15	CCACGCCGTA	AACGATGAGT	GCTAAGTGTT	AGGGGGTTTC	CGCCCTTAG	TGCTGCAGCT	20340
	AACGCATTAA	GCACTCCGCC	TGGGGAGTAC	GACCGCAAGt	TGAAACTCAA	AGGAATTGAC	20400
	GGGGACCCGC	ACAAGCGGTG	GAGCATGTGG	TTTAATTCTGA	AGCAACGCGA	AGAACCTTAC	20460
20	CAAATCTTGA	CATCCTTTGA	CAACTCTAGA	GATAGAGCCT	TCCCCTTCGG	GGGACAAAGT	20520
	GACAGGTGGT	GCATGGTTGT	CGTCAGCTCG	TGTCGTGAGA	TGTTGGGTTA	AGTCCCGCAA	20580
25	CGAGCGCAAC	CCTTAAGCTT	AGTTGCCATC	ATTAAGTTGG	GCACTCTAAG	TTGACTGCCG	20640
	GTGACAAACC	GGAGGAAGGT	GGGGATGACG	TCAAATCATC	ATGCCCTTA	TGATTTGGGC	20700
	TACACACGTG	CTACAATGGA	CAATACAAAG	GGCAGCGAAA	CCGCGAGGTC	AAGCAAATCC	20760
30	CATAAAGTTG	TTCTCAGTTC	GGATTGTAGT	CTGCAACTCG	ACTACATGAA	GCTGGAATCG	20820
	CTAGTAATCG	TAGATCAGCA	TGCTACGGTG	AATACGTTCC	CGGGTCTTGT	ACACACCGCC	20880
	CGTCACACCA	CGAGAGTTTG	TAACACCCGA	AGCCGGTGGA	GTAACCTTTT	AGGAGCTAGC	20940
35	CGTCGAAGGT	GGGACAAATG	ATTGGGGTGA	AGTCGTAAAC	AGGTAGCCGT	ATCGGAAGGT	21000
	GCGGCTGGAT	CACCTCCTTT	CTAAGGATAT	ATTCGGAACA	TCTTCTTCAG	AAGATGCGGA	21060
	ATAACGTGAC	ATATTGTATT	CAGTTTTGAA	TGTTTATTTA	ACATTCAAAT	ATTTTTTGGT	21120
40	TAAAGTGATA	TTGCTTATGA	AAATAAGCA	GTATGCGAGC	GCTTGACTAA	AAAGAAATTG	21180
	TACATTGAAA	ACTAGATAAG	TAAGTAAAT	ATAGATTTTA	CCAAGCAAAA	CCGAGTGAAT	21240
	AAAGAGTTTT	AAATAAGCTT	GAATTCATAA	GAAATAATCG	CTAGTGTTTC	AAAGAACACT	21300
45	CACAAGATTA	ATAACGCGTT	TAAATCTTTT	TATAAAAGAA	CGTAACTTCA	TGTTAACGTT	21360
	TGACTTATAA	AAATGGTGGA	AACATAGATT	AAGTTATTAA	GGGCGCACGG	TGGATGCCTT	21420
	GGCACTAGAA	GCCGATGAAG	GACGTTACTA	ACGACGATAT	GCTTTGGGGA	GCTGTAAAGTA	21480
50	AGCTTTGATC	CAGAGATTTT	CGAATGGGGA	AACCCAGCAT	GAGTTATGTC	ATGTTATCGA	21540

55

EP 0 786 519 A2

	GAGGAAGAGA AAGAAAATTC GATTCCCTTA GTAGCGGCGA GCGAAACGGG AAGAGCCCAA	21660
	ACCAACAAGC TTGCTTGTTG GGGTTGTAGG ACACTCTATA CGGAGTTACA AAGGACGACA	21720
5	TTAGACGAAT CATCTGGAAG GATGAATCAA AGAAGGTAAT AATCCTGTAG TCGAAAATGT	21780
	TGTCTCTCTT GAGTGGATCC TGAGTACGAC GGAGCACGTG AAATCCGTC GGAATCTGGG	21840
	AGGACCATCT CCTAAGGCTA AATACTCTCT AGTGACCGAT AGTGAACCAG TACCGTGAGG	21900
10	GAAAGGTGAA AAGCACCCCG GAAGGGGAGT GAAATAGAAC CTGAAACCGT GTGCTTACAA	21960
	GTAGTCAGAG CCCGTTAATG GGTGATGGCG TGCCTTTTGT AGAATGAACC GGCGAGTTAC	22020
	GATTTGATGC AAGGTTAAGC AGTAAATGTG GAGCCGTAGC GAAAGCGAGT CTGAATAGGG	22080
15	CGTTTAGTAT TTGGTCGTAG ACCCGAAACC AGGTGATCTA CCCTTGGTCA GGTGAAGTT	22140
	CAGGTAACAC TGAATGGAGG ACCGAACCGA CTTACGTTGA AAAGTGAGCG GATGAACTGA	22200
	GGGTAGCGGA GAAATTCCAA TCGAACCTGG AGATAGCTGG TTCTCTCCGA AATAGCTTTA	22260
20	GGGCTAGCCT CAAGTGATGA TTATTGGAGG TAGAGCACTG TTTGGACGAG GGGCCCCCTCT	22320
	CGGGTTACCG AATTCAGACA AACTCCGAAT GCCAATTAAT TTAAGTTGGG AGTCAGAACA	22380
	TGGGTGATAA GGTCCGTGTT CGAAAGGGAA ACAGCCCAGA CCACCAGCTA AGGTCCCAAA	22440
25	ATATATGTTA AGTGGAAGAG GATGTGGCGT TGCCCAGACA ACTAGGATGT TGGCTTAGAA	22500
	GCAGCCATCA TTAAAGAGT GCGTAATAGC TCACTAGTCG AGTGACACTG CGCCGAAAT	22560
30	GTACCGGGGC TAAACATATT ACCGAAGCTG TGGATTGTCC TTTGGaCAAT GGTAGGAGAG	22620
	CGTTCTAAGG GCGTTGAAGC ATGATCGTAA GGACATGTGG AGCGCTTAGA AGTGAGAATG	22680
	CCGGTGTGAG TAGCGAAAGA CGGGTGAGAA TCCCGTCCAC CGATTGACTA AGGTTTCCAG	22740
35	AGGAAGGCTC GTCCGCTCTG GGTAGTCGG GTCCTAAGCT GAGGCCGACA GcGTAGGCGA	22800
	TGGATFAACAG GTTGATATTC CTGTACCACC TATAATCGTT TTAATCGATG GGGGGACGCA	22860
	tAGGATAGGC GAAGcGTGcG ATTGGATTGC ACGTCTAAGC AGTAAGGCTG AGTATTAGGC	22920
40	AAATCCGGTA CTCGTTAAGG CTGAGCTGTG ATGGGGAGAA GACATTGTGT CTTGAGTGC	22980
	TTGATTTTAC ACTGCCGAGA AAAGCCTCTA GATAGAAAAT AGGTGCCCCGT ACCGCAAACC	23040
	GACACAGGTA GTCAAGATGA GAATTCTAAG GTGAGCGAGC GAACTCTCGT TAAGGAACTC	23100
45	GGCAAAATGA CCCCCTAACT TCGGGAGAAG GGGTGCTCTT TAGGGTTAAC GCCCAGAAGA	23160
	GCCGCAGTGA ATAGGCCCAA GCGACTGTTT ATCAAAAACA CAGGTCTCTG CTAAACCGTA	23220
	AGGTGATGTA TagGGcTGAC GCCTGCCCCG TGCTGGAAGG TTAAGAGGAG TGGTTAGcTT	23280
50	CTGCGAAGcT ACGAATCGAA GCCCCAGTAA ACGGCGGCCG TAACTATAAC GGTCCTAAGG	23340

55

	TGTCTCAACG AGAGACTCGG TGAAATCATA GTACCTGTGA AGATGCAGGT TACCCGCGAC	23460
	AGGACGGAAA GACCCCGTGG AGCTTTACTG TAGCCTGATA TTGAAATTCG GCACAGCTTG	23520
5	TACAGGATAG GTAGGAGCCT TTGAAACGTG AGCGCTAGCT TACGTGGAGG CGCTGGTGGG	23580
	ATACTACCCT AGCTGTGTTG GCTTTCTAAC CCGCACCCT TATCGTGGTG GGAGACAGTG	23640
	TCAGGCGGGC AGTTTGA CTG GGGCGGTCGC CTCCTAAAAG GTAACGGAGG CGCTCAAAGG	23700
10	TTCCCTCAGA ATGGTTGGAA ATCATT CATA GAGTGTAAG GCATAAGGGA GCTTGACTGC	23760
	GAGACCTACA AGTCGAGCAG GGTGAAAGA CGGACTTAGT GATCCGGTGG TTCCGCATGG	23820
	AAGGGCCATC GCTCAACGGA TAAAAGCTAC CCCGGGGATA ACAGGCTTAT CTCCCCAAG	23880
15	AGTTCACATC GACGGGGAGG TTTGGCACCT CGATGTCGGC TCATCGCATC CTGGGGCTGT	23940
	AGTCGGTCCC AAGGGTTGGG CTGTTGCCCC ATTAAAGCGG TACGCGAGCT GGGTTCAGAA	24000
	CGTCGTGAGA CAGTTCGGTC CCTATCCGTC GTGGGCGTAG GAAATTTGAG AGGAGCTGTC	24060
20	CTTAGTACGA GAGGACCGGG ATGGACATAC CTCTGGTGTA CCAGTTGTCTG TGCCAACGGC	24120
	ATAGCTGGGT AGCTATGTGT GGACGGGATA AGTGCTGAAA GCATCTAAGC ATGAAGCCCC	24180
25	CCTCAAGATG AGATTTCCTA ACTTCGGTTA TAAGATCCCT CAAAGATGAT GAGGTTAATA	24240
	GGTTCGAGGT GGAAGCATGG TGACATGTGG AGCTGACGAA TACTAATCGA TCGAAGACTT	24300
	AATCAAAATA AATGTTTTGC GAAGCAAAAT CACTTTTACT TACTATCTAG TTTTGAATGT	24360
30	ATAAATTACA TTCATATGTC TGGTGA CTAT AGCAAGGAGG TCACACCTGT TCCCATGCCG	24420
	AACACAGAAG TTAAGCTCCT TAGCGTCGAT GGTAGTCGAA CTTACGTTCC GCTAGAGTAG	24480
	AACGTTGCCA GGCAAAAAAT GGATGCGATG AGCCGCATTG AGACCGCAAG GTCTCTTTTT	24540
35	TTTATGTCTA AAACGTCAAA ATAAAAAGCA AACACAAAGA AAAATGGCTT GGCGAAGTGA	24600
	AAACGTTTGA ATCTGACGAA ACGAGAAAAG ATCGCAACGA GTTTAGTAGA GCTAAATGAG	24660
	TAAGyGAGAG CCGAAGrAGA GGAAAGAAGC AAGCGATTGT CACAAGTCAA GAAAGGTTCT	24720
40	TAGCGAsGAT GGTAGCCAAC TTACGTTCCG CTAGAGTAGA ACTGGAAATG ATAATTTAAT	24780
	AATGTACACT TTCGATTGTC TAAGTATGTA CAACTTTAAT TTTGTGTTTA TATAAATTTA	24840
	AAATGATATC ATCGAAAACA AAATATTGTA TAAATAGAGA AGAGCAGTAA GACGGTATCT	24900
45	AATTGAAAAT GATCTTACTG CTCTTTTATA TACTTTATTG AAATACAAAA AGGAAATTAA	24960
	TTATTATACA ATAGACAAGC TATTGCATAA GTAACACTAA CTTTATCAA AGAAGTGTTA	25020
	CTTTATAATT AATGATTTTA TTAGAGCGTC TACATGCGGT TTTAAAGCAT CATCGTCTAT	25080
50	ACCGCCAAAG CCTAATATAA ATTTAGGGGT TTTCTTATAG TCTTGATCAT CATCAAAATT	25140

55

	TCCATTTTTT ACTGTAATTG TAAAATGCAT ACCCGTTTCA GCACCTTGAA TATCAAGCTG	25260
	CTCTTTGTAA GGTTC AATC TTTTAAAT ATAGGTTAGT TTTCTACGAT AAATTCGTCT	25320
5	CATTTTATTT AAATGCCTTT CAAAACCACC GGAAGATATA AACGTTGCAA TAAGGTTTTG	25380
	CATATGAACA GGTACAGTGT TGCCTTCAAT GTGATTTTGA GAATGATATT TTTTCATTAT	25440
	AGAATAGGGT AACACCATAT ATGCAACTCG ACAGCTAGGA AAAATAGACT TTGAAAATGT	25500
10	ACTGATATAA ATCACTTTTT CTCCTCTTGA ATATAGACCT TGAATTGCTG GAATGGGTTT	25560
	GCCGAAATAT CTAAACTCGG AATCATAATC ATCTTCTATA ATAAATCGTT CTCTTTTTTC	25620
	TTGAGCCCAT TGTATTAATT GAGTTCGTTT TTTAAGTCC ATCACATATC CAGTTGGAAA	25680
15	TTGATGGGAA GGC GTTATAT ATACTATATT TTTTGTGAT TTAATAACTT CATCTACGTT	25740
	TATTCCATTA TCTTCAACTT CAATTTGTTT ATATTCAACT TGTTTTTTAT CTAAATATT	25800
	TTTGATTGGT GGATAACTAG GTTTTTCGAT AATAAATGTT GAAGTATAAA GTAAATCGAC	25860
20	TAATTGATTT ACTAATTGTT CGGTAGATGA GCCAATTATA ATTTGATTAG GATCACAAAT	25920
	TACGCCACGA TTAGTAAATA AATAAAATGC CAGTTGAAAC CGCAAATGTA ATTCTCCTTG	25980
	AAAATGTCCT CTACGTAATT GATTTAAATG ATTTGTATCA TAAAGATCTT TGGAAACTT	26040
25	TCTGAAAAGT TCTATAGGGA AATGTTTCGT ATCTATTTC TCCAAATTAA AAGCATAATC	26100
	ATAAGCTTCA TCACTCGCTT TTGGTTTATA TGAATCATCA TCAAAAAGAG AGGGGATAGG	26160
30	TTGATTGTTT AAAATTGTTA AAGATTCAAT TTCGGACACA AAATATCCAG AGCGAGGTCT	26220
	TGAATAAATG TAACCTTCGT CTAATAGAAG TTGATATGCA TGCTCTACGG TTGTTTGGCT	26280
	AATAGATAAA TGTTTGCTTA ATTGTCTTTT AGAATAAAAT TTATCGCCTT CTTTAAATTG	26340
35	ACCTTCAATT ATTTGTTTTT TTAATTTTTC ATAAAGTTGA TGGTATAAAG TGTTTTTCAA	26400
	TTTTATAACT GACCTCCTAA ATTTATCTTA TTTTGTACCT TTTTAAATAT CAGTTTATAC	26460
	ATTACAATGT ATTTAATCAA CTTGAAAAGG GGTTTTATGT ATAATGAGTA AAATTATTGG	26520
40	ATCAGACAGA GTCAAAAGAG GTATGGCTGA AATGCAAAAA GGCGGCGTTA TTATGGATGT	26580
	CGTTAATGCT GAGCAAGCAA GAATTGCAGA AGAAGCTGGC GCGGTAGCAG TTATGGCATT	26640
	AGAACGAGTA CCTTCTGATA TTAGAGCTGC TGGTGGTGTT GCACGTATGG CAAACCCTAA	26700
45	AATTGTAGAA GAAGTAATGA ATGCTGTTTC TATTCCAGTC ATGGCTAAAG CACGTATTGG	26760
	TCATATCACT GAAGCAAGAG TATTAGAGGC GATGGGTGTT GACTATATTG ATGAATCAGA	26820
	AGTGTTAACA CCAGCAGATG AGGAATATCA CTTAAGAAAA GATCAATTTA CAGTACCATT	26880
50	TGTATGTGGA TGTCGTAATT TAGGTGAAGm TGCGCGTAGA ATTGGTGAAG GTGCTGCTAT	26940

55

	ACAAGTTAAT TCAGAAGTTA GTCGATTGAC TGTAATGAAT GATGATGAGA TTATGACTTT	27060
	TGCGAAAGAT ATCGGTGCGC CTTATGAAAT TTTAAAACAA ATTAAAGACA ATGGTCGTTT	27120
5	ACCGGTAGTT AACTTTGCAG CTGGTGGCGT TGCGACTCCT CAAGATGCTG CTTTAATGAT	27180
	GGAATTAGGT GCTGACGGTG TATTCGTTGG ATCAGGTATT TTTAAATCAG AAGATCCAGA	27240
	AAAAATTGCT AAAGCAATTG TTCAAGCAAC AACACATTAC CAAGACTATG AACTAATTGG	27300
10	AAGATTAGCA AGTGAAGTTG GCACTGCTAT GAAAGGTTTA GATATCAATC AATTATCATT	27360
	AGAAGAACGT ATGCAAGAGC GTGGTTGGTA AGATATGAAA ATAGGTGTAT TAGCATTACA	27420
15	AGGTGCAGTA CGTGAACATA TTAGACATAT TGAATTAAGT GGTCAATGAAG GTATTGCAGT	27480
	TAAAAAAGTT GAACAATTAG AAGAAATCGA GGGCTTAATA TTACCTGGTG GCGAGTCTAC	27540
	AACGTTACGT CGATTAATGA ATTTATATGG ATTTAAAGAG GCTTTACAAA ATTCAACTTT	27600
20	ACCTATGTTT GGTACATGCG CAGGATTAAT AGTTCTAGCG CAAGATATAG TTGGTGAAGA	27660
	AGGATACCTT AACAAGTTGA ATATTACTGT ACAACGAAAC TCATTCGGTA GACAAGTTGA	27720
	CAGCTTTGAA ACAGAATTAG ATATTAAAGG TATCGCTACA GATATTGAAG GTGTCTTTAT	27780
25	AAGAGCCCCA CATATTGAAA AAGTAGGTCA AGGCGTAGAT ATCCTATGTA AGGTTAATGA	27840
	GAAAATTGTA GCTGTTGAGC AAGGTAAATA TTTAGGCGTA TCATTCCATC CTGAATTAAC	27900
	AGATGACTAT AGAGTAACTG ATTACTTTAT TAATCATATT GTAAAAAAG CATAGCTTAA	27960
30	TGTATGCTAA ATCAACGAAT TATTGATATT TATAGATTTG TTGAGAAGAA AATATCTCCT	28020
	TCAAACCTAG CTTTGGAGGA GTTATTTTTT ATGTCAAAAT TAAAAATGAT AAAAAATAAA	28080
35	GCTATACATA AGAAAAAAC CCTTCAAAGA GACTGAGAAT AGTCAAATT TTGAAGGGGT	28140
	TAATTCGATG TTGATGTATT TGTTAAATAA AGAATCCAGC GATTGCAGCT GAAATGAAAG	28200
	ATACTAGTGT TGCACCGAAT AATAATTTCA AACCAAAGCG GGCAACTGTA TCTCCTTTTT	28260
40	TGTCATTAAG TGATTTAATC GCACCTGAAA TAATACCGAT AGAGCTAAAG TTAGCAAATG	28320
	ATACTAAGAA TACAGATGTA ACACCTTTTG CGTGTTGAGA TAAATCACTA AGTTTACCAA	28380
	GTGCTTGCAT TGCTACAAAT TCGTTAGATA ATAGTTTTGT CGCCATAACT GAACCGGCTT	28440
45	GAACTGCATC TTGCCATGGC ACACCGACTA AGAATGCAAA TGGTGCAAAG ACAAACCAA	28500
	TTAATGTTTG GAAATCCCAA GAAATAGCGC CACCTGAAAC TGTACTAAAG ATATTGCTTA	28560
	CAATTCCATT TAATAGAGCG ATAATGGCAA TGTATCCGAT TAACATTGCG CCTACAATGA	28620
50	CAGCTACTTT AAATCCATCT AAAATATATT CTCCTAGCAT TTCGAAGAAT GATTGTTGTC	28680
	TTTCTTCAGT TTCTTCAACT AATAATTTGT CATCTTCTTC ATTAACCTTA TAAGGGTTAA	28740

55

	TAGGTTCAAT TAAGGTAAAG TATGCACCGA TAATTGAAGC AGAAACAGTC GACATTGCTG	28860
	AAGCTGTTAA TGTGTATAAA CGTTGCTTAG GTATGTATGG TAATTGTTTT TTAATTGAAA	28920
5	TAAATACTTC AGATTGTCCC AAAATTGCTG CAGCAACTGC ATTGTATGAT TCTAAACGTC	28980
	CCATACCATT AATTTTAGAA ATTAAGAATC CTAAAACATT AATGATTAAA GGTAAAATCT	29040
10	TTGTGTATTG AAGGATACCG ATAATCGCTG AAATAAATAC GATAGGTAAT AATACACTGA	29100
	AGAAGAATGG TGGTTGCTTA GGATCGATAT ATTGAATACC ACCGAATACA AAGTTAACAC	29160
	CATCTGCTGC TTTTAATAAT AAGTAGTTAA AACCGTTTGA AATACCACCA ATAACCTTGA	29220
15	TTCCCATTGT AGTTTTAAGC AAGATAAATG CAAAGATAAG CTGAATTGCA AGTAAAATTC	29280
	CTACATATTT CCAGCGAATA TTTTCTCTGT CTGAGCTAAA TAGAAACGCA AGTGCTAAAA	29340
	AGAAGATAAT TCCGATAATC CCAATTAGAA TATGCATATA TTTCTCATTC CTTTAGTTTT	29400
20	TTCTACaATc TATCATACAA TAAAATGGAA GGGCTAACAT CATAAATTTT TGAAAATATA	29460
	AAAACAAATT AATTGAAAAA GGTCAAAATA GGTCATATAA TATAGTCAAA GAAGGTCAAA	29520
	AAGGGGTGAT ATACATGCAC AATATGTCTG ACATCATAGA ACAATAaTCA AACGTTTATT	29580
25	TGAAGAGTCG AATGAAGATG TCGTTGAAAT TCAGAGAGCG AATATCGCAC AGCGTTTTGA	29640
	TTGCGTACCA TCACAATTAA ATTATGTAAT CAAAACACGA TTCACTAATG AACATGGTTA	29700
	TGAAATCGAA AGTAAACGTG GTGGTGGTGG TTACATCCGA ATCACTAAAA TTGAAAATAA	29760
30	AGATGCAACA GGTATATTA ATCATTTGCT TCAGCTGATT GGACCTTCTA TTTCTCAACA	29820
	ACAAGCTTAT TATATTATTG ATGGGCTTTT AGATAAAATG TTAATAAATG AACGTGAAGC	29880
35	TAAAATGATT CAAGCAGTTA TTGATAGAGA AACGCTATCA ATGGATATGG TTTCTAGAGA	29940
	TATTATTAGA GCAAATATTT TAAAACGTTT GTTACCAGTT ATAAATTATT ACTAAATGAA	30000
	ATGAGGTGTT GAAGTGCTTT GTGAAAATTG TCAACTTAAT GAAGCGGAAT TAAAAGTTAA	30060
40	AGTTACAAGT AAAAATAAAA CAGAAGAAAA AATGGTGTGT CAAACTTGTG CTGAGGGGCA	30120
	CCATCCGTGG AATCAAGCTA ATGAACAACC TGAaTATCAA GAACATCAAG ATAATTTCGA	30180
	AGAAGCATTT GTTGTTAAGC AAATTTTACA ACATTTAGCT ACGAAACATG GAATTAATTT	30240
45	TCAAGA	30246

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14333 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

	TATTCCCCCA TCGGTTTATT AAATCGTCCA TTTCATACT GTTTTCCCCC AAGATGTGGA	60
5	TAAATCCATT TCAAACGCTT GGACGATATC TTGCATCGTA CATACATTAA TTTCATGTCC	120
	TTTTAATAAT GCTAACTTTT CAACTATGTC TGGGTACTTA CGATATAAAT CAACAACTTG	180
10	CTCAAAATCT TTAGAGCCGC TTCGACTACT ACCAATCAAC GTTAATCCTT TTTCAAGTAC	240
	TAATCGTGTA TTCACTTCCA CGGGTAATTC ACTTACGCCT AACAAAGCAA TACTGCCTTC	300
	TGGTGAAATA TGTTCAACTA TTTGTTGAAG TGCAACTTGA CTTCCTTTAC CTCCAACACA	360
15	TTCAAATGCA TGATCAATTT TAAGATCATC TGGTATTTGA TTTACTGTAA AGATGTCATC	420
	TACAAATGAA AAATGACTTA ATTTATAGTC TGTCTTACCA AATACATAAG TTTTAGCTTC	480
	TGGGTACAAC TTACGTAGCA AAATAGCAGT AATATAACCT AAGTTACCAT CACCCCAAAT	540
20	ACCAAAGCTG GTTTTCAAAG GTATAGATTT ACGTTCAAAT CGTTGTATAG CATGATAACT	600
	TACTGACACT AACTCTGTGT ATGAAATCGT ACTCAAATCA ATGTCATTAG GCAGCGGAAC	660
	GATACGATCA TGTGCCATCA CAACGTAGTC TTGCATAAAA CCATCATAAC CACTAGATCT	720
25	AAAATAACTA GAGGCTAAGT AATTCTCCGC AATAATATGA TGTGCTCTG TAGGTGTATT	780
	CGGTACCATT ACTACTTTTCG TACCTTTTTTC AAATACCCCT TTACTATCAA ATACAACTTC	840
	ACCAACAGCT TCATGAACTA ATGACATTGG TAATTTTTTG CGTAGTACAT TTTCATCTCT	900
30	TCGACCTGTG TAATACCTTT GATCAGCTGC ACAAATAGAC AAGTATAAAG GTCTTACGAT	960
	GACATGATTA CCATAAATAT CAACATTATT ATATGTGACG TCGAACTGTC TCGGTGCAAC	1020
35	GAGTTGATAT ACTTGATTAA TCATCGGCAA TATCACCTTG AATAATGGCA TTTGCTACTT	1080
	TTAAATCATA CGGTGTTGTC ACTTTAATGT TGTATAGTTC TCCaCGTACC AATTTAACTG	1140
	CATGTCCAGA TTCGACAATG ATTTTACATG CATCTGATAA GATTTCTTTT TGTTCACTAC	1200
40	TTAAGGCGCG ATAACATCT TGTAATAATT TAATATTAAA TGATTGTGGT GTTTGGCCTT	1260
	GATACATTTT ATTCCTTACA GGGATACTGT GTATGTTCTG TTTATCTTTA GACATTACAA	1320
	TCGTATCAAT TGCTTCAATG ACTGTATCTA CTGCACCATA TTTTGCTGCT ACTTCAATGT	1380
45	TCTCTTTAAT AATACGTTGA GTTAAAAATG GTCTTACGGC ATCATGAGTT ACAATCACAT	1440
	CATCATTATT AATTCCATTT ACATTGCGAA TATGGTCGAT AATGTTTATA ATTGTTTCAT	1500
	TTGATCCGT ACCACCTGCA ACTACTTTGA CACGTTGATC TGTAATGTTA TATTTTTTTA	1560
50	AAATATCCTG TGTATGGGAA ATCCACTGTG CTGGCGTTGC GATAATAATC TCATTAAATT	1620
	CACTCACTAA AATGAACCTC TCAATTGTAT GGATTAAAAT CGGTTTATTA TCAATATCTA	1680

55

	CTGCATAAAT CATGTTGTCC TCCATTCTGT CATTACATCA TTTCCATTTA TACATTACTG	1800
	ACCTATGCCC GCACATAAGC CTAACCTATT GCTCACTTGC CTCTTTTATT AATCCAAAGA	1860
5	TAGTTGTAC AATAGTGTGA TAATTTTTTA TAAAAATGTA TTTTGTAAAC TGACCAATTCT	1920
	AAGTTGTTTT GCCATGCAGT TAATCATTAA CTCTGACGAT ATTAAATTGT TAAAGGTATT	1980
10	AATGTTTACT CTTTTTCAAA TTCATTATTA CTGCCATCAT TTTACCATAT ATTATAATAA	2040
	ATTTATCTTA TTAAGTGGCT GTACTTGATT TTCACTTTAA AAATTATCAA ATATTGCCAT	2100
	CTCATTTTAA GTATACAAA TGCAAAACAA CCGATTACAA AGCATATTTT ACACAAGTAA	2160
15	ACCGGCTATT TATCAACGTA TATTCGAAGA TGAATTATTT CGATAGTATC TATAGACCAG	2220
	ACGGCATTCC CACTTTTATA GCTATAACTA TACCAGCGTT TTCGTCCTCA AAGGTGCATA	2280
	CTAATAAATC GTAAACATGA CTTTATCAAA TCGTTCTTTC TTGTTAACTA ATTTATCAAA	2340
20	TGTCTCCGGG CCTTTTTCTA ACGGTAAAAA ATGAGAAATA ATAGGCTTTA CATTAAATATC	2400
	TTTCGTCTTC ATATAATGTA AGGTTGCCGT CCACTCTTTG CCCGAAAAT TACTGGACAA	2460
	ACAGTTCCAA GAGCCACATA CTGTCAACTC GTTACGCAGA ATTTTTTCAA AATGAACGCG	2520
25	ATCAATCTCA ATATCATCAT ATGGTATTCC GAGTAATACC ACCTCGCCAC CTTTTTTAGG	2580
	TAGCGTCAAT ATTTGACCAA TCGTAACTTT AGCACCTGAT GATTCTATAG CTAAATCGAT	2640
30	TTGATTGGCG TAATGATTTT CGATGAATTT CTCAAGATTT TCTTCTTTTG AATTGATTGT	2700
	TTGATGTGCG CCCAATGATG TTGCAATATC TAGTTTATGC GCATCTATAT CTATAGCGAT	2760
	GATATGTGCA GCACCAAATA TTCGTGCCCA TTGAATAGCT AACAAACCTA TACTGCCACA	2820
35	CCCCATTACT GCAACAGTCA TACCAGGTTG TATATTCGAT TTATAAAACC CATGCGCAAC	2880
	AACGGCTGAT GGCTCAACCA TTGCTGCTTC AATGTAATCA ACATTGTCTG GAACCTTTAA	2940
	AACAATTTGC GCTGGCAATT TGACATATTC CGCGAACGAT CCAGGTTTAT ATGAGCCAAT	3000
40	GACGAATAAC TTTTCACATC GTGCATATTC ACCTTTTAAA CAATACTCGC ATTGATAACA	3060
	AGGTATTGCT GGGCAACCTG TCACTTTGTC GCCCACATTA ACATGCGTAA CATCACTTCC	3120
	AATGGCATCT ACTACACCTG AAAATTCATG ACCAAATGGC ATACCTTTAA TGTATGGCCC	3180
45	CATTTTTTTG TATCGTGACG TGTCTGAACC ACATATGCCA GTCGCTCGTA CTTTAATAAT	3240
	AACGTCATTC GCACTTTCAA TGAATGGCTT TTCATTATCC TCATACCGTA AATCTTCCAC	3300
50	GCCATATAAT TTCAATGCTT TCACTTGTA ATCACCTCAA ATTTGATTTA ATTCACAACT	3360
	TTTTTCTTTT TAAAAATACC TGTCGCAAAA TAACCTGCAA TGACAATGGA ATTACTTACG	3420
55	AGTAAATGTT CCATATAAAA ATCAGTGATT TGTCTTAATG GCCCAAGCAT AAAAGTTAGC	3480

	TGCTTTAATA CCTTCGCCGG ATTTTAAATG TTGATACGCC TCGTCCCATT TCGAAATATC	3600
	ATATATTTTT GTACCAAAG CTTCAGCATT TACTAAACCA TCCGCCATAA GTTGCAATGA	3660
5	AGGTTCOCOA TCTGCTGGCT TTTGACTTCT ACTACCAACA ACTGTTATTT CTTTTTGAAT	3720
	CACTTTTTTC ATATCAAATG GAATTTTCAGC ATCCTTAAAA ATACCTATTT GACTGTAGAA	3780
	ACCTTTTTTG CGTAAAATAT CCAAACCTTG TCGTGCTGCT GGAAGTGCAC CTGAACATTC	3840
10	AACAACAACA TCTGCACCGT AACCGTCTGT AATTCCATTG ATATACGTTT TTAAGTCTGT	3900
	TTGTTGTAAT TTGACTACAT AATCCATGTG CAATGCTTCT GCTTTATCTA ATCTGACTTT	3960
	GTCATTGTCC AATCCAGTTA CCACAACAGT TCGCGCTTTA CTTTTTAACA CTTGTGCTAC	4020
15	AAGTAATCCG ATTGGCCCAG GTCCCATAC AACTGCTACA TCGCCTGAAT TGACTTGAAT	4080
	CTTAGAAACG CCATGATGTG CACATGCTAA TGGTTCTGTC ATAGCTGCAG ACTGATACGA	4140
20	TATTCGTCTG GAATATGATG CAAACTTTCT TCACGTGCAA TGACATAATT AGTAAATGCG	4200
	CCATCAACTT GTGTTCCAAT ACCTTTTCGA TGGTTGCATA AATTATAGTC TTTTGATTTA	4260
	CAGTATTCAC ACTCATTACA AACATAGAAT GTCGTTTCAG ATGTGACACG GTCACCAACT	4320
25	TTAAATCTT TAACGTCTGC TCCAACCTCA ACGATTTAC CAGAAAATTC ATGACCTAAT	4380
	GTCACTGGAA AATTAACCTT ATAATGACCT TCATAAGTAT GAATATCTGT GCCACAAAT	4440
	CCTGCATAAT GTACTTTAAT CTTTACTTTA TCATCTAGCG GTGTTGCAAC TTCTTTATCA	4500
30	AGAAGTTCTA AGTTGCCATG TCCTTCTCTT GTTTTIACTA AAGCTTTTAC CACAAACACC	4560
	TCGATTTTTA ATTGAATAGA CTAAATAGTT TAAAGATAAG ATAGTTAACG ATATTACCAC	4620
	CTTGATCAAT ACTTGAAATT TCAGATGAAC CTTTGGCAT TTGTACATTC GTACCTTTTCG	4680
35	CCATATCTGT GAAAATGGGT GCTACGTCTG TTGCAATATA TAGTGAAATT GCAATCATAA	4740
	TCGTACCCAC AATGACAGAA TGAATAATGT TTCCTCTTGC TGCACCAACA ATAAACGCGA	4800
40	CAACAAATGG TATCGTTGCT AAGTCACCAA AAGGTAGTAC TTGGTTTCCT GTTAAATATA	4860
	CGGCTAATAA AACAGTGATA GGTACTAAAA TTAATGCTGT CGAAATAACT GCTGGATGAC	4920
	CTAATGCTAC AGCCGCATCC AATCCAATAT AAATTTACAG TTCGCCAAAA CGTTTATTTA	4980
45	GCCATGTTCT TGCAGACTCT GAACTGGCA TTAACCTTC CATTAGATT TTTACCATTC	5040
	TAGGCATTAA TACCATTACT GCAGCCATTG ACATTCCTAA ATTAATGATG TCTCCAGTTT	5100
	TGTAACCTGC TAACACACCA ATACCTAAAC CTAAAATTAA GCCGACAAAT ATAGACTCTC	5160
50	CAATGCGCC AAAACGTTTT TGAATTGTTT CAGGATCAGC ATCTAACTTA TTCAGACCGG	5220
	GTACTTTTTG TAACAATTTA ACTAAGTAAA TACCTGGTGC ATAAGAAATT GTACTTCCTG	5280

55

	CTACTTTCAA ACAGATAATT TGGAAAATAA CTGCTGCTAA TAACGCTTGC CAAATACTGC	5400
	CTGATACGGC ATAAACCATT GCTGCTGTAA ACGTATAATG CCAAAAATTC CAAATATCTA	5460
5	CATTTCATCGT CTTTGTCACT TTAGTTACTA GCAATACAAC GTTAACTATG ATTCCGAGTG	5520
	GAATAATAAA TGCTGCGACA GATGATGCCC AAGCGATAGA TGATGTTGCT GGCCAACCTA	5580
10	CATCAATCAC ATTCACTG ACGCCTAAAT TTTTAACCAT CGCTGTGCT GCTGGCCCTA	5640
	AATTTTTAAC TAATAAATCG ATGACTAAGA AAATCCCTAC AAAAGCCACA CCTATTGTTA	5700
	AACCAGACCT AAATGCCGCT CCAATTTTCT GCCTAAAGAA TAGGCCAAGC AAGAATATGA	5760
15	CAACCGGTAA AATAACAGT GCACCTAAAT CTAAAAATCC CCTTACAAA TCAGTGAAGT	5820
	AACTCATATT TAAACCCTCC CTGTTATATA TGCATTGTCA CGATACTTTC CGATTGTGAT	5880
	TACATTTGAC GTTACAGTCA TTTCAACGAC AACCCTTGCT AAATTCGACT GCAGTCCTTT	5940
20	TGAATTACAG TCACTGCGTT TCTATGTCAT CAACAATCAT TTGTCGTGAT AGTCATTTAT	6000
	ATGCAATTG CATATATTAA TATGTTATCG ACCCACGTTA CATATCAATT CCGTTATTTT	6060
	TGTAACCTG TTAAGATTTG TTGTTTGT TCTCAATAC CAATACCAGT TAAGAAATTA	6120
25	CGTGCGTTGA TAACTGGGAA TTTATATTCT TTTTTGTCA TTGCAGTTGT AACTAATAAA	6180
	TCTGCAGTGT CTTCATAAGG TCCAATTCT GTAATTTTGA TTGTTTAAAT ATCTACTTTA	6240
30	ATATTGTGTT CTTTGCCAT TTCTCAATT GCATTATTTA CTACTGTTGA CGTTGCAATA	6300
	CCTGCACCAC ACGCTACTAA TACTGTTC ATTTC AATT CCTCCAATTA ATTTT TAGTT	6360
	ATATTCCAAA TAATCATTGA TTAGTGTTGC TAAAATTGTT TCATCTTTCG TTCGTAGAAT	6420
35	CTGCTCCAAT TTTTCTTCAC TTTGAAAAAT TTGCATCAAC TGTGTGAACA GCTTAAGTTG	6480
	ATCATCTACT TTATCCATTG CTAACATAAA AACGATTTTC ACTTCTGTCT GTTGATCAAG	6540
	TGTTCCTATT TCAATAAAGC GCACTTCTTT TTCTAGAACA GCCACACCTA TCGTTCTATG	6600
40	GTTAATATGT TCGACATCTG TATGCGGTAT AGCGACCGAA CATAGATGCG TTGGTAAACC	6660
	AGTAGCAAAT TCTTTTCTC TGTCGATGAC TGCATCTTTA AACGTTGACT TCACGAACCC	6720
	ATTTTGAAAT AACACATCTG ACATTTGTGA CAATACGGAT TCTTTATCAG TTGCCGACAA	6780
45	ATTGAGCATT ATATTTTCTT TATGCACTAA TTGCTGTCCC ATCCATTTTC CCTCGCTTCT	6840
	TTATTTGAAT AATTTTTTAA AATCTCATTT ACATCAGAAT TTTTGCGACT TTGTATGATG	6900
50	CGCTTAATTG CGTCATTGTC TTGCGCCACA TCTCTCAATT GTAGTAACGC TCTTAAGTGT	6960
	GTCACTTTAT CAACAGCAGC AATAGGTACA ATAATATGGA TTGCTGTGCC ATCTGACATG	7020
55	TATATTGGTT CTTGTAATAT CAACATACTC ATCGCTGTTT TATGTACATG CTTTTCAGAG	7080

	TGCATCTCAT GAATATATTT AATATCAATA AAATGATTAG CAACTAACAC ATCACTTGCT	7200
5	TTAGCAATAG CTTCATCAAT ATTTTCAACA TGATGCATT CTTTCACGTG CCTTGCCGGT	7260
	ATCAAGTCAG CTAAATCTAA TGyCTWATTT tGTGtGACaA TCGATCCATT AATGGTTGAA	7320
	ATTGAATTAT AATTGGCAAT AAAATCTTCT AAACCATCAC GTAGTcTGTA ATGTCATTAA	7380
10	CTGTCGTTGT GCGTTCAATT AATGCCATTA ACTTGTTTAT TTCCTTATCA ATGTCAGCCG	7440
	ATTCCTTATT AATGTACTTC ATCACTTCTT TACGTAACCT TCGTTGCTCA TTTTCAGATA	7500
	AAGCTACTTT TGTGATAAAT AATTTTTTAT GTGTTAGGAC AAACATTGGT GAAAAGACGA	7560
15	TGTCATAATC TAATGTGTAA TTTTCAAATG TTCTAAGTGA AATCGCATCT AAGAAAATAA	7620
	TTTCTGGAAA TAAGTTTCGC AACTCGTATA ACATCATTTG TGATACTGAC GTGCCTTG TG	7680
	TACACACGAT AATAGCTTTT ATCTTGCCAT CGAAGTTTTC ATCTTGACGT CTCAAACTAC	7740
20	CTCCGAACAA CATGGTTAAA TATGCTATTT CATTATCAGG CAACGATTTT CCGAAATATT	7800
	CAGTTAACGA TTGACATGAT TGTTTCACCA TATGAAATAA GGATTGATAA TTTCTTGTA	7860
	AAGGATTTAT TAATTCATCA CGATCCGTTA AGTTATATTT AATCCTATAA AAAGCAGGCG	7920
25	TTAAATGTAA CAAGAGTTGC TGTGATAATT TCTCCTTATC TTCAATGTTA ATAAAAGTGA	7980
	TTTGTTCAAA ATGGTGAATC ATTTGAGCGA TGGCCATCGT TAAATTCGAT ATGCTATCTG	8040
	ATTCTTGCAA ATCAGTCCAT TGCACACTTG TTGAAAGTAA GTGTAATGTC AAATATAACT	8100
30	TTTCCGCTTC TGGCAAATCC GGCTCATGTT GCGTCATAAT CTCCGTTGCT TGATATTCTT	8160
	TCGTATCCCT CAAATACTGA TAATTAATAT TTAATGGATT CATCACATGA CCACCTTGAA	8220
35	TTCGTCTACG AATCACACAA AGGACATAAG GCAATGAACT AAGTGATTTG TCTATAAAGC	8280
	GACTCTTCAA AAATTGTTCT ACCTGTTTGA TCTTGCTTTT TTGATATGCG ATATCTTCGA	8340
	ATGtTAAGTT GAGCGCCTTT AAAACTTCAC TTTTAGTAAT ATCATGATTC AACCTTTGAT	8400
40	CAATCAACTT AATGAAGAAA CGGCGAACTT CAAATTCATC ACCAACAATT TCATAACCAT	8460
	GTTTTCGAGA ATACTTAAGT GACAAACCAT GATTTTCCAA TTGCTCTTTC ACATGATTTA	8520
	TATCGTGAAT GACAGTATTT TTA CTGACTT GTAAATCAAT TGAAAAATGG TTTAGAGACA	8580
45	TTGCGTTTTT CTTACTAAAA AGCATGAGCA TTAAATAATA ACGACGTGTT TCTATGCTAA	8640
	AAATGACATT GTTGCCGTTT AACATTTGCT GCTCCGATAC ATCTCGCTTG AATAACGTCA	8700
	TGATTTCAGA ACTTACAATA AAATTTCCCTT GGCTTGTTCT TTCAAGTTTT GGATAACCCT	8760
50	CTTGTTCAAG CCACAAATTG ATTTTTTGAA TGCGATATCC TAGTTGTCTA CGAGACAAAC	8820
	CAAATATCGA TTCAAGTTCT TTACCATGAA TAGTAGGATT CAATACAATT TCTCTGAGTA	8880

55

TCAATCGTCA CACCGATGTA CACACTTTGA ACACATATTT TCAAAATGAG CATGTACATC 9000
 ATGTGTATGT TTTAACAACA TTTCAATTAT ATCTATATTT TTTGTGATTT TAATCTTTTA 9060
 5 AAATAAAGCA ATTGAAATTT TTGCATATAT TTTTGTGTTT TGTGTTTTTT TGAAGCATT 9120
 TTAACATACA TATCTCAATC ATTATCAAAT TGTGATGACC ATTGTAACCC AATACAAAAA 9180
 CCCTAAGGAC GCTTATATCA GCGCCTTAG GGTAACTGT ATCTATTAA TTAAGTATTA 9240
 10 TTATTCGTAT GTACGTAATC TATGGTCTAT CAAGTCCAC ACTTCTTCAA CATCAACTGC 9300
 TGTAGCAAAA TAAGCATTGG CAGGCTTACC TGTAACATGA TTTAAATCGA CAGCCATAGT 9360
 GCCATAAGTT AGTGGACTTT GATGTTCAAT GTCGATATTA ACGGGTACCA TTGTAAACAA 9420
 15 TTCTGGTTGT AACAAATACA AAATTGTACA AGCATCATGT ATTGGACCAC CATCCATATT 9480
 AAAGTGAGTC TTGTATGTCT TCTTAAAGAA TTGCAATAAT TCTACGACGA ACTGTGCAAC 9540
 20 AGGATTATTG ATACTTTCAA AGCGTTCAAT CACGTGATCG TCGGCTAAAA CTGTGATGTG 9600
 TACATCTAAA CCAAACACAT TTATAGTAAT CCCACTTTCA AAAACACGCT TCGCTGCTTC 9660
 AGCATCTACC CAAATATTGA ATTCTGCTGT AGGCGTCCAA TTTCCAAATG TACCACCACC 9720
 25 CATCAAAGTA ATAGATTCAA TATGCTCAGC GATTCTTGGC TCACGAATCA ATGCCGTGTC 9780
 TACATTCGTA AGAGGACCTG TCGCTACAAT TGTTACAGGT GTATCACTCG TCATCACTTT 9840
 GTTTATAATC ACATCTGATG CTGGCATTGC AACTGCTTGA CGTGATGGTG TCGACGGTAG 9900
 30 TTTCGGACCA TCTAATCCAG ATTCCCCATG TATTTAGAAA GCAAAGGCAG CTGGTTTAAT 9960
 TAACGGCCTA TCCGCACCTT TCGCTACTGC TATATCTTGG CGTCCATAA TATCCAATAC 10020
 GTTCAAGGCG TTTGTGCTAT TCTTGTCAAC TGATTGATTA CCTGCGACTG TTGTTACAGC 10080
 35 TAATATCTCT AGTGGACTGT CAATTGCCCC CGCTAAAATT AATGCTATTG CATCATCGTG 10140
 TCCTGGATCA CAATCCATAA TAATCTTTCT TTTCATTAT ATATCCACCT TTCTTAAGTT 10200
 GTTATCGATA GCTTATGTAT ATTTATTTAT GTGGTGAATC ATGTTTATTT TGAAAAATAG 10260
 40 TTTTAACTTT CTCATATTTT TGGATACAAA CACTATTTAT CTATTTTATG GCTTATAAAT 10320
 TTATCCGATA TGCCTTATCA ACCTACCTCG CTAAAAATAG GATGTCTACA TATCTATACC 10380
 45 GACTTTTGTG AACTCATTTT CACAACAATA TAAACAGCAA TTTATATGAT TGTTACATGA 10440
 TTCAAACAAT TTTTATGAAA AATATTTTCA TACACAGAAT ATATATTGAT ATTAAATTTT 10500
 TCAAAAGCTA TATTGAGAAT AATTAGGAGG GATGTTGATG AAATCTTTAT TTGAAAAGC 10560
 50 ACAGCAGTTC GGCAAGTCCT TTATGTTACC TATCGCAATC TTACCAGCTG CAGGTCTATT 10620
 GTTGGGTATC GGTGGTGCAT TAAGTAATCC AAACACCGTT AAAGCATACC CTATTTTAGA 10680

55

AAATTTACCG GTCATCTTTG CAATTGGTGT CGCAATCGGA TTATCTAGAA GCGATAAAGG 10800
 TACTGCAGGT tTAGctGCGC TGCTCGGTTT CTTAATTATG AACGCAACTA TGAATGGCTT 10860
 5 ATTAACATATC ACGGGCACAT TGGCAAAAGA TCAGCTTGCA CAAAATGGAC AAGGCATGGT 10920
 GCTCGGTATA CAAACGGTTG AAACCGGTGT TTTTGGCGGG ATTATCACAG GTATTATGAC 10980
 CGCAATACTT CACAACAAAT ATCACAAAGT GGTATTACCA CCGTATTTAG GTTTCCTTGG 11040
 10 TGGCTCTAGA TTTGTCCCTA TTGTCACAGC ATTTGCCGCA ATCTTTTATG GTGTATTGAT 11100
 GTTTTTTCAAT TGGCCAAGCA TACAAGCCGG CATTTATCAT GTTGGTGGAT TTGTAACGAA 11160
 AACAGGTGCC ATCGGTACTT TTGTTTATGG CTTTCATCTTA AGATTGTTAG GTCCACTCGG 11220
 15 TTTACACCAT ATTTTTTACT TACCGTTTTG GCAGACGGCA CTTGGTGGTA CTTTAGAAGT 11280
 CAAAGGGCAC TTAGTTC AAG GTACGCAGAA CATCTTCTTT GCTCAACTTG GTGATCCAGA 11340
 20 TGTGACGAAG TATTATTGAG GTGTGTCACG CTTTATGTCA GGCCGTTTTA TTACGATGAT 11400
 GTTCGGCTTA TGTGGTGGCG CACTTGCAAT TTATCACACA GCTAAACCTG AACATAAAAA 11460
 AGTTGTGCGC GGTTTAATGT TATCCGCTGC ACTCACTTCA TTTTAAACAG GTATTACCGA 11520
 25 ACCTTTAGAG TTTAGTTTCT TGTGTCGC ACCTATTCTT TATGTAATCC ATGCCTTCTT 11580
 TGATGGATTA GCATTTATGA TGGCAGACAT TTTCAACATT ACAATTGGTC AAACCTTCAG 11640
 TGGAGGCTTT ATCGATTCTT TACTCTTTGG TGTGCTACAA GGTAATAGTA AAACAAACTA 11700
 30 CCTATACGTC ATACCTATTG GAATTGTGTG GTTCTGTTT TATTACATCG TTTTCAGATT 11760
 CTTAATTACG AAATTTAATT TCAAAACACC TGGTCGAGAA GATAAAGCTG CAGCACAACA 11820
 AGTTGAGGCT ACTGAAAGAG CACAACTAT TGTGCTGGT TTGGGAGGCA AAGATAACAT 11880
 35 TGAAATCGTT GACTGTTGTG CAACGAGACT ACGCGTCACA CTTTCATCAA ATGACAAAGT 11940
 CGATAAAGTA TTAATCGAAA GTACTGGTGC CAAAGGTGTA ATCCAGCAAG GCACTGGTGT 12000
 GCAAGTAATT TATGGGCCTC ACGTTACAGT TATCAAAAAT GAAATTGAAG AATTGCTCGG 12060
 40 GGATTAAGAC TAACCGAAAT ATCAACAGAA CTAATGGCAA CGATGTACGA AGTAAGAAGT 12120
 GACATCGTTG CTTTTATTTT TAATGTTACA TTTGAAGCAT TAAGTTCATC ATGCACTGTA 12180
 45 GTGAGCCCGC AAATCGCCTC TGCTAGACAA TCATCTTAAT GCTATGATTA AAGCTTAAGT 12240
 GCCAGATTTG AATTTAATTT CAACAACGAC TTTCACTACA TTAAAAATAG GGCCACTCGA 12300
 CACATATAGT TGTATCAAAT AGCCCTTTAT ACAATTTTTT GGGTAAGGTT TTACAATTTT 12360
 50 TGGGATGGTA TAGATTTTAT AAAAGTTAT TTAAGTTCTT CTGCTTCAGC CATAATATCT 12420
 TTTAATGTTT TAGCTGAATG TGCGAACTTG CTTTGTCTT CGTCGTTTAA TGGGATTTCT 12480

	TCCTCATATT CGCCTTCTAA TAATGCTGAT ACAGTCAATA CGGCATCTTC ATTTCTGAAA	12600
	ATCGCTTCAG TAATTCTAGC TAATCCCATT GCAACACCAT AATAAGTGGC ACCTTTAGCT	12660
5	TGAATAATGT CATATGCTGC ATCACGTGTT TGAACAAAAA TTTGTTCAAT TTGCGCTTTG	12720
	CCCTCAGGAC GTTGTTC AAG	12780
	CATACTGGTA ATTCACTGTC ACCATGTTCA CCAATAATTT GAGCATCGAC GCTACGTGGC	12840
10	GCAACATCGn AcgyTcGCTT AACAAATAATC TAAAGCGTGC AGAGTCTAAA ATTGTACCAG	12900
	AACCTATAAC ACGTTCTTTA GGTAAACCAG AGAATTTCCA TGTTGCATAC GCTAAAATAT	12960
15	CAACAGGATT TGCTAGCTACC AAGAAAATAC CATCAAATTT TGATGCCATT ACTTCACCAA	13020
	CAATTGATTT GAATATTTTC AAGTTTTTAG ATACTAAATC TAAACGTGTT TCTCCAGGTT	13080
	TTTGTGCAGC ACCAGCACAG ATGACAACTA GATCCGCATC ATGACAATCA CTGTATTCGC	13140
20	CAGCTTTCAC ACGAACTGTT GTTGAGAAAT ATGGTGTGGC ATGTTTTTAA TCCATAACAT	13200
	CTCCTCGAAC TTTTTCAGTG TCTAAATCAA TGATGACTAA TTCATCAACA ATGCTTTGGT	13260
	TCACTAATGA AAATGCGTAG CTTGAACCTA CTGCACCATT ACCTATTAAT ACAACTTTGT	13320
25	TCCCTTTAAA TTTGTTTATT AAAAAAATC CCTTATGATT AATTCATAA CATACATGTA	13380
	GCTTCAAATA TGTTAGTTTA ATGCTGCTTA TTGACGATAC AAAAGCAAAT AAACATCTCT	13440
	TTTATTTTCA ACGCATAACT TAAAGGTCA TGTGTCATCC GCTTTTAAAG TTGTGATTTA	13500
30	TTTCACATAT AAAATGTAAC ATGCATTAAG TACTGGGTCA ATATTAAAT GTGATTTATT	13560
	TCACATTTTA TTTTAATTTT TACACCTTTT TAATTTGTAT mCGATTACAT CTTAGATGTC	13620
	TTTAGTCTTC GTACTTCGCC AGTGATTATT TACACTTTCA CATTTTTTATT ATCATGTTTA	13680
35	CTTTTTTCTA GGAAAACAAC AATGTTTTTT GAATTAGTCA AATAAATGCG CTCAATCGTC	13740
	GGTGTGCAAA CAGACAATTG TACACAATGC TTATTGATAA GTATTTAAAA AATTAAAAAT	13800
40	GTCATACAAT TATCAAATTT GCCATTTTAT TTATATTTTC TCAAACCAAT TAATTGAATA	13860
	TCGAAATTTT TAGTAGAATA ATCAAATAT ACAGATTAAA GGAGGAGTAT CATGCTTACA	13920
	GAACAAGAGA AAGACATTAT CAAACAAACG GTGCCTTTAC TTAAAGAGAA AGGGACAGAA	13980
45	ATTACGTCAA TCTTTTATCC AAAAATGTTT AAAGCGCATC CTGAACTTTT AAACATGTTT	14040
	AATCAAACGA ACCAAAAACG AGGCATGCAA TCTTCAGCAT TAGCACAAGC TGTAATGGCC	14100
	GCAGCGGTTA ATATCGATAA CTTAAGTGTT ATTAAACCAG TCATTATGCC AGTCGCATAT	14160
50	AAACACTGCG CACTACAAGT TTATGCTGAA CATTATCCAA TTGTGGGGAA AAATTTATTA	14220
	AAAGCCATTC AAGACGTGAC AGGATTAGAA GAAAATGACC CTGTCATTCA AGCTTGGGCA	14280

55

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8779 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

```

GGTATTTTnG GAnGGGTACC TAAAGCAATT CCGGCAAAGG GThAATCCAG GTACCGAAAT      60
GGACTTCCCG TTATCGATAA TACCGACATA TATTGTGACA AGTAGATTTT ATGGACATTT      120
AGGCTTACTT TTACTTGTGA TAATTGCATG TATGTTTACT GGTATTTAtC CaTCaATACA      180
TATCATTCAA TTATTGATAT ATGTACCGTT TTGTTTTTTC TTAAGTGCCT CGGTGACGTT      240
ATTAACATCA AACTCGGTG TGTTAGTTAG AGATACACAA ATGTTAATGC AAGCAATATT      300
AAGAATATTA TTTTACTTTT CACCAATTTT GTGGCTACCA AAGAACCATG GTATCAGTGG      360
TTTAATTCAT GAAATGATGA AATATAATCC AGTTTACTTT ATTGCTGAAT CATACCGTGC      420
AGCAATTTTA TATCAGGAAT GGTATTTTAT GGATCATTGG AAATTAATGT TATACAATTT      480
CGGTATTGTT GCCATTTTCT TTGCAATTGG TCGGTACTTA CACATGAAAT ATAGAGATCA      540
ATTTGCAGAC TTCTTGTAAT ATATTATAT GACGAAACCC CGCTAACCAT TAATAAATGG      600
AAGTGGGGTT CATTTTTGTT TATAATTTAA GTAAATAACA TATTAAGTTG GTGTATTATG      660
AACGTTTTAA TAAAGAAATT TTATCATTGG GTAGTTTCGA TACTTTCTAA AATGATTACG      720
CCTCAAGTGA TTGATAAACC GCATATCGTA TTTATGATGA CTTTTCCAGA AGATATTAAG      780
CCTATCATCA AAGCATTAAA TAATTCGTCTG TATCAGAAAA CTGTTTTAAC AACACCAAAA      840
CAAGCGCCTT ATTTATCTGA ACTTAGCGAC GATGTTGATG TGATAGAAAT GACTAATCGA      900
ACATTGGTAA AACAAATTAA GGCTTTGAAA AGCGCGCAGA TGATTATTAT CGATAATTAT      960
TACCTATTGC TAGGTGGATA TAATAAGACT TCTAATCAAC ACATTGTTCA AACGTGGCAT      1020
GCAAGTGGTG CATTAAAAAA CTTTGGCTTA ACAGATCATC AAGTCGATGT GTCTGACAAG      1080
GCAATGGTTC AGCAGTACCG TAAAGTTTAT CAAGCGACGG ATTTTTACTT AGTGGGTTGT      1140
GAACAAATGT CACAATGTTT TAAACAGTCT TTAGGTGCAA CAGAAGAGCA AATGCTGTAT      1200
TTTGGGCTTC CGAGAATTAA TAAATATTAC ACAGCTGATA GAGCAACGGT TAAGGCAGAG      1260
TTAAAGGATA AATATGGAAT TACAAATAAG TTGGTATTAT ATGTACCAAC ATATAGAGAA      1320
GATAAAGCAG ATAATAGGGC TATTGATAAA GCTTATTTTG AAAAATGTTT ACCAGGATAT      1380

```

	ATCGACACGT CTACATTAAT GCTAATGTCA GATATAATTA TTAGCGACTA TAGTTCGCTG	1500
	CCAATAGAAG CTAGCTTGTT AGATATTCCA ACTATATTTT ATGTGTATGA TGAAGGAACA	1560
5	TATGATCAGG TGAGAGGCCT GAATCAATTT TACAAAGCAA TACCGGATAG CTACAAAGTG	1620
	TATACTGAAG AAGATTTAAT AATGACGATA CAAGAAAAAG AACATCTATT AAGTCCGTTA	1680
	TTTAAAGATT GGCATAAGTA TAATACTGAT AAAAGTTTAC ATCAGCTCAC AGAATATATA	1740
10	GATAAGATGG TGACAAAATG AGGTTTACGA TAATCATACC TACATGTAAT AATGAGGCAA	1800
	CAATTCGACA ATTGTTAATA TCTATTGAGA GTAAAGAACA CTATAGAATC CTTTGTATTG	1860
	ATGGTGGTTC TACTGATCAA ACAATTCCTA TGATTGAACG GTTACAAAGA GAACTCAAGC	1920
15	ATATTTTCATT AATACAATTA CAAAATGCTT CGATAGCTAC GTGTATTAAT AAAGGTTTGA	1980
	TGGATATCAA AATGACAGAT CCACATGATA GTGACGCATT TATGGTCATA AAACCAACAT	2040
20	CAATCGTATT GCCAGGTAAA TTAGATAGGT TAACTGCTGC TTTCAAAAAT AATGATAATA	2100
	TTGATATGGT AATAGGGCAG CGAGCTTACA ATTACCATGG TGAATGGAAA TTGAAAAGTG	2160
	CTGATGAGTT TATTAAAGAC AATCGAATCG TTACATTAAC GGAACAACCA GATTTGTTAT	2220
25	CAATGATGTC TTTTGACGGA AAGTTATTCA GTGCTAAATT TGCTGAATTA CAGTGTGaCG	2280
	AAACTTTAGC TAACAaCATAC AATCACGCAA TACTTGTCaA GGCGATGCAA AAAGCTACGG	2340
	ATATACATTT AGTTTCACAG ATGATTGTCG GAGATAACGA TATAGATACA CATGCTACAA	2400
30	GTAACGATGA AGATTTTAAT AGATATATCA CAGAAATTAT GAAAATAAGA CAACGAGTCA	2460
	TGGAAATGTT ACTATTACCT GAACAAAGGC TATTATATAG TGATATGGTT GATCGTATTT	2520
	TATTCAATAA TTCATTAAAA TATTATATGA ACGAACACCC AGCAGTAACG CACACGACAA	2580
35	TTCAACTCGT AAAAGACTAT ATTATGTCTA TGCAGCATTG TGATTATGTA TCGCAAAACA	2640
	TGTTTGACAT TATAAATACA GTTGAATTTA TTGGTGAGAA TTGGGATAGA GAAATATACG	2700
40	AATTGTGGCG ACAAACATTA ATTCAAGTGG GCATTAATAG GCCGACTTAT AAAAAATTCT	2760
	TGATACAACCT TAAAGGGAGA AAGTTTGCAC ATCGAACAAA ATCAATGTTA AAACGATAAC	2820
	GTGTACATTG ATGACCATAA ACTGCAATCC TATGATGTGA CAATATGAGG AGGATAAATT	2880
45	AATGAAACGT GTAATAACAT ATGGCACATA TGACTTACTT CACTATGGTC ATATCGAATT	2940
	GCTTCGTCGT GCAAGAGAGA TGGGCGATTA TTTAATAGTA GCATTATCAA CAGATGAATT	3000
	TAATCAAATT AAACATAAAA AATCTTATTA TGATTATGAA CAACGAAAAA TGATGCTTGa	3060
50	ATCAATACGC TATGTCGATT TAGTCATTCC AGAAAAGGGC TGGGGACAAA AAGAAGACGA	3120
	TGTCGAAAAA TTTGATGTAG ATGTTTTTGT TATGGGACAT GACTGGGAAG GTGAATTCGA	3180

	TAAATCAAA CAAGAATTAT ATGGTAAAGA TGCTAAATAA ATTATATAGA ACTATCGATA	3300
	CTAAACGATA AATTAACCTA GGTATTATA AAATAAATAT AAAACGGACA AGTTTCGCAG	3360
5	CTTTATAATG TGCAACTTGT CCGTTTTTAG TATGTTTTAT TTTCTTTTTC TAAATAAACG	3420
	ATTGATTATC ATATGAACAA TAAGTGCTAA TCCAGCGACA AGGCATGTAC CACCAATGAT	3480
	AGTGAATAAT GGATGTTCTT CCCACATACT TTTAGCAACA GTATTTGCCT TTTGAATAAT	3540
10	TGGCTGATGA ACTTCTACAG TTGGAGGTCC ATAATCTTTA TTAATAAATT CTCTTGGATA	3600
	GTCGCGTGT ACTTTACCAT CTTGACTAC AAGTTTATAA TCTTTTTTAC TAAATCACT	3660
	TGGTAAACA TCGTAAAGAT CATTTTCAAC ATAATATTTT TTACCATTTA TCCTTTGCTC	3720
15	ACCTTTAGAC AATATTTTTTA CATATTTATA CTGATCAAAAT GAGCGTTCCA TTAATGCATT	3780
	CCCCATCATA TTACGTTGCT TCTCGCCACC AAGGTTTTTA TAGTCTCCTG CACCCATGAT	3840
	AACCTGATTA ATTCTAAATT TACCTCGTTT GGTAAGTATC GTATGGTTGT AATTGCTGT	3900
20	ATCACTTGAT CCAGTTTTTA AACCATCTGT ACCCGGCAAA CTCATTTTTG CACCTTCCAA	3960
	TGAAAAGTTG AATGTGTAAT ACGTAACTGC ATGCGTTGTT GGTGCTAACT GCTTTGTAAA	4020
25	GTCTAATATT TTAGGTGTCT CTTTAATCAC GTGTAAATCT AAAATGGCAT AGTCTCTAGC	4080
	AGTCGTTACA GTACGTTCTT GGTCTTTATA CTTTGTGGT GCAAATGTAC GTAATCTGA	4140
	ATTTTCAGCA CCGTTGGAT TGACGAAATG TGTATTTTTT ATTCCGATAG CTTTAGCTTT	4200
30	GTTATTCATT AAATCAACGA AATCGCTGGT GTTTTTTGAA ACCTTCTTAG CTAAAATTAA	4260
	TGCCGCGGCA TTAGTAGAAT TAGATACTGT AATTGTAAAT AGGTCTGCGA TTGTCCATAC	4320
	TTGTCCAGGA TATAGTTTCG TATTACTCAA CTCAGGTAGT GTAGACATAA TATATTCTTT	4380
35	GTTTCGTCATT GTGACTGTGT CATCAAGTGA AAGCTGCCCC TTATTTACAG CTTCCAATGT	4440
	TAAGTACATT GTCATTAATT TAGTCATAGA CGCTGGATTC CACTTAGTAT CGATATTGTA	4500
	TTGATACAGT AATTGTCCAG TTTGACTTAC ATTAACAGCA CTCGTCGGTT CGTATGCAGC	4560
40	CGACAAACCT GCATAACCAT ATTGATTTGC TGCTTGATCA GGGGTTACGT CACTGTTAGT	4620
	AGCTTGTCGA TATGGTGTCA TAATACTTAA TGTAAACAT AAAATGATGA TAATAGATAT	4680
45	TAAATTTTTT ATAAAGCGTT AATCTTCCCT TTCCAATTC TTAAATATTC CCTAAAAGCA	4740
	ATGGTTATTC CTAATTACGG AAATCATTGC TAATTCACCT CACCTTAATT AAATTGTTGA	4800
	AAATAAAGTT TTCTGCAGTT AATTGAAAA ATAATGCAAA TATATTACGT GTGTAGCTAA	4860
50	AGGTGTTATA ATGTTGTAC GAAGAGCAAA CTTACTCAA AGCGATTAAT TTTCATGTTT	4920
	TAATATAAAG ACTTTGAGAA GTTATTACAA AAAATGCAAT AGAAATATTC TATCATATAA	4980

55

	AAGTATATGA TAGAAATGCA TGTATCTATC TAAATGAATT AACTATAAAT TTCAAACAGA	5100
	AGAGGTAAAA CTATGAAACG AGAAAATCCA TTGTTTTTCT TATTTAAAAA ACTATCATGG	5160
5	CCAGTGGGTC TTATCGTTGC AGCTATCACT ATTTTCATCAC TAGGGAGCTT AAGTGGACTA	5220
	TTAGTGCCAC TGTTTACTGG ACGAATTGTA GATAAATTTT CCgTGAGCCA TATCAATTGG	5280
	AATCtAATCG CATTATTTGG TGGTATCTTT GTCATCAATG CTTTATTAAG CGGATTAGGT	5340
10	TTATATTTAT TAAGTAAAAT TGGTGAAAAG ATTATTTATG CGATACGCTC AGTTTTATGG	5400
	GAGCATATCA TACAATTAAA AATGCCATTG TTTGACAAAA ATGAAAGTGG TCAATTAATG	5460
15	AGTCGATTAA CTGACGATAC GAAAGTGATA AATGAATTTA TTTACAAAA GCTACCTmAC	5520
	TTATTACCAT CAATCGTTAC ATTAGTTGGG TCACTAATCA TGTTATTTAT TTTAGATTGG	5580
	AAAATGACAT TATTAAcATT TATAACGATA CCGATATTG TTTTaATTAT GATTCCTCTA	5640
20	GGTCGTATTA TGCAAAAGAT ATCGACAAGT ACACAATCTG AAATTGCAAA CTTcAGTGGT	5700
	TTGTTAGGGC GTGTCCTAAC TGAAATGCGT CTTGTTAAAA TATCAAATAC AGAGCGTCTT	5760
	GAATTAGATA ATGCACATAA AAATTGAAAT GAAATATATA AATTAGGTTT AAAACAGGCT	5820
25	AAAATTGCGG CAGTTGTACA ACCAATTTCA GGTATAGTTA TGTTGCTAAC AATTGCAATT	5880
	ATTTTAGGTT TTGGTGcATT AGAAATTGCG ACTGGTGCAA TCACTGCAGG TACATTAA TT	5940
	GCAATGATAT TTTATGTTAT TCAGTTATCT ATGcCTTTAA TCAATCTTTC CACGTTAGTT	6000
30	ACAGATTATA AAAAGGCAGT CGGTGCAAGT AGTAGAATAT ACGAAATCAT GCAAGAACCT	6060
	ATTGAACCGA CAGAAGCTCT TGAAGATTCT GAAATGTAT TAATTGATGA CGGTGTATTG	6120
	TCATTTGAAC ATGTAGACTT TAAATATGAT GTGAAGAAAA TATTAGATGA TGTGTCGTTc	6180
35	CAAATCCcAC AAGGTCAAGT GAGTGCTTTT GTAGGCCCTT CTGGGTCTGG TAAAAGTACG	6240
	ATATTTAATC TGATAGAACG TATGTATGAA ATTGAGTCAG GTGATATTAA ATATGGCCTT	6300
40	GAAAGTGTCT ATGATATCCC GTTATCTAAG TGGCGACGCA AAATTGGATA TGTTATGCAA	6360
	TCAAATTcGA TGATGAGTGG TACAATTAGA GACAATATTT TATACGGAAT TAATCGTCAT	6420
	GTTTCAGATG AAGAACTTAT TAATTATGCT AAATTAGCGA ACTGTCATGA TTTTATCATG	6480
45	CAATTTGATG AAGGATATGA CACGCTTGTA GGTGAACGAG GATTGAAACT GTCTGGCGGA	6540
	CAACGTCAAC GTATTGATAT TGCTAGAAGT TTTGTTAAAA ATCCTGATAT TTTGTTACTT	6600
	GATGAAGCAA CAGCTAATCT CGATAGTGAA AGTGAATTGA AAATTCAAGA AGCTTTAGAA	6660
50	ACATTGATGG AAGGTAGAAC AACGATTGTC ATTGCGCATC GTTTGTCTAC AATTAAAAAA	6720
	GCCGGTCAAA TTATATTCTT AGACAAAGGA CAGGTAACAG GTAAAGGTAC GCATTcAGAA	6780

55

	TTTTATATAT ATAAGTAAGC TTGGAGCAAA TACACATATA CCATCGAGGA AATTAAAGTG	6900
	TGGCACATTG ATGGATATAG ATGTTAATAA ATTGCTTCAA GCTTTTGTCT ATTTTAAATC	6960
5	ATTTGAGAAG TTACGACATA ATAATTCTTA AATTAATGAA ATCGATATTT TAAGAAAAAA	7020
	ATGCTCATGG TATAATACAA GTTATAAGCA AACATACATA TATTAAATAC TGTAGCCACG	7080
	AGTCATAATT CTTCATATTT TACATAGCAA TTAACTGAT TTTAGAGTCC ACGGTACAGA	7140
10	AGTTTGATAT TTCAATGTTT CTAAATTTTT AAAAATTAA ATCATAGGTG GGTGCCAAAT	7200
	GTTTTTATTA ATCAACATTA TTGGTCTAAT TGTATTTCTT GGTATTGCGG TATTATTTTC	7260
	AAGAGATCGC AAAAATATCC AATGGCAATC AATTGGGATC TTAGTTGTTT TAAACCTGTT	7320
15	TTTAGCATGG TTCTTTATTT ATTTTGATTG GGGTCAAAA GCAGTAAGAG GAGCAGCCAA	7380
	TGGTATCGCT TGGGTAGTTC AGTCAGCGCA TGCTGGTACA GGTTTTGCAT TTGCAAGTTT	7440
20	GACAAATGTT AAAATGATGG ATATGGCTGT TGCAGCCTTA TTCCCAATAT TATTAATAGT	7500
	GCCATTATTT GATATCTTAA TGTACTTTAA TATTTTACCG AAAATTATTG GAGGTATTGG	7560
	TTGGTTACTA GCTAAAGTAA CAAGACAACC TAAATTCGAG TCATTCTTTG GGATAGAAAT	7620
25	GATGTTCTTA GGAAATACTG AAGCATTAGC CGTATCAAGT GAGCAACTAA AACGTATGAA	7680
	TGAAATGCGT GTATTAACAA TCGCAATGAT GTCAATGAGC TCTGTATCGG GAGCTATTGT	7740
	AGGTGCGTAT GTACAAATGG TACCAGGAGA ACTGGTACTA ACGGCAATTC CACTAAATAT	7800
30	CGTTAACGCG ATTATTGTGT CATGCTTGTT GAATCCAGTA AGTGTGAAG AGAAAGAAGA	7860
	TATTATTTAC AGTCTTAAAA ACAATGAAGT TGAACGTCAA CCATTCTTCT CATTCCTTGG	7920
	AGATTCTGTA TTAGCAGCAG GTAAATTAGT ATTAATCATC ATCGCATTTG TTATTAGTTT	7980
35	TGTAGCGTTA GCTGATCTAT TTGATCGTTT TATCAATTTG ATTACAGGAT TGATAGCAGG	8040
	ATGGATAGGC ATAAAAGGTA GTTTCGGTTT AAACCAAATT TTAGGTGTGT TTATGTATCC	8100
	ATTTGCGCTA TTAATCGGTT TACCTTATGA TGAAGCGTGG TTGGTAGCAC AACAAATGGC	8160
40	TAAGAAAATT GTTACAAATG AATTTGTTGT TATGGGTGAA ATTTCTAAAG ATATTGCATC	8220
	TTATACACCA CACCATCGTG CGGTTATTAC AACATTCTTA ATTTCAATTTG CAAACTTCTC	8280
45	AACGATTGGT ATGATTATCG GTACATTGAA AGGCATTGTT GATAAAAAGA CATCAGACTT	8340
	TGTATCTAAA TATGTACCTA TGATGCTATT ATCAGGTATC CTAGTTTCAT TATTAACAGC	8400
	AGCTTTCGTT GGTTTATTTG CATGGTAATA TGTCGAAGAG TGAATATGAT AATACATTTT	8460
50	AACATAATAA TATGTCCAGG CATGTCGTCT ATTGATATAG GTGAGATGCT TGGACTTTTT	8520
	TATTATTGAT ATAAAGGTAT nTAAATATTT TTAAAGTTAC CGAAATTGAA GCATTATAAA	8580

55

GACAGTAAGG ACTAGGTACA GTCATAGTAC TTCGAGCAAA ATTTGTTTTG TTATTATAAA 8700
 CAACACAAAG GAGATAACTT CTCTANTGAA GAAGTTAAAA ACATTATAGC AGACAATGAA 8760
 5 ATGAAAGTAA ATTAAAAAT 8779

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31096 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GTTGCAGTAG TCAAAGAATT AAACAAGGTG AAGGCGTGTA GCTTGCACAC CCGAAAATGT 60
 20 GCGTAAGTTA aCGGATGCAG GACATAAAGT AATTGTTGAA AAAAATGCTG GCATTGGTTC 120
 AGGATTTTCT AACGATATGT ATGAAAAAGA AGGCGCTAAG ATCGTAACTC ACGAACAAGC 180
 ATGGGAAGCT GATCTTGTTA TCAAAGTAAA AGAACCTCAT GAAAGCGAAT ATCAATATTT 240
 25 CAAAAAGAAT CAAATTATCT GGGGATTTTT ACATCTAGCA TCTTCAAAAG AAATAGTAGA 300
 AAAAATGCAA GAAGTTGGTG TAACTGCGAT TAGTGGTGAA ACCATTATAA AAAATGGAAA 360
 AGCAGAATTA TTAGCGCCAA TGAGTGCTAT AGCAGGTCAA CGCTCAGCAA TTATGGGAGC 420
 30 TTACTACTCT GAAGCACAAC ATGGTGGTCA AGGTACTTTA GTGACTGGTG TACATGAAAA 480
 TGTGGATATA CCTGGTAGTA CATATGTGAT TTTCGGTGGT GGAGTAGCAG CAACAAATGC 540
 AGCAAATGTT GCCTTGGGAC TAAATGCTAA AGTAATCATT ATCGAGTTAA ACGATGACCG 600
 35 CATTAAATAT CTTGAAGATA TGTATGCAGA AAAAGATGTC ACAGTAGTCA AATCAACACC 660
 AGAAATTTTA GCAGAACAAA TTAAGAAAGC AGATGTATTT ATTTCTACAA TTTTAATTTT 720
 AGGTGCGAAA CCGCCAAAAT TGTTACTCG TGAGATGGTT AAATCAATGA AAAAAGGTTC 780
 40 AGTATTAATC GATATAGCTA TTGACCAAGG TGGAACTATT GAAACAATTA GACCAACTAC 840
 AATTTCTGAT CCAGTGTATG AAGAAGAAGG TGTGATTCAT TATGGTGTAC CAAATCAACC 900
 45 AGGAGCAGTC CCAAGAACTT CAACAATGGC ATTAGCACAA GGAAATATTG ATTATATATT 960
 AGAAATTTGT GACAAAGGCT TAGAACAAGC AATTAAAGAT AATGAAGCCT TAAGTACTGG 1020
 TGTAACATT TACCAAGGAC AAGTGACAAA TCAAGGATTA GCTTCATCAC ATGACCTAGA 1080
 50 TTATAAGAA ATATTAAATG TTATCGAATA GATAGTAATT TAAATGAAAT TGAGTGAAAT 1140
 GAATATTTTA AATATAGCAT TATAGTTTGG ACTAAAAATT TACAAAACGG AAGGATGTAA 1200

	TCGAAGAAGC TAAAGCAAGC ATTAAACCAT TTATTTCGTCG AACACCTCTA ATTAAATCAA	1320
	TGTATTTAAG CCAAAGTATA ACTAAAGGGA ATGTATTTCT AAAATTAGAA AATATGCAAT	1380
5	TCACAGGATC TTTTAAATTT AGAGGCGCTA gCAATnAAAA TTAATCACTT AACAGATGAA	1440
	CAAAAAGAAA AAGGCATTAT CGCAGCATCT GCTGGGgAAC CATGCACAAG GTGTTGCTTT	1500
	AACAGCTAAA TTATTAGGCA TTGATGCAAC GATTGTAATG CCTGAAACAG CACCACAAGC	1560
10	GAAACAACAA GCAACAAAAG GCTATGGGGC AAAGGTTATT TTAAAAGGTA AAAACTTTAA	1620
	CGAAACTAGA CTTTATATGG AAGAATTAGC GAAAGAAAAT GGCATGACAA TCGTTCATCC	1680
	ATATGACGAT AAGTTTGTAA TGGCAGGCCA AGGAACAATT GGTTTAGAAA TTTTAGATGA	1740
15	TATTTGGAAT GTGAATACAG TCATCGTACC AGTTGGCGGT GGAGGATTAA TTGCAGGTAT	1800
	TGCCACCGCA TTAAAATCAT TTAACCTTC AATTCATATT ATCGGTGTTT AATCTGAGAA	1860
20	TGTTTCATGGT ATGGCTGAGT CTTTCTATAA GAGAGATTTA ACTGAACATC GAGTGGATAG	1920
	CACAATAGCA GATGGTTGTG ATGTAAAAGT TCCTGGTGAA CAAACATATG AAGTAGTTAA	1980
	ACATTTAGTA GATGAATTTA TTCTTGTTAC TGAAGAAGAA ATTGAACATG CTATGAAAGA	2040
25	TTTAATGCAG CGTGCCAAAA TTATTACTGA AGGTGCAGGC GCATTACCAA CAGCTGCAAT	2100
	TTTAAGTGGA AAAATAAACA ATAAATGGCT TGAAGATAAA AATGTTGTTG CATTAGTTTC	2160
	AGGCGGGAAT GTTGACTTAA CTAGAGTTTC AGGTGTCATT GAACATGGAC TGAATATTGC	2220
30	AGATACAAGC AAGGGTGTGG TAGGTTAAAA CATTTAATCT TAAAAATGAG GTGTAATTAT	2280
	GTCAAATGGT AAAGAATTAC AAAAAAATAT AGGTTTCTTC TCAGCGTTTG CTATTGTTAT	2340
	GGGGACAGTT ATTGGTTCAG GAGTATTCTT TAAATATCA AACGTAACAG AAGTAACAGG	2400
35	AACAGCAGGA ATGGCCTTGT TTGTATGGTT CCTAGGCGGC ATCATTACCA TTTGTGCGGG	2460
	GTTAAACAGCA GCAGAACTTG CTGCTGCAAT CCCTGAAACA GGTGGCTTAA CGAAGTATAT	2520
	AGAATATACA TACGGTGATT TCTGGGGCTT CCTATCAGGT TGGGCGCAAT CATTATTTTA	2580
40	TTTTCCAGCT AACGTAGCAG CATTGTCTAT CGTATTGCG ACACAGCTAA TTAATTTATT	2640
	CCATTTATCT ATAGGTTTCGT TAATACCAAT AGCAATCGCA TCTGCGTTAT CTATTGTGTT	2700
45	GATAAATTTT CTAGGTTCAA AAGCAGGCGG AATTTTACAA TCAGTTACTT TAGTAATTAA	2760
	ACTGATTCCA ATCATCGTTA TTGTAATTTT TGGTATTTT CAATCTGGAG ATATCACTTT	2820
	TTCATTAAAT CCAACTACAG GTAATTCaGG AAATGGCTTC TTTACAGCAA TTGGTAGTGG	2880
50	TTTATTAGCA ACTATGTTTG CATATGATGG TTGGATTCAT GTAGGAAATG TTGCGGGGGA	2940
	ACTTAAAAAT CCTAAACGCG ATTTACCTTT AGCGATTTCA GTTGGTATCG GTTGTATTAT	3000

55

	TGGTAATTTA AATGCAGCTT CAGATACATC AAAAATATTA TTTGGTGAAA ATGGCGGTAA	3120
	GATTATTACA ATCGGTATAT TAATTTCTGT TTATGGTACG ATCAATGGCT ATACTATGAC	3180
5	TGGTATGCGC GTACCATATG CAATGGCTGA AAGAAAATTA TTGCCATTTA GCCATTTATT	3240
	CGCAAAATTA ACAAATCTG GCGCACCATG GTTTGGCGCA ATTATACAAC TTATAATCGC	3300
	TATCATCATG ATGTCAATGG GAGCATTTGA TACAATTACA AATATGTTAA TCTTTGTTAT	3360
10	TTGGTTGTTT TATTGTATGT CATTGTGTGC GGTAAATAATT TTAAGAAAAC GTGAACCAA	3420
	TATGGAACGA CCATATAAAG TACCGTTATA TCCGATCATA CCTTTAATTG CTATTTTGGC	3480
	AGGATCATTT GTATTAATTA ATACACTGTT TACACAATTT ATATTAGCAA TCATTGGAAT	3540
15	TCTAATAACA GCACTTGGTA TACCAGTTTA TTACTATAAA AAGAAACAAA AAGCAGCATA	3600
	AGGTAAGATA ACTAGCATTG AGAATAAATG GATGGACTAC TAATAAATTT AAAGTTTAC	3660
20	ACATTAAAT CAAAAACCAT TCAATTATTC TATGGAACAG ACAAATTTCT GTTATGGAAT	3720
	TTGTCTGTTT TTCAAAGTA TAGGGAGGCA AATAGAGATG GAAAAGCCGT CAAGAGAGGC	3780
	ATTTGAAGGC AATAATAAGT TGTTAATAGG AATTGTTCTA AGTGTAATAA CGTTTTGGCT	3840
25	ATTTGCACAA TCATTGGTTA ATGTTGTACC AATACTTGAA GATAGTTTCA ATACAGATAT	3900
	TGGAACGGTT AATATCGCCG TTAGTATAAC TGCTTTATTT TCAGGAATGT TTGTAGTAGG	3960
	AGCAGGTGGT CTTGCTGATA AATATGGCAG AATTAACTC ACGAACATTG GTATTATCTT	4020
30	AAATATATTA GGTTCATTAT TAATCATTAT TTCAAATATT CCTTTATTAC TTATTATAGG	4080
	AAGATTAATT CAAGGACTTT CAGCAGCATG TATTATGCCT GCAACTTTGT CTATTATTAA	4140
	GTCATATTAC ATTGGGAAAG ATAGACAACG CGCTTTAAGT TATTGGTCAA TTGGCTCATG	4200
35	GGGCGGCTCT GGTGTTTGTT CATTTTTTGG AGGTGCAGTT GCAACGCTTT TAGGTTGGCG	4260
	TTGGATTTTC ATCCTATCAA TTATAATTC ATTAATTGCA CTGTTTCTTA TTAAAGGCAC	4320
	ACCTGAACT AAATCTAAAT CGATTTCTCT AAATAAATTT GACATTAAAG GTCTGGTTCT	4380
40	TTTAGTCATT ATGCTCCTCA GTTTAAATAT TTTAATTACT AAAGGATCAG AATTAGGTGT	4440
	AACCTCACTT CTTTTATTA CTTTATTAGC TATTGCAATT GGATCTTTTA GTTTATTTAT	4500
45	AGTTCCTGAA AAGCGTGCTA CAAATCCTTT AATCGATTTT AAATTATTTA AAAATAAAGC	4560
	TTACACAGGT GCAACAGCTT CAAACTTTTT GTTAAATGGT GTTGCAGGAA CATTAAATAGT	4620
	AGCCAACACA TTTGTTCAA GAGGTTTAGG ATATTCTTCA TTGCAAGCAG GAAGTTTATC	4680
50	AATCACTTAT TTAGTAATGG TACTAATTAT GATTCGTGTT GGTGAAAAGT TACTTCAAAC	4740
	ACTCGGATGC AAGAAACCAA TGTTAATTGG AACAGGAGTT CTTATTGTCT GAGAATGTCT	4800

55

	ATTCTTTGGT TTAGGACTAG GGATATATGC TACACCATCA ACAGATACAG CAATTGCAAA	4920
	TGCACCGTTA GAAAAAGTAG GCGTTGCTGC AGGTATCTAT AAAATGGCTT CTGCATTAGG	4980
5	TGGAGCATT TGGCGTCGCAT TGAGTGGTGC AGTATATGCA ATCGTATCAA ATATGaCAAA	5040
	CATTTATACA GGTGcAATGa TTGnCATTAT GGTtaAATGC AGGTATGGGa ATATTATCaT	5100
	TCGTTATCAT TTTGtTACTT GTGcCTAAAC mAAACGACAC TCAATTATGA TAATTGAGAA	5160
10	TTAAATTGAA ATCATACAAG TCGCTACAAT ATTAAACAAA AATATAAACC GATTCTTATG	5220
	TGTCATTATT TTAAATGAAC ATAGGGATTG GTTTTTTATT ACTCTTTTAC GCTACTTTAT	5280
15	TTATAATTAT TATAAATTGT CACAAATTCA ATTTACCTTA CAATATATTT TGTGTTATTA	5340
	TATTCTGGAG CATAAATAAA TTGTTCAACA CATAGTTGTA ATGTGTTTCA ATACTTTTGT	5400
	GATAGATTGC GAAATTGTAT TGAATCGTCA TCGTTTTTAA TTTTAAATG AGAATGGAAT	5460
20	GAGCATTACA ATACACAAGC AATCAAAAGT AAATACATTC ACAACACAAC AGAGACATAA	5520
	CAACAAGATA AGGAGTGAAC AATAGCTGTG AATTATCGTG ATAAAATTCA AAAGTTTAGT	5580
	ATTCGTAAAT ATACAGTTGG TACATTTTCA ACTGTCATTG CGACATTGGT ATTTTTAGGA	5640
25	TTCAATACAT CACAAGCACA TGCTGCTGAA ACAAATCAAC CAGCAAGCGT GGTAAACAG	5700
	AAACAACAAA GTAATAATGA ACAGACTGAG AATCGAGAAT CTCAAGTACA AAATTCTCAA	5760
	AATTCACAAA ATGGTCAATC ATTATCTGCT ACTCATGAAA ATGAGCAACC AAATATTAGT	5820
30	CAAGCTAATT TAGTAGATCA AAAAGTAGCG CAATCATCTA CTAATAATGA TGAACAACCA	5880
	GCATCTCAA ATGTAAATAC AAAGAAAGAT TCGGCAACGG CTGCGACAAC ACAACCAGAT	5940
35	AAAGAACAAA GTAAGCATAA ACAAACGAA AGTCAATCTG CTAATAAAAA TGGAAACGAC	6000
	AATAGAGCGG CTCATGTAGA AAATCATGAA GCAAATGTAG TAACAGCTTC AGATTCATCT	6060
	GATAATGGTA ACGTACAACA TGACCGAAAT GAATTACAAG CGTTTTTTGA TGCAAATTAT	6120
40	CATGATTATC GCTTTATTGA CCGTGAAAAT GCAGATTCTG GCACATTTAA CTATGTAAAA	6180
	GGCATTTTTG ATAAGATTAA TACGTTATTA GGCAGTAATG ATCCAATAAA CAATAAAGAC	6240
	TTGCAACTTG CATACAAAGA ATTGGAACAA GCTGTTGCTT TAATTCGTAC AATGCCTCAA	6300
45	CGTCAACAGA CTAGCCGACG TTCAAATAGA ATTCAAACGC GTTCGGTTGA GTCAAGAGCT	6360
	GCAGAGCCTA GATCAGTATC AGACTATCAA AATGCAAATT CATCATATTA TGTGAAAAT	6420
	GCTAATGATG GTTCGGGCTA TCCTGTTGGT ACATATATCa ATGCTTCTAG TAAAGGGGCG	6480
50	CCATATAATT TACCAACTAC ACCATGGAAT ACATTGAAGG CCTCTGACTC AAAGGAAATT	6540
	GCTCTTATGA CAGCGAAACA AACTGGAGAC GGGTACCAAT GGGTTATTAA GTTTAATAAA	6600

55

	GTAGGAAGAA CTGACTTTGT AACAGTTAAT TCAGATGGAA CAAATGTACA ATGGAGTCAT	6720
	GGAGCAGGAG CAGGTGCAAA TAAACCACTT CAACAAATGT GGGAATATGG AGTAAATGAT	6780
5	CCTCATCGTT CACATGACTT TAAAATAAGA AATAGAAGTG GCCAAGTAAT ATATGACTGG	6840
	CCAAGTGTCC ATATTTATTC TTTAGAAGAT TTATCTAGAG CGAGTGATTA TTTTAGTGAA	6900
	GCTGGAGCGA CACCTGCTAC TAAAGCTTTT GGTAGACAAA ATTTTGAATA TATTAATGGT	6960
10	CAAAAACCTG CTGAATCACC GGGTGTTCCT AAAGTTTATA CTTTCATCGG TCAAGGTGAT	7020
	GCAAGTTATA CAATTTCAAT TAAAACACAA GGTCCAAGTG TTAATAAATT GTACTATGCA	7080
	GCAGGTGGGC GTGCTTTAGA GTACAATCAA TTATTTATGT ACAGTCAACT ATACGTCGAA	7140
15	TCAACGCAAG ACCATCAACA ACGTCTTAAT GGTTTAAGAC AAGTGGTTAA TCGTACATAT	7200
	CGCATAGGTA CAACTAAACG TGTAGAAGTG AGTCAAGGAA ATGTACAAAC GAAAAAGGTA	7260
	TTAGAAAGTA CAAACCTAAA TATAGATGAT TTTGTTGATG ATCCTTTAAG TTATGTTAAG	7320
20	ACGCCGAGTA ATAAAGTGTT AGGATTTTAT TCGAATAATG CAAATACTAA TGCTTTTAGA	7380
	CCGGGTGGAG CCCAACAATT AAATGAATAT CAATTAAGTC AATTATTTAC TGATCAAAAA	7440
25	TTACAAGAAG CAGCAAGAAC TAGAAACCCA ATAAGATTAA TGATTGGTTT CGACTATCCT	7500
	GATGCTTATG GTAATAGTGA ACTTTAGTTC CTGTTAACCT AACGGTATTA CCTGAAATCC	7560
	AACATAATAT TaAATTCTTT AAAAATGACG ATACTCAAAA TATTGCTGAA AAACCATTTT	7620
30	CAAAACAAGC TGGGCATCCA GTTTTCTATG TATATGCAGG TAACCAAGGG AATGCTTCCG	7680
	TGAATTTAGG TGGTAGCGTA ACATCTATTC AACCATTACG TATTAATTTA ACAAGTAATG	7740
	AGAATTTTAC AGATAAAGAT TGGCAAATTA CAGGTATTCC GCGTACATTA CACATTGAAA	7800
35	ACTCGACAAA TAGACCTAAT AATGCCAGAG AACGCAATAT TGAACCTGTT GGTAACCTAT	7860
	TACCGGGGGA TTACTTTGGA ACGATACGTT TTGGACGTAA AGAACAATTA TTCGAAATTC	7920
	GTGTTAAACC ACATACACCA ACAATTACAA CGACAGCTGA GCAATTAAGA GGTACAGCAT	7980
40	TACAAAAAGT GCCTGTTAAT ATTTGGGGAA TACCGTTGGA TCCATCGGCA TTGGTTTATT	8040
	TAGTTGCACC AACAAATCAA ACTACGAATG GTGGTAGTGA GGCAGATCAA ATACCATCTG	8100
	GTTATACGAT ACTTGCGACT GGTACACCTG ATGGGGTGCA TAATACAATT ACTATACGAC	8160
45	CGCAAGATTA TGTTGTATTC ATACCACCTG TAGGTAAACA AATTAGAGCA GTAGTTTATT	8220
	ATAATAAAGT AGTTGCATCT AATATGAGTA ATGCTGTTAC TATTTTGCCA GATGACATTC	8280
50	CACCAACAAT CAATAATCCT GTTGAATAA ATGCCAAATA CTATCGAGGC GACGAAKCAA	8340
	CTTTACAATG GGTGTCTCTG ATAGACATTC TGGTATAAAA AATACAATA TTACGACATT	8400

55

EP 0 786 519 A2

	TACAGGTAGA GTGAGTATGA ATCAGGCATT TAACAGTGAT ATTACATTTA AAGTGTGAGC	8520
	GACAGaCAAT GTCAATAATA CGACAAATGA TAGTCAATCT AAACATGTTT CAATTCATGT	8580
5	AGGTAAAT AGTGAAGATG CTCATCCGAT TGTATTAGGA AATACTGAGA AAGTTGTAGT	8640
	AGTCAATCCG ACTGCTGTAT CTAATGATGA AAAGCAAAGC ATAATTACTG CCTTTATGAA	8700
	TAAAAACCAA AATATAAGAG GATATTTAGC ATCAACTGAT CCAGTAACTG TCGATAATAA	8760
10	TGGTAATGTC ACATTACATT ACCGTGATGG CTCATCGACA ACGCTTGATG CTACAAATGT	8820
	GATGACATAC GAACCAAGTG TGAAACCTGA ATACCAAAGT GTCAATGCTG CTAAAACAGC	8880
	AACGGTAACG ATTGCTAAAG GACAATCATT TAGTATTGGT GATATTAAAC AATATTTTAC	8940
15	TTTAAGTAAT GGACAACCTA TTCCAAGTGG CACATTTACA AATATTACAT CTGATAGAAC	9000
	TATTCCAACT GCACAAGAAG TTAGTCAAAT GAACGCAGGC ACGCAGTTAT ACCATATAAC	9060
20	TGCTACAAAT GCGTATCATA AAGATAGTGA AGACTTCTAT ATTAGTTTGA AAATCATCGA	9120
	TGTGAAACAA CCAGAAGGCG ATCAACGTGT ATATCGTACA TCAACATATG ATTTAACTAC	9180
	TGATGAAATC TCAAAAGTAA AACCAAGCATT TATTAATGCA AATAGAGATG TAATTACGCT	9240
25	TGCCGAAGGT GATATTCAG TTCAAATAC ACCTAATGGT GCTAATGTAA GTACTATTAC	9300
	AGTAAATATT AATAAAGGTC GATTAACGAA ATCATTGCGG TCAAACCTAG CTAATATGAA	9360
	TTTCTTGGCT TGGGTTAATT TCCCACAAGA TTATACAGTG ACATGGACGA ATGCAAAAAT	9420
30	TGCAACAGA CCAACAGATG GTGGTTTATC ATGGTCTGAT GACCATAAAT CTTTAATTTA	9480
	TCGTTATGAT GCTACATTAG GTACTCAAAT TACGACGAAT GATATTTTAA CAATGTTAAA	9540
	AGCAACAACT ACAGTGCCTG GATTGCGAAA TAACATTACT GGTAATGAAA AATCACAAGC	9600
35	AGAAGCTGGC GGAAGACCTA ACTTTAGAAC GACTGGTTAT TCACAATCAA ATGCGACAAC	9660
	TGATGGTCAA CGTCAATTTA CGTTGAATGG TCAAGTGATT CAAGTGTTAG ACATCATCAA	9720
	CCCTTCAAAC GGTATATGGT GGCAACCTGT TACAAATTCA AATACTCGTG CAAACCATAG	9780
40	TAACCAACT GTTGTAAACG TAAACGAACC GGCAGCTAAT GGTGcTGGCG CATTTACAAT	9840
	TGACCACGTT GTAAAAAGTA ATTCTACACA TAATGCAAGT GATGCAGTTT ATAAAGCACA	9900
45	GTTATACTTA ACGCCATATG GTCCAAAACA ATATGTTGAA CATTTAAATC AAAATACAGG	9960
	AAATACTACT GACGCTATTA ACATTTATTT TGTACCAAGT GACTTAGTGA ATCCAACAAT	10020
	TTCAAGTAGT AATTACACTA ATCATCAAGT GTTCTCAGGT GAAACATTTA CAAATACTAT	10080
50	TACAGCGAAT GATAACTTTG GTGTGCAATC TGTAAGTGT CCAAATACAT CACAAATTAC	10140
	AGGTACTGTT GATAATAACC ATCAACATGT TTCTGCAACG GCACCAAATG TGACATCAGC	10200

55

	GTTCAATGTA ACAGTGAAAC CTTTGCCTGA TAAATATCGA GTTGGTACTT CATCAACGGC	10320
	TGCTAATCCT GTGAGAAATTG CCAATATTTT GAATAATGCG ACAGTATCAC AAGCTGATCA	10380
5	AACGACAATT ATTAATTCGT TAACGTTTAC TGAAACAGTA CCAAATAGAA GTTATGCAAG	10440
	AGCAAGTGGC AATGAAATCA CTAGTAAAC AGTTAGTAAT GTCAGTCGTA CTGGAAATAA	10500
	TGCCAATGTg CACAGTAACT GTTACTTATC AAGATGGAAC AACATCAACA GTGACTGTAC	10560
10	CTGTAAAGCA TGTCATTCCA GAAATCGTTG CACATTGCGA TTACACTGTA CAAGGCCAAG	10620
	ACTTCCCAGC AGGTAATGGT TCTAGTGCAT CAGATTACTT TAAGTTATCT AATGGTAGTG	10680
	ACATTGCAGA TGCAACTATT ACATGGGTAA GTGGACAAGC GCCAAATAAA GATAATACAC	10740
15	GTATTGGTGA AGATATAACT GTAAGTGCAC ATATCTTAAT TGATGGCGAA ACAACGCCGA	10800
	TTACGAAAAC AGCAACATAT AAAGTAGTAA GAACTGTACC GAAACATGTC TTTGAAACAG	10860
20	CCAGAGGTGT TTTATACCCA GGTGTTTCAG ATATGTATGA TCGGAAACAA TATGTTAAGC	10920
	CAGTAAATAA TTCTTGGTCG ACAAATGCGC AACATATGAA TTTCCAATTT GTTGGAACAT	10980
	ATGGTCCTAA CAAAGATGTT GTAGGCATAT CTACTCGTCT TATTAGAGTG ACATATGATA	11040
25	ATAGACAAAC AGAAGATTTA ACTATTTTAT CTAAAGTTAA ACCTGACCCA CCTAGAATTG	11100
	ACGCAAACCTC TGTGACATAT AAAGCAGGTC TTACAAACCA AGAAATTAAA GTTAATAACG	11160
	TATTAAATAA CTCGTCAGTA AAATTATTTA AAGCAGATAA TACACCATTA AATGTCACAA	11220
30	ATATTACTCA TGGTAGCGGT TTTAGTTCGG TTGTGACAGT AAGTGACGCG TTACCAAATG	11280
	GCGGAATTAA AGCAAAATCT TCAATTTCAA TGAACAATGT GACGTATACG ACGCAAGACG	11340
	AACATGGTCA AGTTGTTACA GTAACAAGAA ATGAATCTGT TGATTCAAAT GACAGTGCAa	11400
35	CAGTAACAGT GACACCACAA TTACAAGCAA CTACTGAAGG CGCTGTATTT ATTAAAGGTG	11460
	GCGACGGTTT TGATTTCCGA CACGTAGAAA GATTTATTCA AAACCCGCCA CATGGGGCAA	11520
	CGGTTGCATG GCATGATAGT CCAGATACAT GGAAGAATAC AGTCGGTAAC ACTCATAAAA	11580
40	CTGCGGTTGT AACATTACCT AATGGTCAAG GTACGCGTAA TGTGAAGTT CCAGTCAAAG	11640
	TTTATCCAST TGCTAATGCA AAGGCGCCAT CACGTGATGT GAAAGGTCAA AATTTGACTA	11700
	ATGGAACGGA TCGCATGAAC TACATTACAT TTGATCCAAA TACAAACACA AATGGTATCA	11760
45	CTGCAGCATG GGCAATAGA CAACAACCAA ATAACCAACA AGCAGGCGTG CAACATTTAA	11820
	ATGTCGATGT CACATATCCA GGTATTTTCA CTGCTAAACG AGTTCCTGTT ACTGTTAATG	11880
50	TATATCAATT TGAATTCCTT CAACTACTT ATACGACAAC GGTTGGAGGC ACTTTAGCAA	11940
	GTGGTACGCA AGCATCAGGA TATGCACATA TGCAAAATGC TACTGGTTTA CCAACAGATG	12000

55

TGAATAAACC GAATGTGGCT AAAGTCGTTA ACGCAAATA TGACGTCATC TATAACGGAC 12120
 ATACTTTTGC AACATCTTTA CCAGCGAAAT TTGTAGTAAA AGATGTGCAA CCAGCGAAAC 12180
 5 CAACTGTGAC TGAAACAGCG GCAGGAGCGA TTACAATTGC ACCTGGAGCA AACCAAACAG 12240
 TGAATACACA TGCCGGTAAC GTAACGACAT ACGCTGATAA ATTAGTTATT AAACGTAATG 12300
 GTACGTTGT GACGACATTT ACACGTCGCA ATAATACGAG TCCATGGGTG AAAGAAGCAT 12360
 10 CTGCAGCAAC TGTAGCAGGT ATTGCTGGAA CTAATAATGG TATTACTGTT GCAGCAGGTA 12420
 CTTTCAACCC TGCTGATACA ATTCAAGTTG TTGCAACGCA AGGAAGCGGA GAGACAGTGA 12480
 GTGATGAGCA ACGTAGTGAT GATTTACAG TTGTCGCACC ACAACCGAAC CAAGCGACTA 12540
 15 CTAAGATTG GCAAAATGGT CATATTGATA TCACGCCTAA TAATCCATCA GGACATTTAA 12600
 TTAATCCAAC TCAAGCAATG GATATTGCTT ACGTGAATA AGTGGGTAAT GGTGCAGAAC 12660
 20 ATAGTAAGAC AATTAATGTT GTTCGTGGTC AAAATAATCA ATGGACAATT GCGAATAAGC 12720
 CTGACTATGT AACGTTAGAT GCACAACTG GTAAAGTGAC GTTCAATGCC AATACTATAA 12780
 AACCAAATTC ATCAATCACA ATTACTCCGA AAGCAGGTAC AGGTCACTCA GTAAGTAGTA 12840
 25 ATCCAAGTAC ATTAAGTACA CCGGCAGCTC ATACTGTCAA CACAAGTAA ATTGTGAAAG 12900
 ATTATGGTTC AAATGTAACA GCAGCTGAAA TTAACAATGC AGTTCAAGTT GCTAATAAAC 12960
 GTACTGCAAC GATTAAAAAT GGCACAGCAA TGCCTACTAA TTTAGCTGGT GGTAGCACAA 13020
 30 CGACGATTCC TGTGACAGTA ACTTACAATG ATGGTAGTAC TGAAGAAGTA CAAGAGTCCA 13080
 TTTTCACAAA AGCGGATAAA CGTGAGTTAA TCACAGCTAA AAATCATTTA GATGATCCAG 13140
 TAAGCACTGA AGGTAAAAAG CCAGGTACAA TTACGCAGTA CAATAATGCA ATGCATAATG 13200
 35 CGCAACAACA AATCAATACT GCGAAAACAG AAGCACAACA AGTGATTAAT AATGAGCGTG 13260
 CAACACCACA ACAAGTTTCT GACGCACTAA CTAAAGTTCTG TGCAGCACAA ACTAAGATTG 13320
 ATCAAGCTAA AGCATTACTT CAAAATAAAG AAGATAATAG CCAATTAGTA ACGTCTAAAA 13380
 40 ATAACTTACA AAGTTCTGTG AACCAAGTAC CATCAACTGC TGGTATGACG CAACAAAGTA 13440
 TTGATAACTA TAATGCGAAG AAGCGTGAAG CAGAACTGA AATAACTGCA GCTCAACGTG 13500
 45 TTATTGACAA TGGCGATGCA ACTGCACAAC AAATTTTCTG TGAAAAACAT CGTGTGATA 13560
 ACGCATTAA AGCATTAAAC CAAGCGAAAC ATGATTTAAC TGCAGATACA CATGCCTTAG 13620
 AGCAAGCAGT GCAACAATTG AATCGCACAG GTACAACGAC TGGTAAGAAG CCGGCAAGTA 13680
 50 TTAAGTCTTA CAATAATTCG ATTCTGTCAC TTCAAAGTGA CTTAACAAGT GCTAAAAATA 13740
 GCGCTAATGC TATTATTCAA AAGCCAATAA GAACAGTACA AGAAGTGCAA TCTGCGTTAA 13800

55

EP 0 786 519 A2

	CTGATAATAG TGCTTTAAAA ACTGCTAAGA CGAAACTTGA TGAAGAAATC AATAAATCAG	13920
	TAACACTACTGA TGGTATGACA CAATCATCAA TCCAAGCATA TGAAAATGCT AAACGTGCGG	13980
5	GTCAAACAGA ATCAACAAAT GCACAAAATG TTATTAACAA TGGTGATGCG ACTGACCAAC	14040
	AAATTGCCGC AGAAAAACA AAAGTAGAAG AAAAATATAA TAGCTTAAAA CAAGCAATTG	14100
	CTGGATTAACT TCCAGACTTG GCACCATTAC AAAGTGCATA AACTCAGTTG CAAAATGATA	14160
10	TTGATCAGCC AACGAGTACG ACTGGTATGA CAAGCGCATC TATTGCAGCA TTTAATGAAA	14220
	AACTTTCAGC AGCTAGAAGT AAAATTCAAG AAATTGATCG TGTATTAGCC TCACATCCAG	14280
	ATGTTGCGAC AATACGTCAA AACGTGACAG CAGCGAATGC CGCTAAATCA GCACTTGATC	14340
15	AAGCACGTAA TGGCTTAACA GTCGATAAAG CGCCTTTAGA AAATGCGAAA AATCAACTAC	14400
	AACATAGTAT TGACACGCAA ACAAGTACAA CTGGTATGAC ACAAGACTCT ATAAATGCAT	14460
20	ACAATGCGAA GTTAACAGCT GCACGTAATA AGATTCAACA AATCAATCAA GTATTAGCAG	14520
	GTTCAACCGAC TGTAAGACAA ATTAATACAA ATACGTCTAC AGCAAATCAA GCTAAATCTG	14580
	ATTTAGATCA TGCACGTCAA GCTTTAACAC CAGATAAAGC GCCGCTTCAA ACTGCGAAAA	14640
25	CGCAATTAGA ACAAAGCATT AATCAACCAA CGGATACAAC AGGTATGACG ACCGCTTCGT	14700
	TAAATGCGTA CAACCAAAAA TTACAAGCAG CGCGTCAAAA GTTAACTGAA ATTAATCAAG	14760
	TGTTGAATGG CAACCCAACT GTCCAAAATA TCAATGATAA AGTGACAGAG GCAAACCAAG	14820
30	CTAAGGATCA ATTAAATACA GCACGTCAAG GTTTAACATT AGATAGACAG CCAGCGTTAA	14880
	CAACATTACA TGGTGCATCT AACTTAAACC AAGCACAACA AAATAATTTT ACACCAACAA	14940
	TTAATGCTGC TCAAAATCAT GctGCGCTTG AAACAATTAA GTCTAACATT ACGGCTTTAA	15000
35	ATACTGCGAT GACGAAATTA AAAGACAGTG TTGCGGATAA TAATACAATT AAATCAGATC	15060
	AAAATTACAC TGACGCAACA CCAGCTAATA AACAAGCGTA TGATAATGCA GTTAATGCGG	15120
40	CTAAAGGTGT CATTGGAGAA ACGACTAATC CAACGATGGA TGTTAACACA GTGAACCAAA	15180
	AAGCAGCATC TGTTAAATCG ACGAAAGATG CTTTAGATGG TCAACAAAAC TTACAACGTG	15240
	CGAAAACAGA AGCAACAAAT GCGATTACGC ATGCAAGTGA TTTAAACCAA GCACAAAAGA	15300
45	ATGCATTAAC ACAACAAGTG AATAGTGcAC AAAACGTGCA AGCAGTAAAT GATATTAAAC	15360
	AAACGACTCA AAGCTTAAAT ACTGCTATGA CAGGTTTAAA ACGTGGCGTT GCTAATCATA	15420
	ACCAAGTCGT ACAAAGTGAT AATTATGTCA ACGCAGATAC TAATAAGAAA AATGATTACA	15480
50	ACAATGCATA CAACCATGCG AATGACATTA TTAATGGTAA TGCACAACAT CCAGTTATAA	15540
	CACCAAGTGA TGTTAACAAT GCTTTATCAA ATGTCACAAG TAAAGAACAT GCATTGAATG	15600

55

ATTTAAATAA TGCACAACGT CAAAACCTTAC AATCGCAAAT TAATGGTGCG CATCAAATTG 15720
 ATGCAGTTAA TACAATTAAG CAAAATGCAA CAAACTTGAA TAGTGCAATG GGTAACCTAA 15780
 5 GACAAGCTGT TGCAGATAAA GATCAAGTGA AACGTACAGA AGATTATGCG GATGCAGATA 15840
 CAGCTAAACA AAATGCATAT AACAGTGCAG TTTCAAGTGC CGAAACAATC ATTAATCAAA 15900
 CAACAAATCC AACGATGTCT GTTGATGATG TTAATCGTGC AACTTCAGCT GTTACTTCTA 15960
 10 ATAAAAATGC ATTAAATGGT TATGAAAAAT TAGCACAATC TAAAACAGAT GCTGCAAGAG 16020
 CAATTGATGC ATTACCACAT TTAAATAATG CACAAAAAGC AGATGTTAAA TCTAAAATTA 16080
 ATGCTGCATC AAATATTGCT GCGGTAAATA CTGTTAAACA ACAAGGTACA GATTTAAATA 16140
 15 CAKCGATGGG TAACCTGCAA GGTGCAATCA ATGATGAACA AACGACGCTT AATAGTCAAA 16200
 ACTATCAAGA TCGCACACCT AGTAAGAAAA CAGCATACAC AAATGCGGTA CAAGCTGCGA 16260
 AAGATATTTT AAATAAATCA AATGGTCAAA ATAAAACGAA AGATCAAGTT ACTGAAGCGA 16320
 20 TGAATCAAGT GAATTCTGCT AAAAATAACT TAGATGGTAC GCGTTTATTA GATCAAGCGA 16380
 nCAAaCAGCA AAACAGCAGT TAAATAATAT GACGCATTTA ACAACTGCAC AAAAAACGAA 16440
 25 TTTAACAAAC CAAATTAATA GTGGTACTAC TGTCGCTGGT GTTCAAACGG TTCAATCAAA 16500
 TGCCAATACA TTAGATCAAG CCATGAATAC GTTAAGACAA AGTATTGCCA ACAAGATGC 16560
 GACTAAAGCA AGTGAAGATT ACGTAGATGC TAATAATGAT AAGCAAACAG CATATAACAA 16620
 30 CGCAGTAGCT GCTGCTGAAA CGATTATTAA TGCTAATAGT AATCCAGAAA TGAATCCAAG 16680
 TACGATTACA CAAAAAGCAG AGCAAGTGAA TAGTTCTAAA ACGGCACTTA ACGGTGATGA 16740
 AAACCTAGCT GCTGCAAAAC AAAATGCGAA AACGTACTTA AACACATTGA CAAGTATTAC 16800
 35 AGATGCTCAA AAGAACAATT TGATTAGTCA AATTACTAGT GCGACAAGAG TGAGTGGTGT 16860
 TGATACTGTA AAACAAAATG CGCAACATCT AGACCAAGCT ATGGCTAGCT TACAGAATGG 16920
 TATTAACAAC GAATCTCAAG TGAAATCATC TGAGAAATAT CGTGATGCTG ATACAAATAA 16980
 40 ACAACAAGAG TATGATAATG CTATTACTGC AGCGAAAGCG ATTTTAAATA AATCGACAGG 17040
 TCCAAACACT GCGCAAAATG CAGTTGAAGC AGCATTACAA CGTGTTAATA ATGCGAAAGA 17100
 TGCATTGAAT GGTGATGCAA AATTAATTGC AGCTCAAAAC GCAGCGAAAC AACATTTAGG 17160
 45 TACTTTAACG CATATCACTA CAGCTCAACG TAATGATTTA ACAAATCAAA TTTCACAAGC 17220
 TACAACTTA GCTGGTGTG AATCTGTTAA ACAAATGCG AATAGTTTAG ATGGTGCTAT 17280
 50 GGGTAACTTA CAAACGGCTA TCAACGATAA GTCAGGAACA TTAGCGAGCC AAAACTTCTT 17340
 GGATGCTGAT GAGCAAAAAC GTAATGCATA CAATCAAGCT STATCAGCAG CCGAAACCAT 17400

55

EP 0 786 519 A2

	TGTTAATAAT GCGAAACATG CATTAAATGG TACGCAAAAC TTAAACAATG CGAAACAAGC	17520
	AGCGATTACA GCAATCAATG GCGCATCTGA TTAAATCAA AAACAAAAAG ATGCATTAAA	17580
5	AGCACAAAGCT AATGGTGCTC AACGCGTATC TAATGCACAA GATGTACAGC ACAATGCGAC	17640
	TGAACTGAAC ACGGCAATGG GCACATTAAA ACATGCCATC GCAGATAAGA CGAATACGTT	17700
	AGCAAGCAGT AAATATGTTA ATGCCGATAG CACTAAACAA AATGCTTACA CAACTAAAGT	17760
10	TACCAATGCT GAACATATTA TTAGCGGTAC GCCAACGGTT GTTACGACAC CTTCAGAAGT	17820
	AACAGCTGCA GCTAATCAAG TAAACAGCGC GAAACAAGAA TTAAATGGTG ACGAAAGATT	17880
	ACGTGAAGCA AAACAAAACG CCAATACTGC TATTGATGCA TTAACACAAT TAAATACACC	17940
15	TCAAAAAGCT AAATTTAAAG AACAAAGTGGG ACAAGCCAAT AGATTAGAAG ACGTACAAAC	18000
	TGTTCAAACA AATGGACAAG CATTGAACAA TGCAATGAAA GGCTTAAGAG ATAGTATTGC	18060
20	TAACGAAACA ACAGTCAAAA CAAGTCAAAA CTATACAGAC GCAAGTCCGA ATAACCAATC	18120
	AACATATAAT AGCGCTGTGT CAAATGCGAA AGGTATCATT AATCAAACTA ACAATCCGAC	18180
	TATGGATACT AGTGCGATTA CCCAAGCTAC AACACAAGTG AATAATGCTA AAAATGGTTT	18240
25	AAACGGTGCT GAAAACCTAA GAAATGCACA AAACACTGCT AAGCAAACT TAAATACATT	18300
	ATCACACTTA ACAAATAACC AAAAATCTGC CATCTCATCA CAAATTGATC GTGCAGGTCA	18360
	TGTGAGTGAG GTAAGTGCTA CTAAAAATGC AGCAACTGAG TTGAATACGC AAATGGGTAA	18420
30	CTTGGAACAA GCTATCCATG ATCAAAACAC AGTTAAACAA AGTGTTAAAT TTAAGTATGC	18480
	AGATAAAGCT AAACGTGATG CGTATACAAA TGCGGTAAGC AGAGCTGAAG CAATTCTGAA	18540
	TAAACGCAA GGTGCAAATA CGTCTAAACA AGATGTTGAA GCGGCTATT CAAAATGTTTC	18600
35	AAGTGCTAAA AATGCATTGA ATGGTGATCA AAACGTTACA AATGCGAAGA ATGCAGCTAA	18660
	AAATGCATTA AATAACTTAA CGTCAATTAA TAATGCACAA AAACGTGACT TAACAACTAA	18720
40	AATTGATCAA GCAACAACCTG TAGCTGGTGT TGAAGCTGTA TCTAATACGA GTACACAATT	18780
	GAALACAGCG ATGGCTAACT TGCAAAATGG TATTAATGAT AAAACAAATA CACTAGCAAG	18840
	TGAAAACTAT CATGATGCTG ATTCAGATAA GAAAACCTGCT TATACTCAAG CCGTTACGAA	18900
45	CGCAGAAAAT ATTTTAAATA AAAATAGTGG ATCAAATTTA GACAAAACCTG CCGTTGAAAA	18960
	CGCGTTGTCA CAAGTTGCTA ATGCGAAAGG TGCCCTAAAT GGTAACCATA ATTTAGAGCA	19020
	AGCTAAATCA AATGCAAACA CTACTATAAA CGGACTTCAA CATTTAACAA CTGCTCAAAA	19080
50	AGATAAATTG AAACAACAAG TGCAACAAGC ACAAATGTT GCAGGTGTAG ATACTGTTAA	19140
	ATCAAGTGCC AACACATTAA ATGGTGCTAT GGGTACGTTA AGAAATAGCA TACAAGATAA	19200

55

	TAACAATGCT GTTGATAGTG CTAATGGTGT CATTAAATGCA ACAAGCAATC CAAATATGGA	19320
	TGCTAATGCA ATTAACCAAA TCGCTACACA AGTGACATCA ACGAAAAATG CATTAGATGG	19380
5	TACACATAAT TTAACGCAAG CGAAACAAAC AGCAACAAAT GCCATCGATG GTGCTACTAA	19440
	CTTAAATAAA GCGCAAAAAG ATGCGTTAAA AGCACAAGTT ACAAGTGC GC AACGTGTTGC	19500
	AAATGTAACA AGTATCCAAC AAAGTCAAA TGAAGTTAAT ACAGCTATGG GTCAATTACA	19560
10	ACATGGTATT GATGATGAAA ATGCAACAAA ACAAACTCAA AAATATCGTG ACGCTGAACA	19620
	AAGTAAGAAA ACTGCTTATG ATCAAGCTGT AGCTGCTGCG AAAGCAATTT TAAATAAACA	19680
	AACAGGTTCA AATTCAGATA AAGCAGCAGT TGACCGTGCA TTACAACAAG TAACAAGTAC	19740
15	GAAAGATGCA TTGAATGGTG ATGCAAAACT GGCAGAAGCG AAAGCGGCAG CTAAACAAAA	19800
	CTTAGGCACT TTAAACCATA TTACGAATGC ACAACGTACT GACTTAGAAG GCCAAATCAA	19860
20	TCAAGCGACG ACTGTTGATG GCGTTAATAC TGTA AAAACA AATGCCAATA CATTAGACGG	19920
	CGCAATGAAT AGCTTACAAG GTTCAATCAA TGATAAAGAT GCGACATTAA GAAATCAAAA	19980
	TTATCTTGAT GCGGATGAAT CAAAACGAAA TGCATATACG CAAGCTGTCA CAGCGGCTGA	20040
25	AGGCATTTTA AATAACAAA CTGGTGGTAA CACATCTAAA GCAGACGTTG ATAATGCATT	20100
	AAATGCAGTT ACAAGAGCGA AAGCGCTTT AAATGGTGCT GACAACTTAA GAAATGCGAA	20160
	AACTTCAGCA ACAATACGA TTGATGGTTT ACCTAACTTA ACACAATTAC AAAAAGACAA	20220
30	CTTGAAGCAT CAAGTTGAAC AAGCGCAAAA TGTAGCAGGT GTAAATGGTG TTAAAGATAA	20280
	AGGTAATACG TTAAATACTG CCATGGGTGC ATTACGTACA AGTATCCAAA ATGATAATAC	20340
	GACGAAAACA AGTCAAAATT ATCTTGATGC ATCTGACAGC AACAAAAATA ATTACAATAC	20400
35	TGCTGTAAAT AATGCAAATG GTGTTATTAA TGCAACGAAC AATCCAAATA TGGATGCTAA	20460
	TGCGATTAAAT GGCATGGCAA ATCAAGTCAA TACAACAAA GCAGCGTTAA ATGGTGCACA	20520
	AACTTAGCT CAAGCTAAAA CAAATGCGAC GAACACAATT AACACGCAC ATGACTTAAA	20580
40	CCAAAACAA AAAGATGCAT TAAAAACACA AGTTAACAAT GCACAACGTG TATCTGATGC	20640
	AAATAACGTT CAACACACTG CAACTGAATT GAACAGTGCG ATGACAGCAC TTAAAGCAGC	20700
45	TATTGCTGAT AAAGAAAGAA CAAAGCAAG CGGTAATTAT GTCAATGCTG ATCAAGAAAA	20760
	ACGTCAAGCG TATGATTCAA AAGTGAATA CGCTGAAAAT ATCATTAGTG GTACACCGAA	20820
	TGCGACATTA ACAGTCAATG ACGTAAATAG TGCGGCATCA CAAGTCAATG CGGCTAAAAC	20880
50	AGCATTAAAT GGTGATAACA ACTTACGTGT AGCGAAAGAG CATGCCAACA ATACAATTGA	20940
	CGGCTTAGCA CAATTGAATA ATGCACAAA AGCAAAATTA AAAGAACAAG TTCAAAGTGC	21000

55

	GAAAGGCTTA AGAGATAGTA TTGCGAATGA AGCAACAATT AAAGCAGGTC AAAACTACAC	21120
	TGACGCAAGT CCAAATAATC GTAACGAGTA CGACAGTGCA GTTACTGCAG CAAAAGCAAT	21180
5	CATTAATCAA ACATCGAACC CAACGATGGA ACCAAATACT ATTACGCAAG TAACATCACA	21240
	AGTGACAACT AAAGAACAGG CATTAAATGG TGC CGGAAAC TTAGCTCAAG CTAAGACAAC	21300
	TGCGAAAAAC AACTTGAATA ACTTAACATC AATTAACAAT GCACAAAAAG ATGCGTTAAC	21360
10	GCGTAgcATT GATGGTGCAA CAACAGTAGC TGGTGTAAT CAAGAACTG CAAAAGCAAC	21420
	AGAATTAAAT AACGCAATGC ATAGTTTACA AAATGGTATC AATGATGAGA CACAAACAAA	21480
	ACAAACTCAG AAATACCTAG ATGCAGAGCC AAGTAAGAAA TCAGCTTATG ATCAAGCAGT	21540
15	AAATGCAGCG AAAGCAATTT TAACAAAAGC TAGTGGTCAA AATGTAGACA AAGCAGCAGT	21600
	TGAACAAGCA TTGCAAAATG TGAACAGTAC GAAGACGGCG TTGAACGGTG ATGCGAAATT	21660
20	AAATGAAGCT AAAGCAGCTG CGAAACAAAC GTTAGGTACA TTAACACACA TTAATAATGC	21720
	ACAACGTACA GCGTTAGACA ATGAAATTAC ACAAGCAACA AATGTTGAAG GTGTTAATAC	21780
	AGTTAAAGCC AAAGCGCAAC AATTAGATGG TGCTATGGGT CAATTAGAAA CATCAATTCG	21840
25	TGATAAGAC ACGACGTTAC AAAGTCAAAA TTATCAAGAT GCTGATGATG CTAAACGAAC	21900
	TGCTTATTCT CAAGCAGTAA ATGCAGCAGC AACTATTTTA AATAAAACag CTGGCGGTAA	21960
	TACACCTAAA GCAGATGTTG AAAGAGCAAT GCAAGCTGTT ACACAAGCAA ATACTGcATT	22020
30	AAACGGTATT CamAACTTAG ATCGTGCGAA ACArGCTGCT AACACAGCGA TTACAAATGC	22080
	TTCGGACTTA AATACAAAAC mAAAAGAAGC ATTAaAAgCA CAAGTAACAA GTGCAGGACG	22140
	TGTATCTGCA GCAAATGGTG TTGAACATAC TGCGACTGAA TTAAATACTG CGATGACAGC	22200
35	TTTAAAGCGT GCCATTGCTG ATAAAGCTGA GACAAAAGCT AGTGGTAACT ATGTCAATGC	22260
	TGATGCGAAT AAACGTCAAG CATATGATGA AAAAGTTACA GCTGCCGAAA ATATCGTTAG	22320
	TGGTACACCA ACACCAACGT TAACACCAGC AGATGTTACA AATGCAGCAA CGCAAGTAAC	22380
40	GAATGCTAAG ACGCAGTTAA ACGGTAATCA TAATTTAGAA GTAGCGAAAC AAAATGCTAA	22440
	CACTGCAATT GATGGTTTAA CTTCTTTAAA TGGTCCGCAA AAAGCAAAAC TTAAAGAACA	22500
45	AGTGGGTCAA GCGACGACGT TGCCAAATGT TCAAAGTGT CGTGATAATG CACAAACATT	22560
	AAACACTGCA ATGAAAGGTC TACGAGATAG CATTGCGAAT GAAGCAACGA TTAAAGCAGG	22620
	TCAAAACTAC ACAGATGCAA GTCAAAACAA ACAAACTGAC TACAACAGTG CAGTCACTGC	22680
50	AGCAAAAGCA ATCATTGGTC AAACAAC TAGTCCATATG AATGCGCAAG AAATTAATCA	22740
	AGCGAAAGAC CAAGTGACAG CTAAACAACA AGCGTTAAAC GGTCAAGAAA ACTTAAGAAC	22800

55

	AGATGCAGTG AAACGTCAAA TCGAAGGTGC AACGCATGTT AATGAAGTAA CACAAGCACA	22920
	AAATAATGCG GATGCATTAA ATACAGCTAT GACGAACTTG AAAAATGGTA TTCAAGATCA	22980
5	GAATACGATT AAGCAAGGTG TTAAC TTCAC TGATGCCGAC GAAGCGAAAC GTAATGCATA	23040
	TACAAATGCA GTGACGCAAG CTGAACAAAT TTTAAATAAA GCACAAGGTC CAAATACTTC	23100
	AAAAGACGGT GTCGAACTG CGTTAGAAA TGTACAACGT GCTAAAAACG AATTGAACGG	23160
10	TAATCAAAAT GTTGCGAAG CTAAGACAAC TGCGAAAAAT GCATTGAATA ACCTAACATC	23220
	AATTAATAAT GCACAAAAAG AAGCATTGAA ATCACA AATT GAAGGTGCGA CAACAGTTGC	23280
	AGGTGTAAAT CAAGTGTCTA CAACGGCATC TGAATTAAAT ACAGCAATGA GCAACTTACA	23340
15	AAATGGTATT AATGATGAAG CAGCTACAAA AGCAGCGCTT AATGGTACTC AAAACCTTGA	23400
	AAAAGCTAAA CAACACGCAA ATACAGCAAT TGACGGTTTA AGCCATT TAA CAAATGCACA	23460
20	AAAAGAGGCA TTAAACAAT TGGTACAACA ATCGACTACT GTTGCGAAG CACAAGGTAA	23520
	TGAGCAAAAA GCAAACAATG TTGATGCAGC AATGGACAAA TTACGTCAA GTATTGCAGA	23580
	TAATGCGACA ACAAACAAA ACCAAAATTA TACTGATGCA AGTCAGAATA AAAAGGATGC	23640
25	GTACAATAAT GCTGTCACAA CTGCACAAGG TATTATTGAT CAAACTACAA GTCCAAC TTT	23700
	AGATCCGACT GTTATCAATC AAGCTGCTGG ACAAGTAAGC ACAACTAAAA ATGCATTAAA	23760
	TGGTAATGAA AACCTAGAGG CAGCGAAACA ACAAGCGTCA CAATCATTAG GTTCATTAGA	23820
30	TAACTTAAAT AATGCGCAAA AACAAACAGT TACTGATCAA ATTAATGGCG CGCATACTGT	23880
	TGATGAAGCA AATCAAATTA AGCAAAATGC GCAAACTTA AATACAGCGA TGGGTAAC TT	23940
	GAAACAAGCG ATAGCTGACA AAGATGCTAC GAAAGCGACA GTTAACTTCA CTGATGCAGA	24000
35	TCAAGCAAAA CAACAAGCAT ATAACA CTGC TGTTACAAAT GCTGAAAATA TCATTTCAAA	24060
	AGCTAATGGC GGCAATGCAA CACAAGCTGA AGTTGAACAA GCAATCAAAC AAGTTAATGC	24120
	TGCAAAACAA GCATTAAATG GTAATGCCAA CGTTCAACAT GCAAAAGACG AAGCAACAGC	24180
40	ATTAATTAAT AGCTCTAATG ACCTTAACCA AGCACAAAA GACGCATTAA AACACAAGT	24240
	TCAAAATGCA ACTACTGTAG CTGGTG TAAA CAATGT TAAA CAAACAGCAC AAGAGTTAAA	24300
45	CAATGCTATG ACACAATTAA AACAGGCAT TGCAGATAAA GAACAAACAA AAGCTGATGG	24360
	TAACTTTGTC AATGCAGATC CTGATAAGCA AAATGCATAT AATCAAGCAG TAGCGAAAGC	24420
	TGAAGCATT AATTAGTGCTA CGCCTGATGT TGTCGTTACA CCTAGCGAAA TTACTGCAGC	24480
50	GTTAAATAAA GTTACGCAAG CTAAAAATGA TTTAAATGGT AATACAAACT TAGCAACGGC	24540
	GAAACAAAAT GTTCAACATG CTATTGATCA ATTGCCAAAC TTAAACCAAG CGCAACGTGA	24600

55

EP 0 786 519 A2

	AGCGGCGACA ACGCTTAATG ACGCGATGAC ACAATTGAAA CAAGGTATTG CGAATAAAGC	24720
	ACAAATTAAA GGTAGCGAGA ACTATCACGA TGCTGATACT GACAAGCAAA CAGCATATGA	24780
5	TAATGCAGTA ACAAAGCAG AAGAATTGTT AAAACAAACA ACAAATCCAA CAATGGATCC	24840
	AAATACAATT CAACAAGCAT TAACTAAAGT GAATGACACA AATCAAGCAC TTAACGGTAA	24900
	TCAAAAATTA GCTGATGCCA AACAAGATGC TAAGACAACA CTTGGTACAC TAGATCATTT	24960
10	AAATGATGCT CAAAAACAAG CGCTAACAACT TCAAGTTGAA CAAGCACCAG ATATTGCAAC	25020
	AGTTAATAAT GTTAAGCAAA ATGCTCAAAA TCTGAATAAT GCTATGACTA ACTTAAACAA	25080
	TGCATTACAA GATAAACTG AGACATTAAA TAGCATTAACT TTTACTGATG CAGATCAAGC	25140
15	TAAGAAAGAT GCTTATACTA ATGCGGTTTC ACATGCAGAA GGTATTTTAT CTAAAGCAAA	25200
	TGGCAGCAAT GCAAGTCAAA CTGAAGTGGG ACAAGCGATG CAACGTGTGA ACGAAGCGAA	25260
20	ACAAGCATTG AATGGTAATG ACAATGTACA ACGTGCAAAA GATGCAGCGA AACAAGTGAT	25320
	TACAAATGCA AATGATTAA ATCAAGCAAT GACACAATTG AAACAAGGTA TTGCAGATAA	25380
	AGACCAAAT AAAGCAAATG GTAACCTTGT CAATGCTGAT ACTGATAAGC AAAATGCTTA	25440
25	CAACAATGCG GTAGCACATG CTGAACAAAT AATTAGTGGT ACACCAAATG CAAACGTGGA	25500
	TCCACAACAA GTGGCTCAAG CGTTACAACA AGTGAATCaA GCTAAGGGTG ATTTAAACGG	25560
	TAACCATAAC TTACAAGTTG CTAAAGACAA TGCAAATACA GCCATTGATC AGTTACCAAA	25620
30	CTTAAATCAA CCACAAAAA CAGCATTAAA AGACCAAGTG TCGCATGCAG AACTTGTTAC	25680
	AGGTGTTAAT GCTATTAAGC AAAATGCTGA TCGGTTAAAT AATGCAATGG GTACATTGAA	25740
	ACAACAAATT CAAGCGAACA GTCAAGTACC ACAGTCAGTT GACTTTACAC AAGCGGATCA	25800
35	AGACAAACAA CAAGCATATA ACAATGCGGC TAACCAAGCG CAACAAATCG CAAATGGCAT	25860
	ACCAACACCT GTATTGACGC CTGATACAGT AACACAAGCA GTGACAACCTA TGAATCAAGC	25920
	GAAAGATGCA TTAACCGGTG ATGAAAAATT AGCACAAGCG AAACAAGAAG CTTTAGCAAA	25980
40	TCTTGATACG TTACGCGATT TAAATCAACC ACAACGTGAT GCATTACGTA ACCAAATCAA	26040
	TCAAGCACAA GCGTTAGCTA CAGTTGAACA AACTAAACAA AATGCACAAA ATGTGAATAC	26100
45	aGCaATGAGT AACTTGAAAC aAGGTATTGC aAACAAAGAT ACTGTCAAAG CAAGTGAGAA	26160
	CTATCATGAT GCTGATGCCG ATAAGCAAAC AGCATATACA AATGCAGTGT CTCAAGCGGA	26220
	AGGTATTATC AATCAAACGA CAAATCCAAC GCTTAACCCA GATGAAATAA CACGTGCATT	26280
50	AACTCAAGTG ACTGATGCTA AAAATGGCTT AAACGGTGAA GCTAAATTGG CAACTGAAAA	26340
	GCAAAATGCT AAAGATGCCG TAAGTGGGAT GACGCATTTA AACGATGCTC AAAAAACAAGC	26400

55

	AGCAACGAGC CTAGATCAAG CAATGGATCA ATTATCACAA GCTATTAATG ATAAAGCTCA	26520
	AACATTAGCG GACGGTAATT ACTTAAATGC AGATCCTGAC AAACAAAATG CGTATAAACA	26580
5	GGCAGTAGCA AAAGCTGAAG CATTATTGAA TAAACAAAGT GGTACTAATG AAGTACAAGC	26640
	ACAAGTTGAA AGCATCACTA ATGAAGTGAA CGCAGCGAAA CAAGCATTAA ATGGTAATGA	26700
	CAATTTGGCA AATGCAAAAC AACAAAGCAA ACAACAATTG GCGAACTTAA CACACTTAAA	26760
10	TGATGCACAA AAACAATCAT TTGAAAGTCA AATTACACAA GCGCCACTTG TTACAGATGT	26820
	CACTACGATT AATCAAAAAG CACAAACGTT AGATCATGCG ATGGAATTAT TAAGAAATAG	26880
	TGTTGCGGAT AATCAAACGA CATTAGCGTC TGAAGATTAT CATGATGCAA CTGCGCAAAG	26940
15	ACAAAATGAC TATAACCAAG CTGTAACAGC TGCTAATAAT ATAATTAATC AAACCTACATC	27000
	GCCTACGATG AATCCAGATG ATGTTAATGG TGCAACGACA CAAGTGAATA ATACGAAAGT	27060
	TGCATTAGAT GGTGATGAAA ACCTTGACGC AGCTAAACAA CAAGCAAACA ACAGACTTGA	27120
20	TCAATTAGAT CATTTGAATA ATGCGCAAAA GCAACAGTTA CAATCACAAA TTACGCAATC	27180
	ATCTGATATT GCTGCAGTTA ATGGTCACAA ACAACAGCA GAATCTTTAA ATACTGCGAT	27240
25	GGGTAAGTTA ATTAATGCCA TTGCAGATCA TCAAGCCGTT GAACAACGTG GTAAGTTTCA	27300
	CAATGCTGAT ACTGATAAAC AAAGTCTTAA TAATACAGCG GTAAATGAAG CAGCAGCAAT	27360
	GATTAACAAA CAACTGGTC AAAATGCGAA CCAACAGAA GTAGAACAAG CTATTACTAA	27420
30	AGTTCAAACA ACACTTCAAG CGTTAAATGG AGACCATAAT TTACAAGTTG CTAAAACAAA	27480
	TGCGACGCAA GCAATTGATG CTTTAACAAG CTTAAATGAT CCTCAAAAAA CAGCATTAAA	27540
	AGACCAAGTT ACAGCTGCAA CTTTAGTAAC TGCAAGTTAT CAAATTGAAC AAAATGCGAA	27600
35	TACGCTTAAC CAAGCAATGC ATGGTTTAAG ACAGAGCATT CAAGATAACG CAGCAACTAA	27660
	AGCAATATAGC AAATATATCA ACGAAGATCA ACCAGAGCAA CAAACTATG ATCAAGCTGT	27720
	TCAAGCCGCA AATAATATTA TCAATGAACA AACTGCAACA TTAGATAATA ATGCGATTAA	27780
40	TCAAGCAGCG ACAACTGTGA ATACAACGAA AGCAGCATT CATGGTGATG TGAAGTTACA	27840
	AAATGATAAA GATCATGCTA AGCAAACGGT TAGTCAATTA GCACATCTAA ACAATGCACA	27900
45	AAAACATATG GAAGATACGT TAATTGATAG TGAAACAAC AGAACAGCAG TTAAGCAAGA	27960
	TTTGAAGTGA GCACAAGCAT TAGATCAACT TATGGATGCA TTACAACAAA GTATTGCTGA	28020
	CAAAGATGCA ACACGTGCGA GCAAGTGCATA TGTCAATGCA GAACCGAATA AAAACAATC	28080
50	CTATGATGAA GCAGTTCAAA ATGCTGAGTC TATCATTGCA GGATTAAATA ATCCAATAT	28140
	CAATAAAGGT AATGTATCAA GTGCGACTCA AGCAGTAATA TCATCTAAAA ATGCATTAGA	28200

55

	TCAATTAACA CCAGCTCAAC AACAAAGCGCT AGAAAATCAA ATTAATAATG CAACAACTCG	28320
	TGATAAAGTG GCTGAAATCA TTGCACAAGC GCAAGCATTa AATGAAGCGA TGAAAGCATT	28380
5	AAAAGAAAGT ATTAAGGATC AACCACAAAC TGAAGCAAGT AGTAAATTTA TTAACGAGGA	28440
	TCAAGCGCAA AAAGATGCTT ATACGCAAGC AGTACAACAC GCGAAAGATT TGATTAACAA	28500
	AACAACTGAT CCTACATTAG CTAAATCAAT CATTGATCAA GCGACACAGG CAGTGACAGA	28560
10	TGCTAAAAAC AATTTACATG GTGATCAAAA ACTAGCTCAA GATAAGCAAC GTGCAACAGA	28620
	AACGTTAAAT AACTTGTCTA ACTTGAATAC ACCACAACGT CAAGCACTTG AAAATCAAAT	28680
	TAATAATGCA GCAACTCGTG GCGAAGTAGC ACAAAAATTA ACTGAAGCAC AAGCACTTAA	28740
15	CCAAGCAATG GAAGCTTTAC GTAATAGCAT TCAAGATCAA CAGCAAACGG AAGCGGGTAG	28800
	CAAGTTTATC AATGAAGATA AaCCaCmAAA AGrTGCTTAC CAAGCAGCAG TTCAAAATGC	28860
20	AAAAGATTTA ATTAATCAAA CTAACAATCC AACGCTTGAT AAAGCACAAG TTGAACAATT	28920
	GACACAAGCT GTTAACCAAG CTAAAGATAA CCTACACGGT GATCAAAAAC TTGCAGACGA	28980
	TAAACAACAT GCGGTTACTG ATTTAAATCA ATTAAATGGT TTGAATAATC CGCAACGTCA	29040
25	AGCACTTGAA AGCCAAATAA ACAACGCAGC AACTCGTGGC GAAGTAGCAC AAAAATTAGC	29100
	TGAAGCAAAA GCGCTTGATC AAGCAATGCA AGCATTACGT AATAGTATTC AAGATCAACA	29160
	ACAAACAGAA TCTGGTAGCA AGTTTATCAA TGAAGATAAA CCGCAAAAAG ATGCTTACCA	29220
30	AGCAGCAGTT CAAAATGCAA AAGATTTAAT TAACCACACA GGTAATCCA CACTCGACAA	29280
	ATCACAAGTA GAACAATTGA CACAAGCAGT AACAACTGCA AAAGATAATC TACATGGTGA	29340
	TCAAAAACCTT GCTCGTGATC AACAAACAAGC AGTAACAACT GTAAATGCAT TGCCAACTT	29400
35	AAATCATGCA CAACAACAAG CATTAAGTGA TGCTATAAAT GCAGCGCCTA CAAGAACAGA	29460
	GGTTGCACAA CATGTTCAAA CTGCTACTGA ACTTGATCAC GCGATGGAAA CATTGAAAAA	29520
40	TAAAGTTGAT CAAGTGAATA CAGATAAGGC TCAACCAAAT TAACTGAAG CGTCAACTGA	29580
	TAAAAAAGAA GCAGTAGATC AAGCGTTACA AGCTGCAGAA AGCATTACAG ATCCAACTAA	29640
	TGGTTCAAAT GCGAATAAAG ACGCTGTAGA CCAAGTATTA ACTAAGCTTC AAGAAAAAGA	29700
45	AAATGAGTTA AATGGTAATG AGAGAGTCGC TGAAGCTAAA ACACAAGCGA AACAACTAT	29760
	TGACCAATTA ACACATTTAA ATGCTGATCA AATTGCAACT GCTAAACAAA ACATTGATCA	29820
	AGCGACGAAA CTTCAACCAA TTGCTGAATT AGTAGATCAA GCAACGCAAT TGAATCAATC	29880
50	TATGGATCAA TTACAACAAG CAGTTAATGA ACATGCTAAC GTTGAGCAAA CTGTAGATTA	29940
	CACACAAGCA GATTCAGATA AACAAAATGC TTATAACAA GCTATTGCTG ATGCTGAAAA	30000

TGCAAAACAA GCATTAAATG GTGATGAACG TGTAGCACTT GCTAAAACAA ATGGTAAACA 30120
 TGACATCGAC CAATTGAATG CATTAAACAA TGCTCAACAA GATGGATTTA AAGGTCGCAT 30180
 5 CGATCAATCA AACGATTAA ATCAAATCCA ACAAATTGTA GATGAGGCTA AGGCACCTAA 30240
 TCGTGCAATG GATCAATTGT CACAAGAAAT CACTGACAAT GAAGGACGCA CGAAAGGTAG 30300
 CACGAACTAT GTCAATGCAG ATACACAAGT CAAACAAGTA TATGATGAAA CGGTTGATAA 30360
 10 AGCGAAACAA GCACTTGATA AATCGACTGG TCAAACTTA ACTGCAAAAC AAGTTATCAA 30420
 ATTAATGAT GCAGTCACTG CAGCTAAGAA AGCATTAAAT GGTGAAGAAA GACTTAATAA 30480
 TCGTAAAGCT GAAGCATTAC AAAGATTGGA TCAATTAACA CATCTAAACA ATGCTCAAAG 30540
 15 ACAATTAGCA ATCCAACAAA TTAATAATGC TGAAACGCTA AATAAAGCAT CTCGAGCAAT 30600
 TAATAGAGCA ACTAAATTAG ATAATGCAAT GGGTTCAGTA CAACAATATA TTGACGAACA 30660
 20 GCACCITGGT GTTATCAGCA GCACAAATTA CATCAATGCA GATGACAATT TGAAAGCAAA 30720
 TTATGATAAT GCAATTGCGA ATGCAGCACA TGAGTTAGAT AAAGTGCAAG GTAATGCAAT 30780
 TGCAAAAGCT GAAGCAGAGC AATTGAAACA AAATATTATC GATGCTCAA ATGCATTAAA 30840
 25 TGGAGACCAA AACCTTGCAA ATGCCAAAGA TAAAGCAAAT GCGTTTGTTA ATTCGTTAAA 30900
 TGGATTAAAT CAACAGCAAC AAGATCTTGC ACATAAGCA ATTAACAATG CCGATACTGT 30960
 ATCAGATGTA ACAGATATTG TTAATAATCA AATTGACTTA AATGATGCAA TGGAACATT 31020
 30 GAAACATTTA GTTGACAATG AAATTCCAAA TGCAGAGCAA ACTGTCAATT ACCAAAACGC 31080
 TGACGATAAT GCTAAA 31096

(2) INFORMATION FOR SEQ ID NO: 60:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2243 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

45 ATGACAGAAT GGGAGCGAGG ACTTAGAATG TTTCTAAAT CAGGTTTATT AAATTTTGAG 60
 TTAGCGATAG TAAATCGTTC ATTAAATGAT GATGAAAAG CATTAAATA TGTGCGTAAA 120
 GCATTAAATG CAGACCTTAA AAATACAGAT TATATTAAT TAGAAAAGA GTTGACTAAA 180
 50 TCAAATGAGT CGAAAATAA ATAACCTTTA TGATGTACAA CAGTTATTGA AAAGTTACGG 240
 ATTTCTAATA TATTTTAAAA ATCCAGAAGA TATGTACGAA ATGATTCAAC AGGAGATTTC 300

55

	TAATCAGAGA AGGAATGAAC AGAAATGACA AAAATTATTT TAGCAGCTGA TGTAGGCGGG	420
	ACGACTTGTA AATTAGGTAT TTTCACACCT GAATTAGAAC AATTACATAA ATGGTCTATT	480
5	CACACTGATA CATCTGATAG TACAGGATAT ACACTTTGA AAGGAATTTA TGATTGTTTT	540
	GTTGAAAAAG TAAATGAAAA TAATTATAAT TTTTCAAATG TACTTGCGGT AGGTATTGGT	600
10	GTACCAGGTC CTGTTGACTT TGAAAAAGGT ACAGTAAATG GAGCAGTAAA CTTATATTGG	660
	CCAGAAAAAG TTAATGTACG TGAGATTTTT GAACAAATCG TTGATTGTCC AGTGATATGTA	720
	GATAATGATG CTAACATAGC TGCTTTAGGG GaGAAACACA AAGGTGCTGG TGAAGGTGCC	780
15	GATGATGTTG TTGCCATCAC ACTTGGTACA GGTCTAGGTG GAGGAATTAT TTCCAAATGG	840
	TGAAATCGTA CATGGTCATA ATGGCTCTGG CGCAGAAATA GGTCATTTTA GAgCAGACTT	900
	CgATCAACGA TTTaAATGTA ATTGTGGTCG TTCTGGATGT ATTGAAACAG TTGCTTCaGC	960
20	GACAGGCGTT GTTAACTTAG TTAACTTCTa CTATCCGAAG TTGACGTTTA GATCTTCTAT	1020
	ATTAGAATTG ATTAAAGAAA ATAAGGTtAC aGCAAAAGCT GTTTTTGATG CGGCAAAAGC	1080
	TGGTGACCAa TTCTGTATTT TCATTACTGA AAAGGTTGCA AACTATATTG GATATTTATG	1140
25	TAGTATTATT AGTGTTACAA GTAATCCGAA ATATATCGTT CTAGGTGGAG GAATGTCTAC	1200
	TGCAGGACCT ATTTTAATTG AAAATATTAA AACAGAATAT CATAATTTAA CATTTGCACC	1260
	TGCTCAATTT GAAACTGAAA TTGTACAAGC GAAATTAGGT AATGATGCAG GTATTACAGG	1320
30	AGCAGCAGGA TTAATCAAGA CCTATGTATT AGATAAAGAG GGGGTAAAAT AATGGCTATT	1380
	GTTGATGTGG TTGTTATTCC AGTTGGAACG GAAGGTCCGA GTGTTAGTAA ATATATTGCA	1440
35	GATATTGAGA AAAAaCTTCA AGAATATAAA GCAATGGGTA AAATTGATTT TCAATTAACA	1500
	CCAATGAATA CTCTAATTGA AGGTGAATTA AGCGATGTAT TAGAAGTTGT GCAAGTGATA	1560
	CATGaATTAC CTTTTGATAA AGGTTTAAGT AGAGTTTGTA CAAATATCCG TATTGATGAC	1620
40	CGACGAGACA AATCTAGAAA AATGAATGAT AAACtaACAT CAGTACAAA ACATTTAGAA	1680
	AATAGTGGTG AAAACCTATG AGGATTTCaA GCTTAACTTT AGGCTTAGTT GATACTAATA	1740
	CGTATTTTCAT CGAAAATGAC AAAGCTGTTA TTCTGATTGA CCCTTCAGGT GAAAGTGAAA	1800
45	AAATTATTAA AAAATTAAAC CAAATAAATA AACCGTTAAA AGCTATTTTA TTAACACATG	1860
	CACACTTTGA TCATATCGGA GCAGTCGATG ATATAGTTGA TCGATTGAT GTCCCGTTTT	1920
	ATATGCATGA AGCAGAGTTT GATTTTCTAA AAGATCCCGT TAAAAATGGG GCAGATAAAT	1980
50	TTAAGCAATA TGGATTACCA ATTATTACAA GTAAGGTAAC TCCTGAAAAG TTAAmCGAAG	2040
	GTAGCACAGA AATAGAAGGA TTTAAGTTnT nAy-TGTaCA CACACCTGGA CATTACCAG	2100

55

GAATCGGACG TACAGATTTA TATAAAGGTG ATTATGAAAC GCTAGTTGAT TCTATTCAAG 2220
 ATAAAATATT TGAATTAGAA GGC 2243

5 (2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8009 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

15 TTGGnATCAT tyAcgGTAA AAGAATAAaG CAAGATTtAT TTCATTAGTA CTAATTTGTG 60
 CAATGTTTGC AATTTGTTGG GTTGCATATA TTCAATGGGA GTCTACAATC GCTTCATTTA 120
 20 CACAATCTAT TAATATTTCa ATGGCACAAT ATAGTGTTTT ATGGACAATT AACGGAATAA 180
 TGATTTTAGT AGCACAACCA TTAATTAAAC CGATTCTCTA TCTGTAAAAA GGAAACTTAA 240
 AGAAGCAAAT GTTTGTCGGC ATCATCATT TTATGTTGTC GTTCTTTGTC ACGAGTTTTG 300
 25 CCGAAAACTT TACAATATTT GTTGTGCGTA TGATTATTTT AACTTTTGGA GAAATGTTTG 360
 TATGGCCAGC AGTTCCAAC ATAGCCAATC AGTTAGCGCC AGATGGTAAG CAAGGACAGT 420
 ACCAAGGTTT TGTGAATTCA GCTGCTACAG TAGGAAAAGC ATTTGGTCCA TTTCTTGGTG 480
 30 GTGTATTAGT TGATGCGTTT AATATGCGCA TGATGTTTAT CGGTATGATG CTACTACTTG 540
 TATTTGCATT AATATTATTA ATGGTTTTCA AGGAGAATAA TACGCAACCT AAAAAAATAG 600
 ATGCATAATG AGTAAATAGA ATTAACGTTA TAGACTTGAA ATAAATGTCG TTATAACATA 660
 35 ATATTAATTT GTATAATTTA ATTTGTTTTG GAGCTTTTCT ACAGAAAGCT AGTGATGCTG 720
 AGAGCTAGTG TTAAGGACTA AATGTAAATC GTATTAATTT TAAATTGAAT GAATGACATC 780
 TCTTACTATT AAAATGAGTG CACAATTTTT GTGAAATAGG GTGGTAACGC GGCAATGTC 840
 40 GTCCCTATGT AAATAGAATA GTTAGAGGTG TCTTTTTTAT TGAATAGGAG GAAATGTGTT 900
 GAATTACAAC CACAATCAAA TTGAAAAGAA ATGGCAAGAC TATTGGGACG AAAATAAAAC 960
 45 ATTTAAAACA AATGATAACT TAGGTCAAAA GAAATTTTAT GCTTTAGACA TGTTTCCATA 1020
 TCCATCAGGT GCTGGTTTAC ATGTTGGACA TCCTGAGGGC TATACAGCAA CAGATATCAT 1080
 TTCAAGATAT AAAAGAATGC AAGGATATAA TGTATTACAT CCGATGGGGT GGGATGCATT 1140
 50 CGGATTACCA GCAGAGCAAT ATGCTTTAGA CACTGGCAAC GACCCACGTG AATTTACAAA 1200
 GAAAAATATC CAAACTTTTA AACGACAAAT TAAAGAATTA GGGTTCAGTT ATGATTGGGA 1260

55

GTTATATAAC AAAGGTTTAG CATACGTTGA TGAAGTTGCA GTTAACTGGT GTCCAGCATT 1380
 AGGCACTGTT TTATCTAACG AAGAAGTGAT TGATGGTGTC TCTGAACGTG GTGGACATCC 1440
 5 AGTTTATCGT AAGCCGATGA AACAATGGGT ACTTAAAATC ACAGAATATG CAGATCAATT 1500
 ATTAGCAGAT TTAGATGATT TAGATTGGCC TGAGTCTTTA AAAGATATGC AGCGCAATTG 1560
 GATTGGACGT TCTGAAGGGG CCAAAGTTTC ATTTGATGTA GATAATACGG AAGGAAAAGT 1620
 10 AGAAGTATTT ACGACTAGAC CAGATACAAT CTATGGTGCA TCATTCTTAG TCTTAAGTCC 1680
 TGAACATGCA TTAGTTAATT CAATTACAAC AGATGAATAT AAAGAAAAAG TAAAAGCTTA 1740
 TCAAACAGAA GCTTCTAAAA AGTCAGATTT AGAACGTACA GATTTAGCAA AAGATAAATC 1800
 15 AGGTGTATTT ACTGGTGCAT ATGCAACTAA TCCTTTATCT GGTGAAAAAG TACAAATTTG 1860
 GATTGCTGAT TATGTATTAT CAACATATGG TACTGGAGCA ATTATGGCAG TACCAGCGCA 1920
 20 TGATGACAGA GATTATGAAT TTGCTAAAAA GTTTGATTG CCAATCATTG AAGTCATCGA 1980
 AGGTGGAAAT GTTGAAGAAG CAGCATACAC TGGTGAAGGT AACATATTA ATTCTGGTGA 2040
 ACTTGATGGT TTAGAAAATG AAGCGGCAAT TACTAAAGCT ATTCAATTAT TAGAGCAAAA 2100
 25 AGGTGCTGGC GAAAAGAAAG TTAATTACAA ATTAAGAGAT TGGTTATTCA GTCGTCAGCG 2160
 TTATTGGGGC GAACCAATTC CTGTCAATTCA TTGGGAAGAT GGAACAATGA CAACTGTTCC 2220
 TGAAGAAGAG CTACCATTGT TGTACCTGA AACAGATGAA ATCAAGCCAT CAGGGACTGG 2280
 30 TGAGTCTCCA CTAGCTAATA TTGATTCATT TGTAATGTT GTAGATGAAA AACAGGTAT 2340
 GAAAGGACGT CGTGAAACAA ATACAATGCC ACAATGGGCA GGTAGTTGTT GGTATTATTT 2400
 ACGTTACATC GATCCTAAAA ATGAAAATAT GTTAGCAGAT CCTGAAAAAT TAAACATTG 2460
 35 GTTACCTGTT GATTTATATA TCGGTGGAGT AGAACATGCC GTTCTTCACT TATTATATGC 2520
 AAGATTTTGG CATAAAGTCC TTTATGATTT GGCTATCGTA CCTACTAAAG AACCTTTCCA 2580
 AAAATTATTT AACCAAGGTA TGATTTTAGG AGAAGGTAAT GAGAAGATGA GTAAATCTAA 2640
 40 AGGAAATGTA ATCAATCCTG ATGATATAGT ACAGTCTCAT GGTGCAGATA CTTTGCCTCT 2700
 TTACGAAATG TTTATGGGAC CTTTAGATGC TGCAATTGCA TGGAGTGAAA AAGGATTAGA 2760
 45 TGGGTCTCGT CGATTCTTAG ATCGCGTATG GCGTTTAAAT GTAAATGAAG ATGGGACATT 2820
 GAGTTCAAAA ATTGTAAC TAATAATAA ATCTTTAGAT AAAGTTTATA ACCAACTGT 2880
 TAAAAAGGTA ACAGAAGACT TTGAAACATT AGGATTTAAT ACTGCTATTA GTCAATTAAT 2940
 50 GGTATTTATT AATGAGTGTT ATAAAGTTGA TGAAGTTTAT AAACCTTACA TTGAAGGCTT 3000
 CGTTAAAATG TTAGCACCTA TTGCACCACA TATCGGTGAA GAATTATGGT CAAAATTAGG 3060

	TGATGAAGTA	GAAATCGTTG	TTCAAGTGAA	TGGTAAATTG	AGAGCTAAAA	TTAAAAATTGC	3180
	TAAAGATACA	TCAAAAGAAG	AAATGCAAGA	AATTGCCTTA	TCTAATGACA	ATGTTAAAGC	3240
5	GAGTATTGAA	GGTAAAGACA	TCATGAAAGT	CATCGCTGTT	CCTCAAAAAT	TAGTCAATAT	3300
	TGTAGCTAAA	TAATGTTTTA	AGGAGGACTT	TGAAATGAAG	TCAATTACTA	CAGATGAATT	3360
	AAAAAATAAA	CTTTTAGAAT	CTAAACCACT	TCAAATTGTT	GATGTTTCGT	CTGATGAAGA	3420
10	AACAGCAATG	GGATATATTC	CTAATGCAAA	GTTAATTCCA	ATGGATACCA	TTCCGGATAA	3480
	TTTAAATTCA	TTTAATAAAA	ATGAAATATA	TTATATTGTA	TGTGCTGGTG	GAGTTCGAAG	3540
	CGCTAAAGTT	GTAGAATATT	TAGAGGCAAA	TGGCATTGAT	GCCGTAAATG	TCGAAGGCGG	3600
15	CATGCACGCA	TGGGGCGATG	AAGGTTTGGG	AATAAAAAGT	ATTTAAAGTA	GTGACATAAT	3660
	TTAAATAAAT	ATTACATTTG	TAATGACACC	AAGTAACGTT	TCGTTTGCTT	GGTGTTTTTT	3720
20	GGTATGAATT	ACTTTCTGTT	ACAAAACAAT	CTAAAGCGTT	CTTGTTATGT	TTTATTAAGA	3780
	TTTTAATTAC	AAAACGGAAA	CTAAATTGTA	ATAAAATAAA	ACTTTATTTT	ATAAAATGAT	3840
	GATGATAAAA	TTGAGTGAAC	TTAAAATATT	GTACAAAATA	ATATAGCTAT	AAATATAATA	3900
25	TAGCTATAAA	TATAATATGA	GGGAGCGTAT	ATTTTATAGCA	TAATTCTTAA	CAACACAGCA	3960
	GAGAACAGAC	AACCAGGAGG	AAAATGAAAT	GAATTGTGTA	AAGAAAAATA	AATATAGTAT	4020
	TAGGAAGTAT	AAAGTAGGCA	TATTCTCTAC	TTTAATCGGA	ACAGTTTTAT	TACTTTCAAA	4080
30	CCCAAATGGT	GCACAAGCCT	TAACACGGA	TAATAATGTA	CAAAGCGATA	CTAATCAAGC	4140
	AACACCTGTA	AATTCACAAG	ATAAAGATGT	TGCTAATAAT	AGAGGTTTAG	CAAATAGTGC	4200
	GCAGAATACA	CCTAATCAAT	CTGCAACAAC	CAATCAAGCA	ACGAATCAAG	CATTGGTTAA	4260
35	TCATAATAAT	GGTAGTATAG	TAAATCAAGC	TACGCCAACA	TCAGTGCAAT	CAAGTACGCC	4320
	TTCAACACAA	AACAATAATC	ATACAGATGG	CAATACAACA	GCAACTGAGA	CAGTGTCAAA	4380
	CGCTAATAAT	AATGATGTAG	TGTCGAATAA	TACCGCATT	AATGTACCAA	CTAAAACAAA	4440
40	TGAAAATGGT	TCAGGAGGAC	ATCTAACTTT	AAAGGAAATT	CAAGAAGATG	TTCGTCATT	4500
	TTCAAATAAA	CCAGAGCTAG	TTGCAATTGC	TGAACCAGCA	TCTAATAGAC	CGAAAAAGAG	4560
45	AAGTAGACGT	GCGGCACCGG	CAGATCCTAA	TGCAACTCCA	GCAGATCCAG	CGGCTGCAGC	4620
	GGTAGGAAAC	GGTGGTGAC	CAGTTGCAAT	TACAGCGCCA	TATACGCCAA	CAACTGATCC	4680
	TAATGCCAAT	AATGCAGGAC	AAAATGCACC	TAACGAAGTG	CTGTCATTTG	ATGACAATGG	4740
50	TATTAGACCA	AGTACCAACC	GTTCTGTGCC	AACAGTAAAC	GTTGTTAATA	ACTTGCCGGG	4800
	CTTCACACTA	ATCAATGGTG	GCAAAGTAGG	GGTGTTTAGT	CATGCAATGG	TAAGAACGAG	4860

55

	TCGTATACAT GGAAGTATA CGAATGACCA TGGCGATTTT AATGGTATCG AGAAAGCATT	4980
	AACAGTAAAT CCGAATTCTG AATTAATCTT TGAATTTAAT ACAATGACTA CTAAAAACGG	5040
5	TCAAGGCGCA ACAAATGTTA TTATCAAAAA TGCTGATACT AATGATACGA TTGCTGAAAA	5100
	GACTGTTGAA GCGCGTCCAA CTTTGCGTTT ATTAAAGTA CCTGATAATG TGAGAAATCT	5160
	CAAAATTCAA TTTGTACCTA AAAATGACGC AATAACAGAT GCGCGTGGCA TTTATCAACT	5220
10	AAAAGATGGT TACAAATACT ATAGCTTTGT TGAAGTCTATC GGACTTCATT CTGGGTCACA	5280
	TGTTTTTGTT GAAAGACGAA CAATGGATCC AACAGCAACA AATAATAAAG AGTTTACTGT	5340
15	AACAACATCA TTAAAGAATA ATGGTAATTC TGGTGCTTCT CTAGATACAA ATGACTTTGT	5400
	ATATCAAGTT CAATTACCTG AAGGTGTTGA ATATGTGAAC AATTCATTGA CTAAAGATTT	5460
	TCCAAGTAAC AATTCAGGCG TTGATGTTAA TGATATGAAT GTTACATATG ATGCAGCAAA	5520
20	TCGTGTGATA ACAATTAAAA GTACTGGAGG AGGTACAGCA AACTCTCCGG CACGACTTAT	5580
	GCCTGATAAA ATACTCGATT TAAGATATAA ATTACGTGTA AATAATGTGC CGACACCAAG	5640
	AACAGTAACA TTTAACGAGA CATTACGTA TAAACATAT ACACAAGATT TCATTAATTC	5700
25	AGCTGCAGAA AGTCATACTG TAAGTACAAA TCCATATACT ATCGATATCA TCATGAATAA	5760
	AGATGCATTA CAAGCCGAAG TTGACAGACG TATTCAACAA GCTGATTATA CATTGCGTC	5820
	ATTAGATATC TTTAATGGTC TGAAACGACG CGCACAAACG ATTTTAGATG AAAATCGTAA	5880
30	CAATGTACCA TTAAATAAAA GAGTTTCTCA AGCATATATT GATTCATTAA CTAATCAAAT	5940
	GCAACATACG TTAATTCGAA GTGTTGATGC TGAAAATGCA GTTAATAAAA AAGTTGACCA	6000
	AATGGAAGAT TTAGTTAATC AAAATGATGA ATTGACAGAT GAAGAAAAAC AAGCAGCAAT	6060
35	ACAAGTTATC GAGGAACATA AAAATGAAAT AATTGGTAAT ATTGGTGACC AAACGACTGA	6120
	TGATGCGGTT ACTAGAATCA AAGATCAAGG TATACAGACC TTAAGTGGGG ATACTGCAAC	6180
40	ACCGGTTGTT AAACCAAATG CTAAAAAGC AATACGTGAT AAAGCAACGA AACAAAGGGA	6240
	AATTATCAAT GCAACACCAG ATGCTACTGA AGACGAGATT CAAGATGCAC TAAATCAATT	6300
	AGCTACGGAT GAAACAGATG CTATTGATAA TGTTACGAAT GCTACTACAA ATGCTGACGT	6360
45	TGAAACAGCT AAAAATAATG GCATCAATAC TATTGGAGCA GTTGTTCCCTC AAGTAACTCA	6420
	TAAAAAGCT GCAAGAGATG CAATTAACCA AGCAACAGCA ACGAAAAGAC AACAAATAAA	6480
	TAGTAATAGA GAAGCAACTC AGGAAGAGAA AAATGCAGCA TTGAACGAAT TAACTCAAGC	6540
50	AACCAACCAT GCTTTAGAAC AAATCAATCA AGCAACAACA AATGCTAATG TTGATAACGC	6600
	CAAAGGAGAT GGTCTAAATG CCATTAATCC AATTGCTCCT GTAAGTGTG TTAAGCAAGC	6660

TGATGCGACT CAAGAAGAAA GACAAGCAGC AATTGACAAA GTGAATGCTG CTGTAACTGC 6780
 AGCAAACACA AACATTTTAA ACGCTAATAC CAATGCTGAT GTTGAACAAG TAAAGACAAA 6840
 5 TGCGATTCAA GGAATACAAG CAATTACACC AGCTACAAAA GTAAAAACAG ATGCAAAAAA 6900
 TGCCATCGAT AAAAGTGCGG AACCGCAACA TAATACGATA TTTAATAATA ATGATGCGAC 6960
 GCTCGAAGAA CAACAAGCAG CACAACAATT ACTTGATCAA GCTGTAGCCA CAGCGAAGCA 7020
 10 AAATATTAAT GCAGCAGATA CGAATCAAGA AGTTGCACAA GCAAAAGATC AGGGCACACA 7080
 AAATATAGTA GTGATTCAAC CGGCAACACA AGTTAAAACG GATACTCGCA ATGTTGTAAA 7140
 TGATAAAGCG CGAGAGGCGA TAACAAATAT CAATGCTACA ACTGGCGCGA CTCGAGAAGA 7200
 15 GAAACAAGAA GCGATAAATC GTGTCAATAC ACTTAAAAAT AGAGCATTAA CTGATATTGG 7260
 TGTGACGTCT ACTACTGCGA TGGTCAATAG TATTAGAGAC GATGCAGTCA ATCAAATCGG 7320
 CGCAGTTCAA CCGCATGTAA CGAAGAAACA AACTGCTACA GGTGTATTAA ATGATTTAGC 7380
 20 AACTGCTAAA AAGCAAGAAA TTAATCAAAA CACAAATGCA ACAACTGAAG AAAAGCAAGT 7440
 GGCTTTAAAT CAAGTGGATC AAGAGTTAGC AACGGCAATT AATMATATAA ATCAAGCTGA 7500
 25 TACAAATGCG GAAGTAGATC AAGCGCAACA ATTAGGTACA AAAGCAATTA ATGCGATTCA 7560
 GCCAAATATT GTTAAAAAAC CTGCAGCATT AGCACAAATC AATCAGCATT ATAATGCTAA 7620
 ATTAGCTGAA ATCAATGCTA CACCAGATGC AACGAATGAT GAGAAAAATG CTGCGATCAA 7680
 30 TACTTTAAAT CAAGACAGAC AACAGCTAT TGAAAGTATT AAACAAGCTA ACACAAATGC 7740
 AGAAGTAGAC CAAGCTGCGA CAGTAGCAGA GAATAATATC GATGCTGTTC AAGTTGATGT 7800
 AGTAAAAAAA CAAGCAGCGC GAGATAAAAT CACTGCTGAA GTGGcGAacG TATTGaAGCG 7860
 35 GTTAAACAAA CACCTAATGC AACTGACGAA GAAAAGCAGG CTGCTGTTAA TCAAATCCAA 7920
 TCAACTTTAA AGATTCAAGC AATTTAATCC AAATTTAATC CAAAACCCAA ACAAATGGAT 7980
 TCAGGGTAGG ACACCACTTA CAAATCCAA 8009

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10953 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

ACCCACCCcn TGGGGATAnT TTACCTGGTG GGGCCTTCGA TTGCCTTTAG GTGAAACCaG 60

	AGATGAATGC TAACCATATT CATTCTGCTA AAGATGGTCG TGTTACTGCG ACAGCTGAAA	180
	TTATTCATCG AGGTAAGTCG ACACATGTAT GGGATATAAA AATTAAGAAT GACAAAGAAC	240
5	AATTAATTAC AGTTATGCGT GGTACAGTTG CTATTAAACC TTTAAAATAA AAGAACTGCT	300
	AGCTGAAATG TTATGAGATA TTCATAACTA CGGCTAGCAG TTTTTTTATG CGCTATATTG	360
	TTGTAGTTTT AGAAATGCTT GTTCAATGCG TTCGGCAGCT TTACGGCCAC CCATAACATT	420
10	TCTACCAAAT GGTCTTAATT CTAAGTCTGC AAAGCATCCT GCGACAAATA GATTGGTAT	480
	CCATTCTAAT TTTTCGGAAA TAACAGGGTA ATTACATTCG TTGATAGGTG CATCATAATT	540
	TTGTATTAAAT TGCTTAATAA GTGGTTGTGA CATAAAATCT TGTTCAAAAC CAGTTGCAAC	600
15	CATAATCTGT TGATATGGAA CAGAATCATT TTCAGTGTTA ATTACACCAC CACTAATTTG	660
	AGTGATAGGT GTTTTATGCa CATTTATACG ACCATTTTTA ATATGTTTTT TAAGGCGTAA	720
20	GTACAGTTCG TGAGGCATTG ATCCTTTATG ACGTTCGCGT TGTACAATGG CATTTCTTTC	780
	AGGCATGCTT TTAGTACTTA AAAATGAAGA CATATTTTTC GGACCTAACC AACCAGGATC	840
	AGCATCAAAG TCATGTATTT CAATATCTTT ATTTAGCCAT AAATGAATCT TTTTATCGTT	900
25	ATCATGATTT AACAAITTTAA GTGCAAGATG TGCAGCAGTa ATGCCGCTAC CAACGATATG	960
	ATCGGTCTTA TCATATACTA CTTGATCAAG TTCTTTCTCG AAGATATGAT TTACATTCTG	1020
	TTTGTCTTTT AAAATGTCAG GCATAAACGG AATATTTGTA CTGCCTATTG CAATAACGAC	1080
30	GCAATCTGTA GTGATAATTT GTCCATCTTC TAACTTGATA TGCCATTTGT CTCTTGTTTT	1140
	ATCTAAAGTT TGAACATAAC CTTGAACCAA GCAATCCTCT AATTGATATT GTTTAGAAGC	1200
	ATGTGCAATA TGATCCATAA ACATTGTCAA TTCAGGTCGT TGATAAGGAC CATAAAAAGC	1260
35	ATTTGTATAT TGGTGCTGTT TAGCGAATTG TTTTAGATGG AACGGTTGTG GATGTACGTG	1320
	ATGTACAATC GGTGATCTTA AATAAGGCAT TTCTATTCTGA TTTGTATATG AGTTAAACCT	1380
40	TTGGCAAAAA GTTTCGTGTG GGTCAATGAT TGTTAATCGG TCTGTTGTTA ATCCGCTTGA	1440
	TAATAGTTTT TGTGCGATTG CAGTTCCTCG TATGCCACCG CCGATAATTG TCCAATGCAT	1500
	AATAAACCT CTCTCTTTTT AAAACGTAAT AGTTACGATT TATAATTATT ATTATCATAA	1560
45	TACATAACGA CATGAAAGGC AATTAAATTA AAGAGATATA TGTAGATAGG GCGAATCTGT	1620
	AGTCAAAGAA AAAATCATTG AAAAGAGGT AACAATGTCA AAAGAWACA GCAGTAAAT	1680
	CATTCCTAAT TTGGAATCAT CTTACTGCTG TTTGTTGTG ATTTATATTC ATGATTTTGT	1740
50	TATATAATCT ACAATTTTGT GTCTTTTAAG TCTTCCGAAA TTTTCATCGAC TTTAGTCTTT	1800
	TTAGTATAAG GCGTTTTAAT ATTATATGCT GCTTTCATAA TCATATGACT TGAAAGAGGA	1860

55

	GCAATAAAAT ATAAAAACGT ACCAAATAGT AATGACATTG CACCTAATGT TGATGCTTTT	1980
	CCGGCAGCAT GTGCACGTGA ATATACATCT TCAAGTCTCA ATAATCCTAT AGCTGCTAGG	2040
5	GCGCTAATTA AAGCACCGAT GATAACAAAG ATAAGTGCAA GACTAATCAG TATGATTTTG	2100
	ATCATGTTCA ATCACCTTAC CTTTGTCCAT AAATTTAGAG AATACTGCAG TACCTAAAAA	2160
	AGCTAATATA CCAATCATCA TAATAACGAC AATCATGTAT TTAATATTTA ATAAAATACT	2220
10	GAATAATGCT ATAACTGCCA TTAATTGAAG ACCAATCGCA TCTAATGCGA CAACACGATC	2280
	GGCAAGTGAT GGGCCTAGCA CAACGCGAAT GAGCATAGCT AACATAGAAA TGACAACTAT	2340
15	GATTAATGCA ATAACGATAA TAACATTATG ATTCAATTATA TTTCGCCCAC CTCTCTTACA	2400
	ATTTTCTCTA ATGATGTTTT AATACTTTCT ACTTCTTGCT CTTTAGTTGA AAAATCTATG	2460
	GCATGAATAT AAATTTTGT ACGATCGTCA CTTACACCAA GCACTACAGT ACCAGGTGTT	2520
20	AATGTAATTA AATTAGACAG CAAGACAATT TGCCAATCTT TTTTAAATC TGTGTGATAA	2580
	ACAAAGAATC CTGGTTCATT TTTAATCGAA GGTTTAATAA TAATTTTCAA AACATCAAAA	2640
	TTAGCTTTAA TCAGTTCGAT TAAGAAAATA ATAECTAATT TAATAATACG ATATAGCGTG	2700
25	ATGACATAAA ATCTACCTGG TAACACTCTG TGTAAGAGGT AAACAAGAAC TAGGCCAAAG	2760
	ATGAAACCTA ACACAAAGTT ATTTGTTGTG TAACTATTTG TCACAAACAA CCAAAACACT	2820
	GCGATAATAA AGTTTAATAC TAATTGTACA GCCATGTTAT TTACCTCCTA ATACAGCTTT	2880
30	AACGTAGGTT GATGGATTGT AGAATGTTTC TGCACCAGCT TTTACCATTG GATATAAGTA	2940
	ATCTGCTGAC AATCCATATA AAACAGTTAT CACAACCTGCA ACGATTGCAA TCGTAGTTAA	3000
	ATATTTGACG TCGACTTTGT TATTAAGATC ATATCCTTTT GGTTGACCGA AAAAGCCTTG	3060
35	TAGGAATATG CGAATGACAG AATATAATAC GACTAAACTT GATAATAAGA CGATGACACC	3120
	ACTTAAATAA AATCCTCTTT CAAATGTTGA TTGGACAATA AAAAATTTTC CATAAAAGCC	3180
	ACTGAGTGGG GGAATGCCAG CTAAACTTAA TGCTGCGATA AAGAATGACC AACCAGTAC	3240
40	AGGATATCGT TTAATTAAGC CACCAAATTG TCTTAAATCA GCAGTGCCCTG TAATTTTAAT	3300
	CATAATTCCG ATAAGCAAGA ATAATGCAAG TTTTACTAAC ATGTCGTGCA ATGTATAGTA	3360
45	AATAGCCCCA ATCATACCTG ACTCTGTCAT CATTGCAACG CCGACTAAGA TCACACCTAC	3420
	AGCAATCATG ACATTGTATA GGATGATTTT TTTAATGTTG GCATATGCAA CAGCACCGAC	3480
	ACAACCAAAG ATGATCGTTA ATAGTGCTAA GAATAAAATG ACATAATGTG AAAAGCTTAC	3540
50	ATTATCACTA AAGAATAGGC TCAATGTTCT AGCGATTGCA TAAACACCAA CTTTGTTTAA	3600
	CAAAGCACCA AAGAATGCAA TGATTGGAAT TGGTGGCAT AGTATGCACT AGGTAACCAA	3660

55

	ATATTGACTA AGCCACTGTC ATGCGCTGAA AGGTTAGCTA ATTTATTGCT TATATCTGCT	3780
	AGATTCAATG TTCCTACTAC TGAATATAAA ATCCTACAC CCATTACGAA GAAGGATGAC	3840
5	GATACAACGT TAACAAGAAC ATATTTTATT GTTCTTGTA GTTGAATTTT TGTAGAACCA	3900
	ATTACTAATA AGAAATAAGA TGACATTAAA AATACTTCGA AAAATACGAA TAGGTTGAAA	3960
	ATGTCACCAG TTGTGAATGC ACCAATGATA CCTATTAACA TAAATAGTAC TGAAAAATAA	4020
10	TAATAATATC TTTCACGTTC AATACCAATT GTTTGGTATG AATATAAAAT CACAATAGCT	4080
	GTAATAATAA TACTAGTAAT TATTAGTAGG GCACTGAATA TGTCTAATAC AAAGACAATA	4140
	CTGTATGGTG CTTTCCATGA ACCTAGCTCT ACGCGTATTG GTCCATGTTT AACACATT	4200
15	GCTAAATTGA TAATTGCCGC GACCAAGGTT AATAATGTAC CGCCTAGTGC GACATAACGC	4260
	TTTATAATAG GACGCTTTCC AATAAAGACA AGTAATATGG CTGTAATTAC TGGAATAACT	4320
20	AGCGTTAACA CAAGCATATT ACTTTCAATC ATCTTCTGGA ACTCCTTTCA TACTCTCAAC	4380
	GTTATCTGTG CCTAATTCTT TATATGTTCT AAATGCTAAT ACTAAGAAAA AGGCTGTTGT	4440
	CGCAAGCGCA TAACGATTGC TGTTAAAATA AGTGCTTGCG GGaTAGGaTC AACATAGCTT	4500
25	TTTACGTTTC CTTCATAAAT TGGAACAGTA CCATGTTTAA GTCCGCCCAT AGTTATTAAA	4560
	AATAAATTTG CTGCATGTGT TAATAGTGTA GTTCCATAA CAATTCGTAT CAGACTTTTA	4620
	GACAAAACGA GATAGACACT AATTGCTGTG AGAATACCAC TAACAAAAAT CATAATAATT	4680
30	TCCACTATTG GTTCTCTCCA ATCGAAATAA TAATTGTCTAT GACAGTACCA ACTACTGCAC	4740
	ATAAAACACC GAAATCAAAG AATACTGCTG TTGTCATATG AACAGGTTCT AATATAAATA	4800
	ACGGTATATC AAATGTGACA TGCGTAAAGA AATTTTTGCC TAAAAACCAA CTTGCGATAG	4860
35	GCGTCGCAAT ACAAAAAACT AATCCGATAC CTATCAAGAT TTTAAATCT AATGGGAAAA	4920
	TTTTACGCAT TGTTTCTATA TCAAATGCAA TCGTAATGAT AACAAAGTGAA CTTGCGAATA	4980
40	ATAATCCGCC GACGAAACCG CCACCAGGTG TATAATGTCC TGCTAAGAAA AGTGAAAAAC	5040
	CAAAGACCAT TACCATGAAA AAGATAATAA CTGCAGCAAA TTGCAAAATT AGATCATTTT	5100
	GTTGTCTATT CATGATTTTT CACCTCGTTA CCTTGCCTTT GACGCTTTTT ACGTAATTTA	5160
45	ATCATTGTAT ATACAGCTAA TCCTGCGATA CCAAGCACAG ATGACTCGAA TAAAGTATCC	5220
	ATACCACGGA AATCAACAAG TATGACGTTT ACCATGTTTT TACCGTGAGC tAAATCATAA	5280
	ACGTGCTCTT GATAAACTT AGATATCGAT TCAAAATGTC TATTTCCGTA TGCAATTAAA	5340
50	CCGATAATAA TGACGGACAA ACCAACACCA CCAGCAATTA AAGCATTAGT AAGCTGGAAT	5400
	GAGCGCTTTT CATTATAACG ATTTAAATTT GGTAAGTGGT AGAAGCATAA TAAGAACAAT	5460

55

	ATAACAATA CAGACACAGC ATATCCAAC TCACTTAACA TAATGATGCT AAATAATCTT	5580
	GATTTAGCGA AAAGAATTAA AAAGGCAGCA CTTAATAATA AAATTACGAT ACAAACCTCG	5640
5	AAAATTCTAA TCGGACTAAC GTCTTTAAAA TTAATGTTGA AAGGTACTGA GAATATAGTG	5700
	ACAAATGTTA ATAAAATTAA TGCACCAAAA ATGATAACTA AATTATTACG TGAATAATCG	5760
	GTAACATAGC TATTCGTCAT CTTTTTCAGAG TAGTTTGGA TAACATTTGC ACTTCTGTTG	5820
10	TACCAATAAT TGAATGTTAG TTTACCAGGT TGTCGTTGCA ACAATTTTAC CCAATAACTA	5880
	AATGTCACAA TTAGTAAGAT ACCTAAAATA TAAATCACTA ATGTTGATAA AAAGGCAGGC	5940
	GTAAATCCAT GGAACATATG GAATTCAACA TCATCAATTA CCGTATGATT AATCGAAGag	6000
15	TnAGCTGGTT CAATAATCGA ATTAGTTAAA ATGCCAGGGA ATAAACCAAA TACAATTACT	6060
	AATGTAGCTA AAATAGCTGG TGATAAAAGC ATTAATATTG ATACTTCGTG TGCTTTTTTA	6120
20	GGTAATTGTT CAGGTTTATA TTGTCCGAAA AATATATGCA TTATAAATTT AATTGAATAT	6180
	ACAAATGTGA AGACACTGCC CACTATACCA ATGATTGGGA ATAGGTAGCC TAATGTATCA	6240
	ACACTGAATA AATTTGCTTG GCTTGCTGTA AATGTTGTTT CTAAAAATGA TTCTTTTGAT	6300
25	AAGAAACCAT TGAACGGTGG TACACCAGCg CATACTTAAT GCTGTAATAA CAGTGATTGT	6360
	AAATGAAATA GGCATAATTG TTAGTAAGCC ACCTAATTTT TTAACATCAC GTGTACCAGT	6420
	AGAATGATCC ACTGCACCTG TAATCATAAA TAGGGCACCT TTAAATGTTG CATGGTTGAT	6480
30	TAAATGGAAT ATTGCAGCCG TAAATGCAGC AGCATATATT TTGCTATCAT CGCCTTGATA	6540
	GTGATAACTA ATGGCACC GA TTCCAAGCAT CGCCATAATC ATACCTAATT GGGATACTGT	6600
	TGAAAATGCC AGTATACCTT TCAAGTCTTG TTGTTTTGTT GCGTTTAGCG AAgCCCAGAA	6660
35	TAATGTAATT AAACCAACGA GTGTGACAGT CCATACCCAA CCTTGCGATG CTGCGAAGAT	6720
	TGGTGTcATT CGAGCGATTA AATATAACCC TGCTTTAACC ATTGTTGCTG AATGAAGATA	6780
	AGCACTGACT GGTGTAGGTG CTTCCATTGC ATCTGGTAGC CAAATATAAA ATGGAACTG	6840
40	AGCAGATTTT GTAAAAGCAC CAATCATGAT TAAAATCATC GCAAAAATGA AGAATGGGCT	6900
	ATTTTGAATT TCAGAAGCAT GTTGAATCAT GTACTGAATG CTAAATGATT GTGTTGGTAT	6960
45	AGCGAGTAAG ATGATACCAC CTAATAATGA TAGACCACCA AATACTGTGA TTATGAGCGA	7020
	TTTTTGAGCA CCATATATAG ATGCTTGTCG TTCGCGCCAG AATGAAATAA GTAAAAAACT	7080
	AGAAAATGAC GTTAGCTCCC AGAATAAATA TAGAATAATA ACATTATCTG AAAGTACGAC	7140
50	ACCTAACATT GCACCCATAA ATAGTAATAA ATAACAATAA AAATTCCTA GTTGTTCTGA	7200
	CTTACTTAAG TAGCCGATTG AATATAATAC TACTAACTG CCGATTCCCTG AAATAAGCAA	7260

55

	CCAATTTAAG GTTTTCAATTA CAGTATTACC TGACATCGTC GTTTTAATTA ATGTAAGCAT	7380
	ATAAATAAAT ATGACGATAG GGACAGGTAA TACGAACCAT CCTAAATGTA TACGTTTAAA	7440
5	AAATCTATAC AGGATAGGAA TAATGAGTGC GAATATTAAC GGTAATATCA CCGCAATATG	7500
	TAACAACTC ACTATGTTGT CCTCCTTTAA AAAATATTTA TGTTATTCAT TATACATGAA	7560
	TGATATAGTT CTGAAAAACG TACACACTCC TTGTTGTGCT TTATTTTCAG AaGTATTTAA	7620
10	ATAAGAAGAA ACACGTCATT TTTTATTTAA AATTTTCTTT GTATTGAAGT GAATAATCTT	7680
	CTTTTAAGCG TGCTAAACTA GCTAAAGACA TTTCAGCATG TTTTGTTTGC TGAGCTTTAA	7740
	GTTTAGTTTC TAAATCTGTA ATTGCTTGTT GAAGTGAATC TTCATAGCGC AATACATCAA	7800
15	CATTGAAGTC GCGTAATTGT GAACGTTTCG TATAGCGTTT TTCAAAATGG CTTAATGCTT	7860
	TGCGGTCATG GAAAAATACA CCTTCAGTTT CAGTAGGGTT ATGTAAATCA CCTTGTTTCG	7920
20	GGTGTTTGAT AACTTGTTCA ACTTTAACAA GGACATCGTC TCCATTTTCT TCAACAATCG	7980
	TGACACCATA GCTACCTGTT TTGTGTGAAA ATCGATATAG CTTCATGCTA TTTTCCTCCC	8040
	TTAAAGTAT GTTAATATAT ATGTATCATA ACATGAATGG AGAATATAAA TGGCTAACTA	8100
25	TCCACAGTTA AACAAAGAAG TACAACAAGG TGAAATCAAA GTGGTTATGC ACACAAATAA	8160
	AGGTGACATG ACATTCAAAT TATTTCCAAA TATTGCACCA AAAACAGTTG AAAATTTTGT	8220
	GACACATGCA AAAAATGGTT ATTATGATGG AATCACATTC CACCGTGTCA TTAATGACTT	8280
30	CATGATTCAA GGTGGCGATC CAACAGCTAC TGGTATGGGT GGCGAAAGTA TTTATGGCGG	8340
	TGCTTTTGAA GATGAATTTT CATTAAATGC ATTTAACTTA TATGGCGCAT TATCAATGGC	8400
	TAATCAGGA CCTAATACTA ATGGTTCACA ATTTTTCATT GTTCAAATGA AAGAAGTACC	8460
35	TCAAAATATG TTAAGTCAAC TTGCAGATGG TGGCTGGCCT CAACCAATCG TTGATGCATA	8520
	TGGCGAAAAG GGTGGTACAC CATGGTTAGA TCAAAAACAT ACAGTATTCT GTCAAATCAT	8580
40	TGATGGTGAA aCTACATTAG AAGATATTGC AAATACAAAA GTGGGACCAC AAGATAAACC	8640
	ACTTCATGAT GTTGTAATTG AATCTATTGA TGTTGAAGAA TAATATCTAA ACATAATTAA	8700
	CTACCAACAT TTAAACTCG GATAAAGCTA ATTTATGAAT GGATTAGTAT ATATTCCAAC	8760
45	gAAAAATAAT AAATAATAT GATGAGCAAT CTCAATATAT TTATCaAGAA AGCACAGTTT	8820
	TTAAATAGAT GTGTATTTTA AAGATAATAG TTGAGGTTGC TTTTATGTT TTTACAGAGA	8880
	ATTGCTATTC AAATAGTAAA TAAATTGAAA ACAAAGTAGC TGGATATCAT ATTGATTTAG	8940
50	ATAGGAATTT GTTGCTAATT TTATTTGTAA ATCCAAGTTT GTAGAATTCT TATTCATTTA	9000
	TAAATAATA TCGTATGAT TTGATTTTTT AATTAGTCCA CCATTCGAT TTGTGCTATG	9060

55

	AACATATCAA GGTGCGTGTA CTGGTATTCA ACCATACGGT GCGTTTGTTG AGACCCCTAA	9180
	TCATACTGAA GGACTGATTC ATATATCAGA AATTATGGAT GACTACGTTC ATAATTTGAA	9240
5	GAAATTTCTA TCAGAAGGCC AAATTGTTAA AGCTAAAATT TTGTCTATAG ATGATGAAGG	9300
	AAAGCTTAAT CTATCATTAA AGGATAATGA TTACTTCAAA AATTATGAGC GTAAGAAGGA	9360
	AAAACAATCA GTATTAGATG AAATCAGAGA AACAGAAAAA TATGGGTTTC AAACACTTAA	9420
10	AGAACGCTTA CCAATCTGGA TAAACAGTC AAAGCGAGCA ATTCGAAACG ACTAAAGGAA	9480
	CAGATAAATC GTACCGAAAA TCATACAAAG GGTCTGAAAT GAAAGTTTCT TAGACTATAA	9540
15	AAGAGATTAG TATCTATTAA ATTTTATTAG ATACTAATCT CTTTTTGTCT ACGATAACGT	9600
	AATATGaTTG ATTCTATTTA CACGTACAAA TGGTTTAAGG TGACATATCC ATTATCTTTG	9660
	TTAGATAGAA TCGTTGATTT GCaATATTGT ATGTGGATTT GTTTTTTTTA TTTATTTTAG	9720
20	AAATGAGAAC TACAACCTAA AGTATTAAAC GAATTGCAAC TATATAAACA GATAATTGGA	9780
	GAATGAAAAA ATTACATGTT ATAGTCAACT CAATAATTTT AAGGAGGAAT TAAGTAATGA	9840
	AAAGTAAATA CGAACCATTG TTTGATAAAG TAGAATTACC AAATGGAGTA GAGTTGAGAA	9900
25	ATCGATTTGT GTTAGCCCCCT TTAACACATA TTTCTTCAAA TGATGATGGT ACTATTTTCAG	9960
	ATGTAGAACT TCCTTATATT GAAAAGCGTT CACAAGATGT TGGTATTACA ATTAATGCTG	10020
	CGAGTAATGT GASTGATGTC GGAAAAGCAT TTCCAGGACA GCCATCAATC GCGCATGACA	10080
30	GTAATATTGA AGGACTAAAA CGATTAGCTA CAGCAATGAA GAAAAACGGT GCCAAAGCAC	10140
	TCGTACAAAT ACATCATGGC GGTGCACAAG CATTGCCTGA ATTAACACCT GATGGAGACG	10200
	TCGTAGCACC AAGTCCAATT TCTTTAAAAA GTTTTGGTCA GAAACAAGAA CATAGTGCTA	10260
35	GAGAAATGAC GAATGAAGAG ATTGAACAAG CAATCAAGGA TTTTGGTGAA GCAACGCGAC	10320
	GTGCaATTGA AGCAGGGTTT GATGGTGTG AAATACATGG CGCGAATCAT TACTTAATTC	10380
	ATCAATTTGT ATCACCATAC TATAATAGAA GAAATGATGT ATGGGCAAAT CAATATAAAT	10440
40	TCCCGGTCGC TGTGATTGAA GAAGTACTTA AAGCGAAAGA AGCGTATGGC AATAAAGACT	10500
	TTATAGTTGG ATACAGATTA TCTCCAGAGG AAGCGGAGTC TCCAGGAATC ACAATGGAAA	10560
45	TTACAGAGGA ACTCGTTAAT AAAATTAGCC ATATGCCAAT CGACTATATT CATGTTTCAA	10620
	TGATGGATAC GCATGCAACG ACACGTGAAG GTAAATACGC TGGACAAGAA AGACTGCCTT	10680
	TAATTCACAA ATGGATAAAT GGTCGTATGC CACTTATCGG TATTGGTTCA ATTTTCACAG	10740
50	CTGACGAAGC TTTAGATGCA GTTGAAAAATG TTGGTGTGTA CTTAGTAGCC ATTGGTAGAG	10800
	AGCTACTACT GGATTATCAA TTTGTTGAAA AAATTAAAGA TGGACGGGAA GATGAAATTA	10860

55

AATTTAATGA AGGGTTTTAT CCATTACCAC GTA

10953

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8155 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

15	TTTGATATAA AACTGAATTA ATTAAATGTA TCGATTCAAC CTAATGAAGT GAATTTACAA	60
	GTTAAAGTAG AGCCTTTTAT CAAAGAGTT AAAGTAAATG TTAAACAGAA AGGTAGTTTA	120
	GCAGATGATA AAGAGTTAAG TTCGATTGAT TTAGAAGATA AAGAAATTGA AATCTTCGGT	180
20	AGTCGAGATG ACTTACAAAA TATAAGCGAA GTTGATGCAG AAGTAGATTT AGATGGTATT	240
	TCAGAATCAA CTGAAAAGAC TGTAAAAATC AATTTWCCAG AACATGTCAC TAAAGCACAA	300
	CCAAGTGAAA CGAAGGCTTA TATAATGTA AAATAAATAG CTAAATTAAA GGAGAGTAAA	360
25	CAATGGGAAA ATATTTTGGT ACAGACGGAG TAAGAGGTGT CGCAAACCAA GAACTAACAC	420
	CTGAATTGGC ATTTAAATTA GGAAGATACG GTGGCTATGT TCTAGCAGAT AATAAAGGTG	480
	AAAAACACCC ACGTGTAATT GTAGGTCGCG ATACTAGAGT TTCAGGTGAA ATGTTAGAAT	540
30	CAGCATTAAAT AGCTGGTTTG ATTTCAATTG GTGCAGAAGT GATGCGATTA GGTATTATTT	600
	CAACACCAGG TGTTGCATAT TTAACACGCG ATATGGGTGC AGAGTTAGGT GTAATGATTT	660
	CAGCCTCTCA TAATCCAGTT GCAGATAATG GTATTAAATT CTTTGGATCA GATGGTTTTA	720
35	AACTATCAGA TGAACAAGAA AATGAAATTG AAGCATTATT GGATCAAGAA AACCCAGAAT	780
	TACCAAGACC AGTTGGCAAT GATATTGTAC ATTATTCAGA TTACTTTGAA GGGGCACAAA	840
40	AATATTTGAG CTATTTAAAA TCAACAGTAG ATGTTAACTT TGAAGGTTTG AAAATTGCTT	900
	TAGATGGTGC AAATGGTTCA ACATCATCAC TAGCGCCATT CTTATTTGGT GACTTAGAAG	960
	CAGATACTGA AACAAATTGA TGTAGTCCTG ATGGATATAA TATCAATGAG AAATGTGGCT	1020
45	CTACACATCC TGAAAAATTA GCTGAAAAAG TAGTTGAAAC TGAAAGTGAT TTTGGGTTAG	1080
	CATTTGACGG CGATGGAGAC AGAATCATAG CAGTAGATGA GAATGGTCAA ATCGTTGACG	1140
	GTGACCAAAT TATGTTTATT ATTGGTCAAG AAATGCATAA AAATCAAGAA TTGAATAATG	1200
50	ACATGATTGT TTCTACTGTT ATGAGTAATT TAGGTTTTTA CAAAGCGCTT GAACAAGAAG	1260
	GAATTAAATC TAATAAAACT AAAGTTGGCG ACAGATATGT AGTAGAAGAA ATGCGTCGCG	1320

	CTGGTGATGG TTTATTAACT GGTATTCAAT TAGCTTCTGT AATAAAAAATG ACTGGTAAAT	1440
	CACTAAGTGA ATTAGCTGGA CAAATGAAAA AATATCCACA ATCATTAAAT AACGTACGCG	1500
5	TAACAGATAA ATATCGTGTT GAAGAAAATG TTGACGTTAA AGAAGTTATG ACTAAAGTAG	1560
	AAGTAGAAAT GAATGGAGAA GGTGGAATTT TAGTAAGACC TTCTGGAACA aACCATTAGT	1620
	TCGTGTCATG GTTGAAGCAG CAACTGATGA AGATGCTGAA aGATTTGCAC AACAAATAGC	1680
10	TGATGTGGTT CAAGATAAAA TGGGATTAGA TAAATAAATA CTGTATTACA AATGAGCCGA	1740
	TGCGTATGcA nTcgtTTTTT GTGTTTGTAG AAATAATTTA TAGTACAAAC GTAAAATGAT	1800
	ATAAACAAAA TAAAAACAAA GTAATCAATA TGTAATATAA AATACACTGG TACTCAATAT	1860
15	ATAATGATGA TAAAATTAAT TTTAATTAGA TAGAGTTGCT TTGTGTTTTT AACGCAGATG	1920
	CTACTACTTA TCTTAACAGT TGATTAAGTG AAATCATTTA ACAGCGAGAA TAATCAACCA	1980
20	GGAGGATGAC TTAATGAATT TATTCAGACA ACAAAAATTT AGTATCAGAA AATTTAATGT	2040
	CGGTATTTTT TCAGCTTTAA TTGCCACTGT TACTTTTATA TCTACTAACC CGACAACAGC	2100
	GTCTGCAGCA GAGCAAAATC AGCCTGCACA AAATCAACCA GCACAACCAG CTGATGCCAA	2160
25	TACACAGCCT AACGCAAATG CTGGTGCTCA AGCTAATCCT ACAGCACAGC CAGCTGCACC	2220
	TGCCAACCAA GGACAACCAG CAGTACAACC AGCAAACCAA GGTGGACAGG CTAATCCAGC	2280
	AGGAGGAGCA GCACAACCAA ATACACAACC AGCTGGACAA GGTGATCAAG CTGATCCGAA	2340
30	TAACGCTGCA CAAGCACAAC CTGGAAATCA AGCAACACCG GCAAACCAAG CAGGTCAAGG	2400
	AAATAACCAA GCAACACCTA ATAATAATGC AACACCGGCA AATCAAACAC AGCCAGCGAA	2460
	TGCTCCAGCA GCAGCGCAAC CAGCAGCACC TGTAGCAGCA AACGCACAAA CTCAAGATCC	2520
35	AAATGCTAGC AATACTGGTG AAGGCAGTAT TAATACGACA TTAACATTTG ATGATCCTGC	2580
	CATATCAACA GATGAGAATA GACAGGATCC AACTGTAACT GTTACAGATA AAGTAAATGG	2640
	TTATTCAATTA ATTAACAACG GTAAGATTGG TTTCGTTAAC TCAGAATTAA GACGAAGCGA	2700
40	TATGTTTGAT AAGAATAACC CTCAAAATA TCAAGCTAAA GGAAACGTGG CTGCATTAGG	2760
	TCGTGTGAAT GCAAATGATT CTACAGATCA TGGTAACTTT AACGGTATTT CAAAACTGT	2820
45	AAATGTAAAA CCAGATTCAG AATTAATTAT TAACTTTACT ACTATGCAAA CGAATAGTAA	2880
	GCAAGGTGCA ACAAATTTAG TTATTAAAGA TGCTAAGAAA AATACTGAAT TAGCAACTGT	2940
	AAATGTTGCT AAGACTGGTA CTGCACATTT ATTTAAAGTA CCAACTGATG CTGATCGTTT	3000
50	AGATTTACAA TTTATTCCTG ACAATACAGC AGTTGCTGAT GCTTCAAGAA TTACAACAAA	3060
	TAAAGATGGT TATAAATACT ATTCATTCAT TGATAATGTA GGTCTATTCT CAGGATCACA	3120

55

	TAATACTGAA ATCGGTAACA ATGGTAATTT TGGTGCTTCA TTAAAAGCAG ATCAATTTAA	3240
	ATATGAAGTA ACATTACCAC AAGGTGTAAC TTACGTAAAT AATTCATTAA CTACAACATT	3300
5	CCCTAATGGT AATGAAGACA GTACAGTATT GAAAAATATG ACTGTTAATT ATGATCAAAA	3360
	TGCAAAATAAA GTTACATTTA CAAGCCAAGG TGTGACAACG GCACGTGGTA CACACACTAA	3420
	AGAAGTTTTA TTCCAGATA AATCTTTAAA ATTATCATAT AAAGTTAATG TTGCGAATAT	3480
10	CGATACACCT AAAAATATTG ATTTTAATGA AAAATTAACA TATCGTACTG CTTCAGATGT	3540
	TGTAATTAAT AATGCGCAAC CAGAAGTACA CTAACGTGAG ATCCATTTTC AGTAGCGGTT	3600
	GAAATGAACA AAGATGCGTT GCAACAACAA GTAAACTCAC AAGTTGATAA TAGTCATTAC	3660
15	ACAACAGCAT CAATTGCAGA ATACAATAAA CTAAACAAC AAGCAGATAC TATTTTAAAT	3720
	GAAGATGCGA ATCATGTTAA AACTGCAAAT CGTGCATCTC AAGCGGATAT TGATGGTTTTA	3780
20	GTAACTAAAT TACAAGCTGC ATTAATTGAT AATCAAGCAG CAATTGCTGA ATTAGATACT	3840
	AAAGCTCAAG AAAAGGTTAC AGCAGCACAA CAAAGTAAAA AAGTTACGCA AGATGAAGTT	3900
	GCAGCACTTG TAACTAAAAT TAACAATGAT AAAAAATATG CAATCGCAGA AATTAATAAA	3960
25	CAAACTACAG CACAAGGTGT CAACTGAA AAAGATAATG GTATCGCAGT GTTAGAACAA	4020
	GATGTGATTA CACCAACAGT TAAACCTCAA GCGAAACAAG ATATTATCCA AGCAGTTACA	4080
	ACTCGTAAAC AACAAATTAA AAAGTCAAAT GCATCATTAC AAGATGAAA AGATGTAGCA	4140
30	AATGATAAAA TTGGTAAAAT TGAACAAAG GCAATTAAAG ATATTGATGC AGCAACAACA	4200
	AATGCACAAG TAGAAGCCAT TAAAACAAAA GCAATCAATG ATATTAATCA AACTACACCT	4260
	GCTACAACAG CTAAAGCAGC AGCTCTTGAA GAATTTGACG AAGTTGTTCA AGCACAAATT	4320
35	GATCAAGCAC CTTTAAATCC TGATACAACA AATGAAGAAG TAGCGGAAGC TATTGAACGT	4380
	ATTAATGCAG CTAAAGTTTC TGGTGTTAAA GCAATTGAAG CGACAACGAC TGCACAAGAT	4440
	TTAGAAAGAG TTAAAAACGA AGAAATCTCA AAAATTGAAA ATATTACTGA CTCTACGCAA	4500
40	ACAAAAATGG ATGCCTATAA TGAAGTTAA CAAGCTGCAA CAGCTAGAAA AGCTCAAAAT	4560
	GCTACAGTTT CAAATGCAAC AAATGAAGAA GTAGCAGAAG CTGATGCAGC AGTAGATGCA	4620
45	GCTCAAAAGC AAGGTTTACA TGACATCCAA GTTGTTAAAT CAAAACAGGA AGTTGCTGAT	4680
	ACAAATCAA AAGTATTAGA TAAATCAAT GCAATTCAAA CACAAGCAA AGTTAAACCT	4740
	GCAGCTGATA CGGAAGTAGA AAACGCATAT AATACACGTA AACAAGAAAT TCAAAATAGC	4800
50	AATGCTTCAA CTACAGAAGA AAAACAAGCT GCATATACAG AATTAGATAC TAAAAAGCAA	4860
	GAAGCAAGAA CAAATCTTGA TGCTGCAAAT ACAAACAGTG ATGTAACAAC AGCTAAAGAC	4920

55

	GCGGAAATCG	CTCAAAAAGC	AAGTGAACGT	AAAACAGCAA	TTGAAGCAAT	GAATGATTCTG	5040
	ACTACTGAAG	AACAACAAGC	AGCGAAAGAC	AAAGTGGATC	AAGCAGTAGT	TACTGCAAAC	5100
5	GCTGATATAG	ATAATGCTGC	AGCAAACAAT	GATGTGGATA	ATGCAAAAAC	TACAAATGAA	5160
	GCTACAATCG	CAGCCATTAC	ACCTGATGCA	AATGTTAAAC	CAGCAGCAAA	ACAAGCAATT	5220
	GCAGATAAAG	TACAAGCTCA	AGAAACAGCA	ATTGATGGAA	ATAACGGCTC	AACAACTGAA	5280
10	GAAAAAGCAG	CTGCTAAACA	ACAAGTTCAA	ACTGAAAAAA	CAACAGCTGA	TGCCGCAATA	5340
	GATGCAGCAC	ATACAAATGC	GGAAGTTGAA	GCGGCTAAAA	AAGCAGCAAT	TGCTAAAATT	5400
	GAAGCGATTG	AGCCAGCAAC	AACAACATAA	GATAATGCCA	AAGAAGCAAT	TGCTACGAAA	5460
15	GCGAATGAAC	GTA AACAGC	AATCGCTCAA	ACGCAAGACA	TTACTGCTGA	AGAAATTGCA	5520
	GCGGCTAATG	CGGACGTAGA	TAATGCTGTG	ACACAAGCAA	ATAGCAACAT	TGAAGCTGCT	5580
	AATAGTCAAA	ATGATGTAGA	CCAAGCGAAA	ACGACAGGTG	AAAATAGTAT	TGATCAAGTA	5640
20	ACACCAACAG	TTAATAAAAA	AGCAACTGCA	CGTAATGAAA	TCACAGCAAT	TTTAAATAAC	5700
	AAATTGCAAG	AGATTCAAGC	TACGCCAGAT	GCAACAGATG	AAGAAAAACA	AGCAGCTGAT	5760
25	GCTGAAGCAA	ATACTGAAAA	TGGTAAAGCA	AATCAAGCCA	TTTCAGCAGC	AACTACTAAC	5820
	GCACAAGTTG	ATGAAGCTAA	AGCAAATGCA	GAAGCAGCGA	TTAATGCGGT	AACACCAAAA	5880
	GTTGTGAAGA	AACAAGCGGC	TAAAGATGAA	ATTGATCAAT	TACAAGCAAC	GCAAACAAAT	5940
30	GTTATCAATA	ATGATCAGAA	CGCTACAACA	GAAGAAAAAG	AAGCAGCTAT	TCAACAATTA	6000
	GCAACAGCAG	TTACAGACGC	GAAAAATAAT	ATTACAGCTG	CAACTGATGA	TAATGGTGTA	6060
	GATCAGGCGA	AAGACGCTGG	AAAGAATTCA	ATTCAAAGCA	CGCAACCAGC	AACAGCGGTT	6120
35	AAATCAAATG	CTAAAAATGA	TGTTGATCAA	GCTGTGACAA	CTCAAAATCA	AGCAATTGAT	6180
	AATAEAACTG	GTGCTACAAC	TGAAGAGAAA	AATGCAGCAA	AAGATTTAGT	TTTAAAAGCT	6240
	AAAGAAAAAG	CGTATCAAGA	TATCTTAAAT	GCACAAACAA	CTAATGATGT	TACGCAAATT	6300
40	AAAGATCAAG	CAGTTGCTGA	TATTCAAGGT	ATTACTGCAG	ATACAACAAT	TAAAGATGTT	6360
	GCGAAAGATG	AATTAGCAAC	AAAAGCAAAC	GAACAAAAAG	CGCTTATTGC	ACAAACTGCA	6420
45	GATGCGACTA	CTGAAGAAAA	AGAACAAGCA	AATCAACAAG	TAGACGCACA	ATTAACACAA	6480
	GGTAATCAAA	ATATTGAAAA	TGCACAGTCA	ATCGATGATG	TAAACACTGC	AAAAGATAAT	6540
	GCAATTCAAG	CAATTGACCC	AATTCAAGCA	TCAACAGATG	TTAAAACGAA	TGCAAGAGCG	6600
50	GAATTGCTAA	CTGAAATGCA	AAATAAAATA	ACTGAAATAC	TTAATAATAA	TGAGACTACT	6660
	AATGAAGAAA	AAGGTAACGA	TATTGGACCA	GTTAGAGCAG	CATATGAAGA	AGGTTTAAAT	6720

55

	AAAGTTCAAC AACTTCATGC AAATCCTGTT AAGAAACCAG CAGGTAAAAA AGAATTAGAT	6840
	CAAGCTGCAG CTGATAAGAA AACACAAATA GAACAAACAC CAAATGCATC ACAACAAGAA	6900
5	ATTAATGATG CAAAACAAGA AGTTGATACT GAATTAAATC AAGCGAAAAC AAATGTCGAT	6960
	CAATCATCAA CAAATGAATA TGTGATAAT GCAGTTAAAG AAGGAAAAGC TAAAATTAAT	7020
	GCAGTTAAAA CATTTAGTGA GTACAAAAAA GATGCTTTAG CTAAATTGA AGATGCATAT	7080
10	AATGCTAAAG TAAACGAAGC GGATAACTCT AACGCATCGA CTTCAAGTGA AATTGCTGAA	7140
	GCGAAACAAA AACTTGCTGA ATTAATAACAA ACTGCCGATC AAAATGTTAA TCAAGCTACT	7200
	TCTAAAGATG ACATTGAAGT TCAAATTCAT AATGACTTAG ATAATATTAA CGATTACACA	7260
15	ATTCCAACAG GTAAAAAAGA ATCAGCTACA ACAGATTTAT ATGCTTATGC AGATCAGAAG	7320
	AAAAATAATA TTTGAGCTGA CACTAATGCA ACACAAGATG AAAAGCAACA AGCAATTAAG	7380
	CAAGTTGACC AAAATGTTCA AACTGCATTA GAAAGCATT AATATGGTGT GGATAATGGT	7440
20	GACGTTGATG ATGCATTAA ACAGGTAAC GCAGCAATTG ATGCTATTCA AGTAGATGCT	7500
	ACTGTTAAAC CTAAAGCGAA CCAAGCTATT GAAGTTAAAG CAGAAGATAC GAAAGAATCT	7560
25	ATTGATCAAA GTGACCAAGT AACTGCTGAA GAAAAAAGT AAGCATTAGC AATGATTAAA	7620
	CAAATTACAG ATCAAGCTAA ACAAGGTATT ACTGATGCAA CAACAAGTGC TGAAGTTGAA	7680
	AAAGCGAAAG CTCaAGGACT TGAAGCATTT GATAACATTC AAATCGACTC AACAGAAAAA	7740
30	CAAAAAGCTA TCGAAGAATT AGAACTGCA CTAGACCAGA TTGAAGCAGG TGTAAATGTC	7800
	AACGCTGATG CTACAACTGA AGAAAAAGAA GCGTTTACGA ATGCTTTAGA AGACATTTTA	7860
	TCAAAAGCAA CTGaAGATAT TTCTGATCAA ACTACAAATG CAGAAATCGC TACTGTCAA	7920
35	AATAGTGCGC TTGAACAACT TAAAGCACA CGTATTAATC CTGAAGTTAA GAAAAATGCT	7980
	TTGGAAGCAA TCAGAGAAGT GGTTAACAAG CAAATAGGAA tAATTAAAAA TGCAGATGCA	8040
	GATGCATCGG CGGAAAGANA TTGCACGTAC GGGATTTAGG TAGATATTTT GGACCGATT	8100
40	GCTGGATAAA TTTAGGGTna AACCCCAACC AATGCCGAAG TTGCCTGAAT TACCA	8155

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1630 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

	CTGTTTTATT TGCAGCACCC ATACTGGAAA TCACTTTAAT CCCTCGGTCA AGACACTCTT	120
	TCATTAAGTG TACTTTGTAC ATTATTGTAT CACTTGCATC TACAAAATAA TCTATATCGT	180
5	AGTTATCGAA AATTCTCTCA TATGTCTCTT CTGTATAAAA CATATGTAAG GCGGTGACTT	240
	TACAATCTGG ATTAATTAAT TTAATACGTT CTTCCATCAA AGAAACTTTA CTTTGTCTTA	300
	CCGTTGTAGT TAAAGCGTGT AATTGTCTGT TTACATTGT AATATCAACA TCATCTTTAT	360
10	CTATTAATAT AATATGACCA ATATTCTGTC TTGCTAATGC TTCAGCAGCA AATGAACCAA	420
	CACCTCCAAC GCCAAGTATG ACAACAGTTT GTTGCTTCAA TAAATCTAAA CCTTGTGTGC	480
	CAATCGCTAG TTCATTTCTT GAAAATTGAT GTTTCATTAT TTTACCTCTT TCACTGATTT	540
15	ATACATAAGT ACATAGTAAC TTAAAATTTT ATATTTAGCA TTATCACTTT GATTATTTTC	600
	CCAAAATTCA ACGAGGAAAC ATTTATTAAA CGCTATAAAA CCCAACTAAT TCTTTATTAA	660
	AAACTTAAAG AAACGCATAA AAATACGCAA GACAAAGTCT TGCATATCGA TAGAGTCCGT	720
20	ATTGCCGTAG TTATAATAGC TTGATCATTC GGCCTGTTAT ATACAGGTGG GTGCCCTGTT	780
	TCTTGTTTTG TACGTCTTC ATATAAGGCG TGTACGCTGC AAGAAAACCC ATTGGGCTCC	840
25	CTTGATCAAA GAGTGTTAGG CCCAAATTAA AAAGCAAAC TACGAACAAC TCAGATGACT	900
	ATCTTATGAT GTTATATTAC CACATAATTA AAATTAATGA AATTATAACA AACCAAAGTT	960
	TATTGATTTT TAAAATTTA GTGACGAATT CGCAAAGAAA GTTCTTCTAA TTGTTTATCA	1020
30	GAAACTTCAC TAGGCGCATT CGTTAATAAA CATGTAGCAG ATGCTGTTTT AGGGAATGCG	1080
	ATTGTATCTC TCAAGTTTGT TCTATTAGTC AATAACATGA CTAATCGGTC TAATCCTAAT	1140
	GCAATACCGC CATGTGGTGG TGCACCATAT TTAATGCAT CTAGTaaGAA GCCGAACGTG	1200
35	TCCTgTGCTT GTTCTTTAGT AAATCCAAGA ACTTCGAACA TTTTCTTTG TAACTCACCA	1260
	TCATGAATTC TGATTGAACC GCCACCTAAT TCATAACCAT TTAATACTAT GTCATAAGCA	1320
	TTTGCCTCAG CTTCTtCTGG CGCAGTGCCA AGCTTAGCAA TATCAGCTTC TTTTGGAGAT	1380
40	GTAAATGGAT GATGTGCTGC AACGTAACGT TTCGCATCTT CATCATATTC TAATAATGGC	1440
	CAATCTGTCA CCCATAAGAA GTTTAATTTT GTTTCATCGA TTAAACCTAA TTCTTTAGCT	1500
45	AATTTGACAC GTAATGCACC TAACTTTGT GCAACGACAT TTGGTtctGTC TGCAACAAAC	1560
	ATTACTAAGT CACCAGCTTC AGCACCAGTT AATGTAAGTA ATGTTTCAAC ATTTTCTGTT	1620
	CAAAGAAACG	1630

50 (2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 732 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

	CAATTGGACA TCTTGTATGA AAAGGACAAC CTTGCGGCGG ATTACTTGGC GAAGGTAATT	60
10	CTCCTTTTAA TATAATTCTA TTGTTATTAT GTTTATCAAT TTGTGGTATT GATGAAATCA	120
	ACGCTTTTGT ATATGGATGT TTGGGATTTT CATAAATTC TTTATCAGAT GCGATTTCAA	180
	CTATATGACC TAAATACATA ACTCCAATGA CATCACTTAT ATGTTTTACT ACACTTAAAT	240
15	CATGTGCGAT AAATAAATAG CTTAAGTTAA ATTGTTCTTG TAAATCTTTT AATAAATTCA	300
	GTACTTGAGA TTGAACAGAT ACATCTAATG CACTTACAGG CTCATCAGCA ACAATTAAAC	360
	TGGACGCAA AGCCAATGCT CTTGCAATTC CCACTCTTTG TCTCTGTCCA CCTGAAAATT	420
20	CATGTGCATA TTATAATAT GCATCTTCAC TTAGGCCAAC ACATTTTAAT AAATATAGTA	480
	CTTCTTTTTT TATTTCTTCT TTTGGCAATT TTTTATAATT TAAAATAGGT TCTGAAATGA	540
	TATCTCCAAC CATTTGCATC GGATTCAATG ATGCATACGG ATCTTGAAAT ATCATCTGAT	600
25	ATTGTTGTCG TGATTTTCTG AGTTTTTTAC CTTGTAATCT TGTATATCT TCACCATTAA	660
	CAATTATTGA GCCTGAAGTT GCATCTTCAA GCCTGATAAT CACTTTACCT AACGTTGACT	720
30	TACCACAACC CG	732

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

35	(A) LENGTH: 5838 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

	AATATATTCA TATGTTTCAT CAACAATATT AGCTGCTTTT TGAATTAAAG CAATTTCGTC	60
45	AGCATCTTTG ACGTCTCTAA TTTTATCTAC AGTATTAGAA ATGCTTATTA ATGATATACG	120
	GCTTTTATTT AATTCAAGGT ATGTATCATA ACTTACATGA TGCCCCCTCAA AACCTACATT	180
	TTCAAAATTT TCTTGGTGTA GCAATTCTTT AATCTCACCA ATAATAGTAG ATTTACGATT	240
50	AATAATTTC AATTGTCG CTTGCTTAGT TGCTTGATCA ATATATCTAA AGTCTGTTAT	300
	CAAATATTGT TTATCTTTAG ATATGATAAG TGCTCCACTG GTACCAGTAA AACCTGATAA	360
	ATATCTTCTA TTGTAATCCG AAAGAATGAT AATCGCATCT AAATGTTTTT GTTCTAAAT	420

55

	CAACTTTATA CATTAAAATA ATATCATAAT AAGGATAAAA AATAATAGAT ATTGATTTTA	540
	GGGAGATAGT AATGAAAAAA TTGGTTTCAA TTGTTGGCGC AACATTATTG TTAGCTGGAT	600
5	GTGGATCACA AAATTTAGCA CCATTAGAAG AAAAAACAAC AGATTTAAGA GAAGATAATC	660
	ATCAACTCAA ACTAGATATT CAAGAACTTA ATCAACAAAT TAGTGATTCT AAATCTAAAA	720
	TTAAAGGGCT TGAAGAGGAT AAAGAAAACA GTAAAAAAC TGCATCTAAT AATACGAAAA	780
10	TTAAATTGAT GAATGTTACA TCAACATACT ACGACAAAGT TGCTAAAGCT TTGAAATCCT	840
	ATAACGATAT TGAGAAAGAT GTAAGTAAAA ACAAGGCGA TAAGAATGTT CAATCGAAAT	900
15	TAAATCAAAT TTCTAATGAT ATTCAAAGTG CTCACACTTC ATACAAAGAT GCTATCGATG	960
	GTTTATCACT TAGTGATGAT GATAAAAAAA CGTCTAAAAA TATCGATAAA TTAAACTCTG	1020
	ATTGAATCA TGCATTTGAT GATATTAAAA ATGGCTATCA AAATAAGAT AAAAAACAAC	1080
20	TTACAAAAGG ACAACAAGCG TTGTCAAAAT TAAACTTAAA TGCAAAATCA TGATAGGAGT	1140
	CTTTTAATGC GTAATATAAT ATTTTATCTT GTACTTATTA TTGCTGCGAT TGGATTAGTA	1200
	ATGAATCTAG ATGCCTTTAT TTTTCAATC GTCAGAATGT TAATCAGCTT TGcGTAAaTAG	1260
25	CTGGTATTAT TTATCTGATT TATTATTTCT TCATCTTAAC TGAAGACCAA CGCAAATATC	1320
	GCAAAGCAAT GCgTrAaGTA TAAAAGAAAT CAAAGAAGAA AATAGATAAA AAAACGGAAG	1380
	CACTTGTAGG TAAAATAGTC TACGTGCTTC CATTTTTTAT TCTAAAACT ACTTTCTAAA	1440
30	CATCCATTCA TCTGAACGAT ATTTTTCAGT TAATTCTTCC ACTTCTGCCA ATTGAGCTTC	1500
	TGtTAATTCA AGTGGCTTTA ATTCTATATT TAAACCTTTC TTAAACCTT TCTCGAAAGC	1560
	TTCTTCCATT TGAATAATAG TAATGTGTTT ATCTGAAATA TCATTGATGG CAACTGCTTT	1620
35	TTCAACGAAT GCCTCTTTCA TTTTAAATTT TAATCTTTCA TTTTATAAA TrAACATATC	1680
	AAACaGTTCA TCAATATCAA TATCTTGTA AATCGAACCG TGTTGGAGGA TTACGCCCTT	1740
40	TTGTCTCGTT TGAGCACTCC CAGCAATCTT ACGGCCTTCA ACAACTAGCT CATACCAACT	1800
	TGGTGCATCA AAACACACTG AACTTCGAGG TTGTTTTAAT TTTTGACGCT CTTCAGGCGT	1860
	TTTAGGTACC GCAAAATAAG TATCAAATCC TAAGTTTTTA AATCCTTCTA ATAATCCTTG	1920
45	TGAAATCACT CTGTACGCTT CTGTAAGTGT AGAAGGCATA TTCGGATGCG ATTACAGGCAC	1980
	AATCACACTG TAAGTTAACT CTTTATCATG TAGCACCCCA CGGCCACCAG TTTGACGCCT	2040
	TACGAGACCA AAACCTTTCT CTTTAACTT ATCAATATCA ATTTCTTTT GTAGCCTTG	2100
50	GAAATACCCT ATTGATAATG TTGCAGGATT CCATGTGTAA AAACGTATAA CTGGATCAAT	2160
	TTCACCTCTA GAGACAAAAT TTAATAACGC TTCATCCATT GCCATATTAT AATATGGGTC	2220
55		

	AAATGTATAA	TATTTGATTC	GCTAATTAAT	CAATTTAACT	AAATGAATAA	TAATTGCAAT	2340
	TCTTTAGTGA	AATATTTTGA	TAATTTGACC	TAACAGTCTT	ATAATTATAT	TATCGTTTAA	2400
5	TTAGGGAGGA	TGCAAGATGA	GTGCTAGTTT	GTACATCGCA	ATAATTTTGA	TTATAGCAAT	2460
	TATTGCTTAT	ATGATTGTTT	AACAAATTCT	TAACAAGCGA	GCTGTTAAAG	AATTAGATCA	2520
	AAATGAATTC	CATAATGGGA	TTAGAAAAGC	TCAAGTCATC	GATGTTAGAG	AGAAAGTTGA	2580
10	CTATGACTAC	GGTCACATTA	ATGGGTCTCG	CAATATTCCT	ATGACAATGT	TCAGGCAACG	2640
	ATTCCAAGGA	TTAAGAAAAG	ATCAACCGGT	ATACTTATGT	GATGCCAATG	GGATTGCTAG	2700
15	CTATAGAGCC	GCTCGTATTT	TGAAAAAGAA	TGGATATACA	GATATCTATA	TGTTAAAAGG	2760
	CGGCTATAAA	AAATGGACTG	GAAAAATAAA	GTCTAAAAAA	TAGTTTTTGT	AAATTTAATA	2820
	TACGATTTAA	TAAAATCTGA	GTGTTAATTG	ATCATCAATA	ACAATACTCA	GATTTTAAAT	2880
20	TTTTAACAAA	GTCTGTTACT	ATATTTCTCT	AGCTTCACCT	ATCATTAAAC	TTAGTTTCAG	2940
	CATAATAAAG	AAAGTTCAGC	TCATTTTCAA	TACGATTCAA	TTACCGCAAT	CTAAAAAATG	3000
	AAAAGACAAT	TTCTATGAAA	GAATAATACC	AAACCCTAAG	AGTTATTACT	TCGGTTTAGT	3060
25	TTTCTTGTTT	AAATAGAAAT	TGTCTTTTTT	AATTGATTTT	GAAACCATTA	TCCTTAAATC	3120
	TTCATACAAA	GTTAGAATAA	TAATTCTCGG	AATATGTGTT	TAATACTTTA	TTTTTCCTGT	3180
	TTAAGATTTT	CAAACTTTAA	TATTGGTTTA	CGAGCAGCTG	TAGCTTCGTC	TAATCGATCA	3240
30	ATCACAGTTG	TATGTGGTGC	TTCTAGCacT	TTATCAGGAT	CATTTTTAGC	TTCTTCAGCA	3300
	ATACTAATTA	ATGTATCGAT	AAAATAATCA	AGTGTTTCTT	TAGACTCTGT	CTCAGTCGGT	3360
35	TCAATCATCA	TACCTTCTTC	AACATTTAAT	GGGAAGTATA	TTGTTGGTGG	ATGTACACCG	3420
	AAATCTAATA	ATCGCTTAGC	CATGTCTAAA	GTACGTACAC	CAAATTCTTT	TTGACGCACA	3480
	CCACTTAACA	CAAACTCGTG	TTTACAATAT	TGTTTATAAG	GTATTTCAAA	GTGTTTAGAT	3540
40	AAACGTGCTT	TAATATAATT	CGCATTAAAG	ACCGCTGCTT	CAGAAACCTC	TTTAAGTCCA	3600
	GTTGCTCCCA	TAGTTCGAAT	ATACGTATAA	GCTCTTAAGT	AAATACCAAA	GTTACCATAA	3660
	AATGGTTTTA	CACGTCCGAT	AGAATTTTTA	ATGTCATTAT	CATATTTAAA	TTGTGCGCCA	3720
45	TCTTTAATAA	CCATTGGCTT	TGTAAGTAA	CTTGCTAGTT	CTTTTACTAC	ACCGACTGGA	3780
	CCTGAACCAG	GACCGCCACC	ACCATGTGGA	CCAGTAAATG	TTTTATGCAA	GTTTAAATGA	3840
	ACAGCATCAA	ATCCCATATC	TCCTGGGCGA	ACTTTGTCCA	TAATAGCGTT	TAAATTCGCA	3900
50	CCATCATAAT	ATAATAGACC	ACCAGCATT	TGGACGATTT	CACGGATTTC	CATAATATTT	3960
	TTTTCGAAAA	TACCTAAAGT	GTTTGGATTA	GTTAACATAA	TAGCTGCTGT	ATTTTCATTT	4020

55

	GATTTAAATC CTGCAAATGa AGCTGAGGCT GGaTTTCGTAC CATGCGCAGA ATCTGGcACA	4140
	ATGACTTCAT CACGATGACC TTCACCATTa TTCTCATGGT AAGCTTTAAA TATCATCAAT	4200
5	GCAGTCCATT CACCATGTGC GCCAGCAGCT GGTGTGAATG TCACCTCATC CATACCAGTA	4260
	ATTCTCTTTA ATTCTTCTTG CAAACTATAA ATAATTTCTA ATGAACCTTG AACTTGATCT	4320
	TCATCTTGTA ATGGATGTGA TTCACTAAAT CCTGGTATTC TAGCAACCTT TTCATTAATT	4380
10	TTAGGGTTAT ACTTCATCGT ACATGAACCC AATGGATAAA ATCCGTTGTC TACACCGAAA	4440
	TTTTTATTTG AAAGTTcAGT ATAATGACGT ACTAAGTCTA GTTCAGCAAC TTCAGGAAAC	4500
	TCCGCTTTGT TTTTACGAAT AAATTTATCA TCTAACAATG ACTCAACAGA ATTTGTTTTA	4560
15	ATATCACTTT TTGGTAATGA ATATGCATAT CTGCCTTCAC GAGATCTTTC AAAAATTAAT	4620
	GGACTTGATT TACTAGTCAT TTAACTCACC AGCCTTTTCT ACAAATGTAT CGATTTTCATC	4680
20	TTTTGTTCTT AATTCAGTTA CAGCTATTAA CATGTGATTT TTAAAGTCGT CTGAAACAAC	4740
	ACCTAAATCA AAACCACCGA TAATATTGTA CTTCACTAAT TCCTCGTTAA CTTGTTGAAT	4800
	TGGTTTGTCa AATTTGACTa CAAACTCATT GmnaAGnTGT ACCATCTAAT ACTTCAAAAC	4860
25	CTTTTTTAAT AAATTGTTGT TTAGCATAGT TAGCATGTTT TATATTTTGA ACTGCAATAT	4920
	CATAGATACC TTGTTTACCA AGTGCTGACA TTGCAATTGA TGaCGcTAAa GCATTTAATG	4980
	CTTGGTTAGA ACAAATATTA GATGTCGCTT TATCGCGTCG AATATGTTGT TCACGTGCTT	5040
30	GTAATGTTAA TACAAAGCCA CGATTACCTT CATCATCTTG TGTTTGACCG ACTAATCTAC	5100
	CTGGCACTTT ACGCATTAAc TTTTTCGTCG TTGCAAAATA TCCACAATGT GGCCCACCGA	5160
	ATTGAGCAGG AATTCCGAAT GGCTGAGTAT CACCTACAAC AATATCTGCA CCAAATGAAC	5220
35	CTGGAGGTGT AAGTAATCCC AATGCTAATG GATTTGCATA TACGATAAAT AATGCTTTTT	5280
	TATCTTCAAT AAAGCTATGA ATCTTTTCAA GATCTTCAAT TGAACCGTAA AAGTTTGGAT	5340
40	ATTGTACTGC AACAGCTGCT GTTTCATCAT CCACTGCTGC TTCTAATTTT TTCAAATCTG	5400
	TAACAGTGCC ATCTAAATCG ATTTCCACTA CTTCGAATTC CTTACGCGTC TTAGCATAAG	5460
	TATGAAGTAC TTGTAATGCT TGATAATGTA AACCTTTTGA GACTACAATT TTATTTTCT	5520
45	TTGTTTGAAT AAATGCTAAG ATACATGCTT CAGCAAAGCT AGTCATCCCA TCATACATAG	5580
	AAGAATTTGC TACATCCATA TCTGTTAATT CACAAATTAA AGTTTGAAC TCAAAAATGG	5640
	CTTGTAATTC ACCTTGAGAA ATTTCCGGTT GATATGGCGT ATATGCTGTG TAAAATTCTG	5700
50	ATCTTGAAAT CATAGCATCC ACAACTGATG GCGCGTAATG ATCATAAACA CCAGCACCCA	5760
	TAAATGATGT ATGCGTTTCT TTAGTGATAT tCTTGCTkGC AATGGGGATT TAAACnTCTA	5820

55

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18355 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

5	ATnATAATTG GCTTTGCTAA TAATTACTTC CCTGAATTAC aAGTATTAGC AAACGAAATA	60
10	AAATCTGATA TGGCTAGTTC ATTAAACAA TGATATTTT ATTAAATTT TTAAGCTTT	120
15	GTACGAAATT GTACAAAGCT TTTTGGTGC GTATTGTATG GGCAACAACT TGACGATGAA	180
20	AATCCGTAC AGGATTGGTA ATAGGAAATG TTAGCGAAAG ACAAGGGTAT CCATTGTAGA	240
25	TTAACAAAAG GACGTTTCCA CAAGTGTGGG TTATTCTCAC TAAAGCAATA CGCAGAGACA	300
30	ACTTACGTAA AATTTTGAAC TGA CTAGAAC GGAAC TTCTA CTCAATTATT GATAAAAATT	360
35	TTCAAAAAGA CTTGAATGTG CTGAGAATAC GAAGTTTATG GAAGGATTAT CAAAATATAA	420
40	ATGTGCATTC ATTTACAACC TTTATTGACA ATGATTCTCA ACTAATATAG TATATAATCA	480
45	AATCGTAATA GTTACGATTT GTTTCTGCA ACTTTTTTGA AGTTTTAGTT GAGGTGAAAA	540
50	CAATAAAAGC ATCTAAGTGA ATGTAGTTAA CGGACAAC TG CATTGCTTG TAGAGCCACA	600
55	AGAAGCAACT TTAAATAAGG TTTACGGTTG CATTTTGATA CAACAACCGA TTA CTAAAGTC	660
60	ATGCTTTCCA CTTTGCGGGT TAGCATGACT TACCTAATAG ATAGAGCTAT TAGGTT CAGC	720
65	TTCTAAAAAA TTACAGTTTT AGAGGAATAC AGTTGCTTGc tTCGCAACAA CTGCATAAGA	780
70	GCCATGGTTT TCGCTTTTGC GAATTAGCAT GACTTACCTA CTAGATAGAG CTATTAGGTT	840
75	CATCTTCTAA AAAATTACAG GTTTAGAGGA ATACAGTTGT TTGCTTCGCA ACAACTGCAT	900
80	AAGAGCCTCT AGTAATTAAA ATTACAGAGG CTCTAAAAAT ACATCTAAAG GAGTGTTCGTA	960
85	TGAATCGGCA GGTATAGAA TTTTCTAAGT ATAATCCTTC GGGGAATATG ACGATACTTG	1020
90	TTCAATCAAA ACATGATGCT AGTGAATATG CATCTATCGC CAATCAGTTG ATGGCCGCAA	1080
95	CACATGTATG CTGTGAACAG GTAGGCTTTA TAGrATCAAC ACAAATGAT GATGGTAATG	1140
100	ATTTTCACTT AGTTATGAGC GGTAATGAAT TTTGCGGTAA TGCGACGATG TCATATATAC	1200
105	ATCATTTGCA GGAAAGTCAT TTGCTTAAAG ACCAACAGTT TAAGGTGAAG GTGTCTGGCT	1260
110	GTTGCGATTT AGTGCAATGC GCAATTCATG ATTGCCAATA CTATGAAGTT CAAATGCCAC	1320
115	AAGCCCATCG TGTGTGCCA ACAACAATTA ATATGGGTAA TCATTCATGG AAAGCAATAG	1380

	TTCAACATTT GGTGAAGCG TTTGTGCGTG AgcAACAAATG GAGTCACAAA TATAAAACAG	1500
	TAGGTATGAT GCTTTTTGAT GAACAACGTC AATTTTTTACA GCCATTAATC TATATACCAG	1560
5	AAATTCAAAG TTTAATTTGG GAAAATAGCT GTGGTTCTGG TACAgCATCA ATTGGGGTTT	1620
	TTAATAATTA TCAACGTAAT GACGCATGCA AAGATTTTAC AGTACATCAG CCAGGGGGCA	1680
	GTATTTTAGT GACATCAAAG CGATGTCATC AATTGGGATA TCAAACCTCA ATTAAAGGAC	1740
10	AGGTTACAAC TGTAGCTACA GGaAAAGCAT ATATAGAATA AGGAGCCTAC AATGAATAAC	1800
	TTTAATAATG AAATCAAATT GATATTACAA CAATATTTAG AAAAGTTTGA AGCGCATTAC	1860
	GAGCGTGTAT TACAAGACGA TCAATATATC GAAGCATTAG AAACATTGAT GGATGACTAT	1920
15	AGTGAATTTA TTTTAAATCC TATTTATGAA CAACAATTTA ATGCTTGGCG TGACGTTGAA	1980
	GAAAAAGCAC AATTaATAAA ATCACTGCAA TATATTACAG CGCAGTGTGT TAAACAAGTG	2040
20	GAAGTCATTA GAGCGAGACG TCTATTAGAC GGACAGGCGT CTACCACAGG TTACTTTGAC	2100
	AATATAGAAC ATTGTATTGA TGAAGAGTTT GGACAATGTA GTATAGCTAG CAATGACAAA	2160
	TTATTGTTAG TTGGTTCAGG TGCATATCCA ATGACGTTAA TTCAAGTAGC AAAAGAAACA	2220
25	GGTGCTTCAG TTATCGGTAT TGATATTGAT CCACAAGCCG TTGACCTAGG GCGCAGAATC	2280
	GTTAACGTCT TAGCACCAA TGAAGATATA ACAATTACGG ATCAAAAGGT ATCTGAACCT	2340
	AAAGATATCA AAGATGTGAC GCATATCATA TTCAGCTCGA CAATTCCTTT AAAGTACAGC	2400
30	ATTTTAGAAG AATTATATGA TTTAACAAAT GAAAATGTCG TAGTTGCAAT GCGCTTTGGT	2460
	GATGGCATCA AAGCAATATT TAATTATCCG TCACAAGAAA CAGCGGAAGA TAAGTGGCAA	2520
	TGTGTGAATA AACATATGAG ACCACAGCAA ATTTTGTGATA TAGCACTTTA TAAAAAGCA	2580
35	GCTATAAAGG TAGGTATTAC GGATGTCTAA ATTATTAATG ATAGGCACTG GTCCgGTCCG	2640
	AATGCAATTA GCGAATATTT GCTATTTAAA ATCAGATTAT GAGATTGATA TGGTTGGACG	2700
40	TGCCTCAACA TCAGAAAAAT CAAAACGCTT ATATCAAGCG TATAAAAAAG AGAAACAATT	2760
	TGAAGTCAAA ATACAAAACG AGGCGCATCA ACATCTGGAA GGTAAGTTTG AAATTAATCG	2820
	TTTGTATAAA GATGTTAAAA ACGTTAAGGG TGAATACGAA ACGGTTGTCA TGGCATGCAC	2880
45	AGCAGATGCT TATTATGACA CACTACAGCA ATTGTCGTTA GAACTTTGC AAAGTGTCAA	2940
	ACATGTCATT TTAATATCAC CGACATTTGG TTCGCAAATG ATTGTCGAAC AATTTATGTC	3000
	TAAATTTAAT AAAGATATCG AAGTGATTTT ATTCTCAACT TATCTTGGCG ATACACGTAT	3060
50	TGTTGATAAA GAAGCGCCTA ATCATGTGTT GACAAACAGGT GTAAAAAAGA AATTGTACAT	3120
	GGGATCGACA CATTCAAACCT CAACAATGTG TCAACGAATC TCTGCTTTAG CTGAGCAATT	3180
55		

	TTATGTGCAC CCACCACTAT TTATGAATGA CTTTTCATTG AAAGCCATTT TCGAAGGAAC	3300
	AGATGTACCG GTTTATGTGT ATAAGTTATT TCCTGAAGGA CCGATAACGA TGACACTAAT	3360
5	CCGTGAAATG CGTTTAAATGT GGAAGGAAAT GATGGTTATT TTACAAGCAT TTAGAGTGCC	3420
	GTCAGTCAAC CTGCTTCAAT TTATGGTGAA GGAAAATTAT CCAGTACGTC CTGAAACTTT	3480
	GGATGAAGGT GATATTGAGC ATTTTCGAAAT CTTGCCAGAT ATCTTACAAG AATATCTGCT	3540
10	TTATGTAAGA TATACCGCAA TCCTCATTGA TCCATTTTCA CAGCCAGACG AAAACGGACA	3600
	TTACTTTGAT TTTTCAGCTG TACCATTTAA GCAAGTCTAT AAAAATGAAC AGGATGTTGT	3660
	TCAAATCCA AGAATGCCAA GTGAAGATTA TTACAGAACG GCGATGATTC AGCATATTGG	3720
15	GAAAATGCTA GGTATCAAAA CGCCAATGAT TGATCAGTTC CTAACGCT ATGAAGCAAG	3780
	TTGCCAGGCG TACAAGGATA TGCATCAAGA TCAACACTTA TCTTCTCAAT TTAATACAAA	3840
20	TCTATTTGAA GGAGATAAAG CACTCGTCAC AAAATTTTTG GAAATCAATA GAACGCTTTC	3900
	ATAATAAGGG TTTGAAGTTT TATAATAGAA AAAAATTATT GAATTATGTT TGACATTTAC	3960
	ATAAAAATAA GCAAATAATT GAGAAAAATA ATCATTACGA TTTGATTAAG TAATGCAACT	4020
25	TATCAATTTA GAAAGAGGAA AAGCAAATGA GAAACTAAC TAAATGAGT GCAATGTTAC	4080
	TTGCATCAGG GCTAATTTTA ACTGGTTGTG GCGGTAATAA AGGTTTAGAG GAGAAAAAAG	4140
	AAAACAAGCA ATTAACGTAT ACGACGGTTA AAGATATCGG TGATATGAAT CCGCATGTTT	4200
30	ACGGTGGATC AATGTCTGCT GAAAGTATGA TATACGAGCC GCTTGACGT AACACGAAAG	4260
	ATGGTATTAA GCCTTTACTA GCTAAAAAGT GGGATGTGTC TGAAGATGGG AAGACATACA	4320
	CGTTCCATTT GAGAGATGAC GTTAAATTCC ATGATGGTAC GCCATTTGca TGctGACGCA	4380
35	GTTAAGAAAA ATATTGACGC AgTTCAAGAA AACAAAAAAT TGCATTCTTG GTTAAAGATT	4440
	TCGACATTAA TTGACAATGT TAAAGTTAAA GATAAGTACA CGGTTGAATT GAATTTGAAA	4500
40	GAAGCATATC AACCTGCATT GGCTGAATTA GCGATGCCTC GTCCATATGT ATTTGTGTCT	4560
	CCAAAAGACT TTaAAAACGG TACAaCAAAA GATGGCGTTA AAAAGTTCGA TGGTACTGGT	4620
	CCATTTAAAT TAGGTGAACA CAAAAAAGAT GAGTCTGCAG ACTTTAACAA AAATGATCAA	4680
45	TACTGGGGCG AAAAGTCTAA ACTTAACAAA GTACAAGCAA AAGTAATGCC TGCTGGTGAA	4740
	ACAGCATTCC TATCAATGAA AAAAGGTGAA ACGAACTTTG CCTTCACAGA TGATAGAGGT	4800
	ACAGATAGCT TAGACAAAGA CTCTTTAAAA CAATTGAAAG ATACAGGTGA CTATCAAGTT	4860
50	AAGCGTAGTC AACCTATGAA TACGAAAATG TTAGTTGTCA ATTCTGGTAA AAAAGATAAC	4920
	GCTGTGAGTG ACAAACAGT CAGACAAGCG ATTGGTCATA TGGTAAACAG AGATAAAATT	4980

55

	ACAGACATTA ATTTGATAT GCCAACACGT AAGTATGACC TTAaaaaAGC AGAATCATT	5100
	TTAGATGAAG CTGGTTGGAA GAAAGGTAAA GACAGCGATG TTCGTCAAAA AGATGGTAAA	5160
5	AACCTTGAAA TGGCAATGTA CTATGACAAA GGTTCCTCAA GTCAAAAAGA ACAAGCAGAA	5220
	TACTTACAAG CAGAATTTAA GAAAATGGGT ATTAAGTTAA ACATCAATGG CGAAACATCA	5280
	GATAAAATTG CTGAACGTCG TACTTCTGGT GATTATGACT TAATGTTCAA CCAAACCTGG	5340
10	GGATTATTGT ACGATCCACA AAGTACTATT GCAGCATTTA AAGAGAAAAA TGGTTATGAA	5400
	AGTGCAACAT CAGGCATTGA GAACAAAGAT AAAATATACA ACAGCATTGA TGACGCATTT	5460
15	AAAATCCAAA ACGGTAAAGA GCGTTCAGAC GCTTATAAAA ACATTTTGAA ACAAATTGAT	5520
	GATGAAGGTA TCTTTATCCC TATTTACAC GGTAGTATGA CAGTTGTTGC ACCaAAAGAT	5580
	TTAGAAAAAG TATCATTAC ACAATCACAG TATGAATTAC CATTCAATGA AATGCAGTAT	5640
20	AAATAAAGGA GCAATTAGAT GTTCAAATTT ATCTTAAAC GTATTGCGCT CATGTTTCCA	5700
	TTGATGATTG TAGTAAGTTT TATGACATTT CTATTGACGT ATATTACAAA TGAAAATCCA	5760
	GCTGTGACAA TTTTACATGC ACAAGGGACG CCAAATGTAA CACCAGAGTT GATTGCAGAA	5820
25	ACGAATGAGA AGTACGGTTT CAATGATCCA TTATTAATTC AATATAAAAA TTGGTTACTT	5880
	GAAGCGATGC AATTTAATTT TGGTACAAGC TACATTACAG GTGACCCAGT TGCTGAACGT	5940
	ATTGGTCCAG CATTTATGAA TACATTGAAA TTAACAATAA TTTCAAGTGT TATGGTGATG	6000
30	ATTACATCAA TTATTTTAGG TGTAGTTAGT GCATTAAAAA GAGGAAAGTT CACTGATCGT	6060
	GCGATACGTT CAGTGGCTTT CTTTCTAACT GCATTACCAT CATATTGGAT AGCTTCAATA	6120
	CTTATTATTT ACGTTTCAGT GAAGTTAAAC ATATTGCCGA CTTCTGGATT AACAGGTCCA	6180
35	GAAAGTTACA TATTGCCAGT GATCGTTATT ACGATTGCCT ATGCTGGTAT TTAATTTAGA	6240
	AATGTTAGAC GCTCGATGGT GGAACAATTA AATGAAGATT ATGTACTTTA TTAAAGAGCA	6300
40	AGCGGTGTGA AATCTATCAC ATTAATGTTG CATGTGTTGC GTAATGCTTT ACAAGTTGCG	6360
	GTATCAATCT TTTGTATGTC TATACCAATG ATAATGGGTG GACTAGTTGT TATCGAGTAT	6420
	ATCTTTGCAT GGCCTGGACT AGGTCAATTA AGTTTAAAG CAATACTTGA ACACGATTTT	6480
45	CCAGTCATTC AAGCATATGT ATTAATTGTA GCGGTATTAT TTATTGTATT TAATACATTA	6540
	GCAGATATCA TTAATGCGCT ATTAAATCCA AGATTAAGGG aGGGCGCACG ATGATAATTT	6600
	TAAAmCGATT ATTmCArgWT AAAGGTGCAG TAATTGCTTT AGGCATTATT GTATTATATG	6660
50	TCTTTTtagg ATTAGCAGCA CCACTTGTGA CATTTTATGA TCCTAACCAT ATCGATACAG	6720
	CAAACAAATT TGCTGGCATG AGTTTCAAC ATCTACTAGG TACTGACCAT TTAGGTAGAG	6780

55

TATTTGTTTC TGTACTTATT GGATCTATTT TAGGATTCTT ATCAGGATAT TTCCAAGGGT 6900
 TTGTTGACGC CTTAATCATG CGTGCCTGTG ATGTTATGTT GGCATTCCCA AGTTATGTTG 6960
 5 TAACGTTAGC ATTAATTGCA TTGTTTGGAA TGGGTGCCGA AAATATTATC ATGGCATTTA 7020
 TTTTGACGCG TTGGGCATGG TTCTGTCGTG TTATACGTAC AAGTGTTATG CAGTACACTG 7080
 CTTCTGACCA TGTAAGATTT GCTAAAACAA TCGGTATGAA TGATATGAAA ATTATTACAA 7140
 10 AACATATTAT GCCATTAACA TTAGCAGATA TTGCTATCAT CTCTAGTAGC TCGATGTGTT 7200
 CAATGATCTT GCAAATATCT GGCTTTTCAT TTTTAGGATT AGGTGTCAA GCGCCTACTG 7260
 CAGAGTGGGG CATGATGCTT AACGAaGCTA GAAAAGTGAT GTTTACACAT CCTGAAATGA 7320
 15 TGTTTGCGCC AGGTATTGCC ATAGTGATTA TAGTGATGGC ATTTAACTTC TTATCCGATG 7380
 CTTTACAAAT TGCTATTGAT CCCC GCATCT CTCTAAAGA TAACTTCGT TCTGTGAAAA 7440
 20 AAGGAGTGGT GCAATCATGA CATTGTTAAC AGTTAAACAT TTGACGATTA CAGATACCTG 7500
 GACAGATCAA CCACTCGTGA GTGATGTGAA TTTTACATTA ACTAAGGGTG AAaCTTTAGG 7560
 CGTTATTGGA GAAAGTGTA GTGGTAAATC AATCACTTGT AAATCGATTA TTGGTTTGAA 7620
 25 TCCCGAACGA CTCGGGGTGA CAGGTGAAAT TATCTTTGAT GGTACaCAA TGTGTGCATT 7680
 ATCTGAATCG CAATTGAAAA AGTACCGTGG TAAAGACATT GCGATGGTCA TGCAACAAGG 7740
 TAGTCGTGCC TTTGACCCAT CAACTACTGT CGGTAAACAA ATGTTTGAGA CTATGAAAGT 7800
 30 ACATACGTCA ATGTCTACAC AAGAAATTGA AAAGACATTG ATTGAATATA TGGATTATTT 7860
 AAGTTTGAAA GATCCTAAAC GTATATTAAA ATCATACCTT TACATGTTAT CAGGAGGAAT 7920
 GTTACAGCGA TTGATGATTG CTTTAGCGTT AgcTTTgAAA CCAAAGTTAA TCATTGCTGA 7980
 35 TGAGCCGACA ACGGCTTTAG ATACAATTAC ACAATATGAT GTACTGGAAG CATTATAGA 8040
 TATTAAAAAA CACTTTGACT GTGCGATGAT TTTCAATTCA CATGATTTAA CGGTTATTAA 8100
 CAAGATTGCA GACCGTGTG TTGTGATGAA AAATGGTCAG CTTATTGAAC AAGGGACACG 8160
 40 TGAATCAGTC TTGCATCATC CAGAACATGT TTATACGAaT ATTKtATTAT CAACGAAGAA 8220
 GAAGATTAAT GATCATTTTA AACATGTGAT GAGGGGTGAT GTACATGATT AAAATTAAAG 8280
 ATGTTGAAAA GTCATATCAA AGCGCACATG TTTTAAAGCG TCGTCGAACA CCTATCGTGA 8340
 45 AAGGTGTGTC ATTTGAGTGT CCAATCGGTG CGACGATTGC GATTATCGGA GAAAGTGGTA 8400
 GCGGTAAATC GACGTTGAGT CktATGATAT TAGGTATTGA GAAACCGGAT AAAGTTGTG 8460
 50 TAACCTTAAA TGATCAACCG ATGCATAAGA AGAAAGTGAG ACGTCATCAA ATTGGTGCTG 8520
 TATTTCAAGA TTATACGTCA TCATTACATC CATTTAGAGT TGTTAGAGAA ATCTTATTTG 8580

55

	TGTTGGAAGA AGTCGGTCTA TCTAAGGCAT ACATGGATAA ATATCCTAAT ATGTTATCAG	8700
	GTGGAGAGGC GCAACGTGTT GCGATTGCGC GTGCAATATG TATTAACCCT AAATATATTT	8760
5	TGTTTGATGA AGCCATTAGT TCACTCGACA TGTCAATTCA AACACAAATA TTAGATTTAT	8820
	TGATTCATTT ACGTGAAACG CGTCAGTTGA GTTATATTTT TATCACACAT GATATTCAAG	8880
	CTGCCACGTA TTTATGTGAT CAATTAATTA TTTTAAAAA CGGAAAAATA GAAGAACAAA	8940
10	TTCCGACAAG CGCATTGCAT AAAAGTGACA ATGCTTATAC AAGAGAATTA ATAGAAAAAC	9000
	AATATCATT CTAAGGAGTG AGATAATGAA AGGTGCAATG GCTTGGCCCT TTTTGAGATT	9060
15	ATATATATTA ACATTGATGT TCTTTAGTGC CAATGCAATC TTAAACGTGT TTATACCTTT	9120
	ACGAGGGCAT GATTTAGGCG CAACGAATAC GGTATTCGGT ATCGTTATGG GGGCATACT	9180
	GTTAACAGCA ATGGTATTTT GACCATGGGC AGGACAAATT ATTGCTCGTG TCGGTCCCAT	9240
20	TAAAGTATTA AGAATTATTT TGATTATCAA TGCCATAGCT TTAATTATTT ATGGTTTTAC	9300
	TGGCTTAGAA GGTATTTTCG TAGCACGTGT TATGCAAGGT GTGTGTACGG CATTCTTTTC	9360
	TATGTCTTTA CAGCTAGGTA TTATTGATGC ATTACCAGAG GAACATCGTT CTGAAGGTGT	9420
25	ATCATTGTAC TCGCTATTTT CAACGATTCC AAACCTAATC GGACCATTAG TTGCCGTAGG	9480
	TATTTGGAAT GCAAATAATA TTTCATTTT TGCAATTGTC ATTATCTTTA TCGCATTAA	9540
	AACAACATTC TTTGATATCG CGTGACCTTT GCTGAACAGG AACCCGATAC GTCAGATAAG	9600
30	ATTGAAAAAA TGCCGTTTAA CGCTGTAAT GTTTTTGCGC AATTTTTCAA AAATAAGAG	9660
	TTGTTGAACA GTGGTATTAT CATGATTGTT GCATCGATTG TATTTGGTGC AGTTAGTACA	9720
	TTTGTTACCGT TATACACAGT GAGTTTAGGA TTCGCGAATG CGGGAATCTT TTTGACAATA	9780
35	CAGGCCATCG CAGTTGTTGC GGCAAGATTT TACTTAAGGA AATACATTCC GTCAGATGGT	9840
	ATGIGGCATC CTAAATATAT GGTATCTGTA CTATCATTAT TAGTAATCGC GTCATTTGTA	9900
	GTGGCATTG GTCCGCAAGT AGGTGCAATT ATTTTCTATG GTAGTGCAT ATTAATAGGA	9960
40	ATGACGCAAG CAATGGTGTA CCCAACATTA ACATCATACT TAAGCTTCGT CTTACCAAAA	10020
	GTAGGTCGTA ATATGTTGTT AGGTTTATTT ATTGCCTGTG CAGACTTAGG TATATCGTTA	10080
45	GGTGGCGCAT TGATGGGACC TATTTCCGAT TTAGTAGGAT TTAAATGGAT GTATCTAATT	10140
	TGTGGTATGT TAGTCATTGT AATAATGATT ATGAGTTTCT TGAAAAAGCC AACACCACGT	10200
	CCAGCGAGTA GTCTTAAATG AAGTGAATTA AAGCATATTA AGTTAATGAA TATTTAATT	10260
50	TTAAAAGGTA TATTGAGCAT GGCGATTCAT GTGCTTCATG CTAGGACATG AAACATTCTA	10320
	TATGGCTCGT TTTTAGAACG ACATATATCT AAATAAGCA CGCTTAAG TGAGTTTTGA	10380

55

	TTACATGAAA ATATGCAAAA CGAGTATAAC TGCTAATTGA TAGAAATAGC TCACCATAAA	10500
	ATTACGGTAT GATTTTAAAT ATAAGTAAGT CGCACTACCT GCTAGTATCA ATGCTGGAAT	10560
5	GAATTCACAC CATGTATTAA TGTATGGATA GTAGAACAGA GTTTCAGGA TAATGGACAA	10620
	TACTATTGTA ATCTTTAAAG GTATTAATCT GCTTAATTCT TGAATTAAAA TATGACGGAA	10680
	AATAAGTTGA CAAATCAAAG TATTTAATAT AATGGTTAAC GAAAATATAG CTATTAAACT	10740
10	GATGGA ₂ CCA TACCCTTTAA TGAGCGGGTA AATGTCAAAG ACAGTAAAGG AATCTACATT	10800
	TAGTGC ₂ AAA ATATTGAAAT GATTTAAAAG TAAAAAGAGT ACGACACTTA GTGTAAATGA	10860
	TATAAGAATA TGCCATTTAT ATTTAGCACT AGCAACGATT TGCGAACGTA TCATTGGAAT	10920
15	AAACGCATCT TCATGCATCA GACGAAAAAT AGCTAGTGAA ATAATAACTG CGAGTAAATA	10980
	GCTAATGTTT ATTGAAATAG GAAAAGAGAA ACCCCACGGA GCTTGTGAG TGAATACAGC	11040
20	TACTAACCCA AAAGTTAAAA AGACGATAAT GATCGGCAAG ATGTTAACCA AAAATATGTA	11100
	AAGGAAAATA AATCCAATAT CACGTTTGAA AAAACGCGAT TGTTCCGGTAG CGTATTCTTC	11160
	TTCTATGTAA TGTTTATTTG TATTTGACAT AGTATACCTC TTAAATAGTT GTATTATATA	11220
25	GATACTTTAG CACATATTAC TTTGTATTGT ATGTTTTATA CATTAAAATT TAAAATGAAA	11280
	AACATATCAT AAAATTGTTT TATAAAATGA AGCGCTTCCA TTGTGTTTTG TTTTGTAAGG	11340
	TGTATCATAA ATATTGAATT GAAATTTTGG GGGGAGGTAT TGTAATGACG TTTCTTACAG	11400
30	TCATGCAATT TATAGTTAAC ATTATCGTTG TAGGATTCAT GCTTACGGTT ATTGTTATCG	11460
	GGCTTATTTG GTTAATTAAA GATAAAAGAC AATCACAACA TAGTGTATTA AGGAATTATC	11520
	CTTTACTAGC ACGTATTAGA TATATTTTCTG AAAAAATGGG ACCGGAATTA CGTCAGTATT	11580
35	TATTTTCTGG GGATAATGAA GGGAAACCTT TTTCACGTAA TGATTATAAA AATATCGTTT	11640
	TGGCTGGAAA ATATAACTCT CGTATGACCA GCTTCGGTAC TACTAAAGAT TATCAAGACG	11700
40	GCTTTTACAT ACAGAACACA ATGTTTCCGA TGCAACGTAA TGAGATTTCA GTAGATAATA	11760
	CAACATTGTT ATCAACATTC ATTTATAAAA TCGCGAATGA GCGTTTATTT AGTCGTGAAG	11820
	AATATCGTGT GCCGACAAAG ATTGATCCGT ATTACTTAAG TGATGACCAT GCAATAAAAT	11880
45	TAGGTGAACA TTTAAAACAT CCATTTATTT TAAACGTAT CGTAGGACAA TCTGGTATGA	11940
	GTTATGGCGC TTTAGGAAAA AATGCCATTA CAGCTTTATC TAAAGGTCTA GCTAAAGCGG	12000
	GCACTTGGAT GAATACAGGT GAAGGTGGCT TATCAGAATA TCATTTAAAA GGTAATGGGG	12060
50	ATATCATTTT CCAAATTGGT CCCGTTTAT TTGGTGTTCG TGATAAAGAA GGTAATTTTA	12120
	GTGAAGGTTT ATTTAAAGAG GTTGACAGT TATCTAACGT ACGCGCATTT GAGCTGAAGT	12180

55

	TTGCTAAAAT CCGAAATGTT GAACCTTATA AAACAATCAA TTCACCTAAC CGTTACGAAT	12300
	TTATTCATAA TGCTGAAGAT TTGATTCGTT TCGTCGATCA GTTGCGAGCAA TTAGGTCAAA	12360
5	AACCACTAGG ATTCAAAATT GTAGTAAGCA AAGTTTCAGA AATTGAAACA CTTGTACGTA	12420
	CGATGGTGGA ACTAGATAAG TATCCAAGCT TTATTACGAT TGATGGTGGT GAAGGTGGTA	12480
	CTGGTGCAAC ATTCCAAGAA TTACAAGATG GTGTTGGCTT ACCGCTATTT ACAGCTCTAC	12540
10	CTATTGTGTC TGGCATGTTA GAAAAATATG GTATTGAGA TAAAGTGAAA TTGGCGGCAT	12600
	CTGGTAAGTT AGTGACACCA GATAAAATTG CGATTGCACT AGGTTTAGGT GCAGATTTTG	12660
	TAAATATCGC ACGTGGGATG ATGATTAGTG TCGGTTGTAT AATGAGTCAA CAATGTCACA	12720
15	TGAATACGTG TCCTGTAGGT GTTGCAACGA CAGATGCGAA GAAAGAAAA GCATTGATTG	12780
	TTGGAGAAAA GCAATATCGT GTCACAACT ATGTAACAAG TTTGCATGAA GGCTTATTCA	12840
20	ATATTGCAGC AGCTGTTGGC GTATCCAGTC CTACAGAAAT TACTGCTGAT CATATTGTAT	12900
	ATCGAAAAGT CGATGGTGAG TTACAAACGA TACATGATTA TAAATTAAAA CTCATTAGTT	12960
	AACCTAATTA TTTTCGGGAAA TTGAAAGCAG CGGATTTTAG CGTACTGCA AATAATTTTA	13020
25	TATTAGTAGT GGATGCTGGT CACACAAGAA CTTCAAATAT TAAAGCCCTC AGAATATGAA	13080
	TTAAGGTTTG TAACCTTAGT CTTATCTGAG GGCATTTTTA AGTTATAAAC TATTTGTCGT	13140
	CCATTTTATC TTTTCTTTT AAACCTCTGT GCTTTAATTG CTTTTCAAGT TTTTCAAAC	13200
30	TAATATCTTT ATTTTCTTTA GTCGAAACAC CAAGACGTTT ATTTAATTTT TTCATGTCAA	13260
	CTTCTGTGTA ATCTATGTCT AAGTGyTCAA TTGCTTTTTT ATCTTTATAG TCTACTTTGT	13320
	ATTTTACGCC TTTAAGGTCT TTGAAAATAC TTTTCAGATT GGCGAATAAC TTTTGGCTT	13380
35	CGTCTTTATC CATACCTAGA TCGTCATATT TAATTGTGTT GATTGTAGAC TGTTTTAAAA	13440
	CTTTATCATC TTTATATGTG ATAGAAGTTA GTACATGTTT ACCACTAACA TCACCWTCAT	13500
	ATGTTTTGGT TTGTTCTTTA CCACAAGCTG ATAATGCAAT GATACAACT AATGCTACTA	13560
40	CAATTAATGA ACATAATTTT TTCAAAGTCA GTCGCCTTCT TTCGATATTT GTATTATAAA	13620
	GAAATTATAA CATTTACTAA AAAATGATGT TATTCAAAAA TTAAATTTT GTCATTTTTT	13680
45	TTGAAGATAT GAGTTTTTTT AAGCGGATTC CTCACAAAT TTAAAAATA TTTAAGCCTK	13740
	AAAATGATAA AGCGKTAGGG AACGTTTTTC TGAAAGTTAG TGATACAATA GTTTTAAGTT	13800
	GAAATACAGG AGGATGAATA ACATGAATCA GTCAGTCAAA TTACTTAAAC ATTTAACAGA	13860
50	TGTAAACGGC ATTGCTGGTT ATGAAATGCA AGTTAAAGAA GCAATGCGTa ACTATATAGA	13920
	GCCTGTCACT GATCAAATTA TTGAAGATAA CTTGGGTGGC ATTTTTGGAA AGAAAAATGC	13980

55

AACAAAGATT GATAAACATG GTTTTATTTT ATTTACGCCA kTgGTGGATG GTGGAATCAA 14100
GTCATGCTAT CTCAAAAAGT AACGATTACA ACAGATTCCG GCAAAGAAAT TAGAGGTATC 14160
5 ATCGGTTCTA AACCGCCACA TGTCTTAACG CCTGAAGAAC GTAAAAAGCC AATGGAAATC 14220
AAAAATATGT TTATAGATAT TGGTGTTAGT AGCAAGGAAG AAGCTGAAGA AGCTGGCGTT 14280
GAAGTAGGCA ATATGGTTAC GCCATATAGT GAATTTGAAG TGCTTGCAAA TGATAAATAT 14340
10 TTAAGTGCAG ATCATTGAT AATCGCTATG GCTGTGCATT AGCTATTGAG GTATTAAAAC 14400
GTTTAAAAGA TGAAAATATT GGCATTAAC TATACAGTGG TGCCACAGTG CAAGAAGAAG 14460
TTGGTTTGGC TGGTGCGAAA GTGGCAGCGA ATACGATTAA ACCAGACTTG GCGATAgcTG 14520
TcGATGTAGG TATTGCTTAT GATACCCAG GTATGTCAGG TCAAACGAGC GATAGTAAAC 14580
TAGGCGGTGG TCCAGTTGTC ATTATGATGG ATGCTACAAG TATTGCTCAC CAAGGTTTGC 14640
20 GAAAgcATaT TAAAGATGTA GCTAAGGAAC ATAACATCGA AGTACAATGG GATACGACAC 14700
CAGGTGGAGG TACAGATGCG GGAAGTATTC ATGTCGCAAA TGAAGGTATT CCAACGATGA 14760
CAATCGGTGT TACGCTGCGA TACATGCATT CTAATGTTTC AGTGCTCAAT GTAGATGATT 14820
25 ATGAAAATTC TATCCGTCTT GTTACTGAAA TTGTCGGTTC ATTGAATGAT GAAAGTTATA 14880
AAAATATCAT GTGGTAATCA AATCCATAAA TAATAAGAA TCCTTTTAAT ATGGTAGGTT 14940
GTTAAACAAT TGTCTAATTT TAATTCTTAG TCATTAGACA GTATCCATGT TAATAGGATT 15000
30 TTTTGTTTTT AATTTAAATG CTGAAAATCA ATTATGCCTA AATTTTGATA TTACAAGAAA 15060
ATGATTTTTT CTAAATGTA ATTGCACTAA AAACCAAAAA AACGGGAATA ATATACCTGA 15120
TATATTACAT GAGGAGCGGT GCAAATGTTG TTAGAAATTA AAGATTTAGT GTATAAGCG 15180
35 AGCGATAGAA TCATACTAGA TCATATCAGT CTAAAAGTAG ATAAAGGCGA GAGTATTGCC 15240
ATTATAGGTC CATCAGGTAG TGGTAAAGT ACATTTCAA AGCAAATATG TAATTTGTTT 15300
AGTCCAATA GTGGAGAACT TTATTTTAAA GGTAAACCCT ATAATGATTA TGACCCGGAA 15360
40 GAATTGCGTC AACGAATCAG TTATTTGATG CAGCAAAGTG ACTTGTGTTG TGAAACGATT 15420
GAAGATAACA TGATATTCCC ATCACTTGCA CGTAATGATA AATTTGATAG AAAACGTGCA 15480
45 AAGCAATTAA TTAAAGATGT CGGTTTGGGA CATTATCAAT TAAGTTCGGA AGTGGAAT 15540
ATGTCGGGTG GTGAGCGGCA AAGAATTGCT ATAGCGCGCC AACTGATGTA TACACCGGAT 15600
ATTCTTTTAT TAGATGAATC GACCAGTGCA TTAGACGTTA ATAATAAGA AAAGATAGAA 15660
50 AATATCATTT TTAAATTAGC AGATCAAGGC GTGGCAATTA TGTGGATTAC CCACAGCGAT 15720
GACCAAAGTA TCGCACACTT TCAAAGCGT ATAACAATTG TTGATGGTCA AATTTCTAAT 15780

55

	CATTCCGATT ATCAITTCAT ATAAAGAAGG TTTACATATT ATTAAAGATT TAATTGTTGC	15900
	GACATTACGA GCAGTTGTGC AATTAATCAT TTTGGGATTT TTGCTGCATT ATATTTTAA	15960
5	AATAAACGAT AAATGGCTGC TTATTTTATG TGTATTGGTC ATTATTATTA ATGCATCATG	16020
	GAATACAATT AGTCGAGCAT CACCAGTGAT GCATCATGTG TTTTGGATAT CATTTCTAGC	16080
	TATCTTCATT GGAACGGCAT TACCGCTTGC AGGTACTATT GCGACAGGGG CCATTCAATT	16140
10	TACCGCAAAT GAAGTTATAC CTATCGGCGG CATGCTTGCA AATAATGGCT TGATTGCAAT	16200
	TAATTTAGCT TACCAGAATT TAGATCGTGC ATTCGTACAA GATGGTACTA ATATTGAATC	16260
	TAAATTATCA CTTGCAGCTA CACCTAAATT GGCTTCTAAA GGTGCAATAC GTGAAAGTAT	16320
15	TCGTTTAGCT ATAGTGCCAA CTATTGATTC GGTAAAACA TATGGGCTTG TGTCGATTCC	16380
	TGGTATGATG ACAGGCTTAA TTATTGGTGG CGTACCACCT TTACAAGCGA TTAAATTTCA	16440
20	ATTGTTAGTC GTGTTTATTC ATACAACTGC GACCATTATG TCTGCTTTGA TTGCGACATA	16500
	TTTAAGCTAT GGTCAATTTT TCAATGCAAG ACATCAATTA GTAGCACGAA ATACTGATGT	16560
	TAAGAGTGAA TCATGATAGA TTTTACTGCA TCAGATTAG GCATTAGTTT TAATTGGAAA	16620
25	TGAAGTGACG CGCACATATA GTATCGCTAT TCATTAGCGC AGCGAAAATA TTCATAAAGG	16680
	CACGCATACT TTGTAGTCAG TTATCTGTTC TGACATATAA AGCGTGCGTG CTTTTTTGGA	16740
	GTTATTGTTG AACTGAAGT AATTATACAT AATTATTAAA TGACATACTT GTGTTAATTT	16800
30	TTCAAACTACT GAAAAACAAT TTCaATAATT TTCCaATTAA GCACAGAAAA TTAAAGCAAA	16860
	ATATTATATA ATAGAACGGT TATATATAaA nATTngTgCA CACATTTTTT AATAAATCGT	16920
	TATTCTAAGG GAAATGAATA TCGGAAATTT TGTTTGAAAG GAGTTTTTAA TTGTCAATCA	16980
35	TGCGACTATT TACATTCATT TTAAGTATTT TTATCGTAGG AATGGTTGAA ATGATGGTTG	17040
	CAGGAATTAT GAACTTGATG AGTCAGGACT TACATGTATC AGAAGCTGTC GTTGGTCAAT	17100
	TAGTGACAAT GTACGCTTTA ACATTTGCGA TATGTGGACC TATTCTGGTT AAATTAACGA	17160
40	ACCGTTTTTC ATCAAGGCCT GTATTATTAT GGACATTACT TATATTTATC ATTGGTAATG	17220
	GCATTATTGC TGTAGCGCCA AATTTTTCaA TATTAGTAGT TGSTAGAATT ATCTCATCTG	17280
45	CAGCAGCAGC ACTAATTATC GTAAAAGTAT TAGCTATTAC AGCGATGTTA TCAGCACCTA	17340
	AAAATCGTGG TAAAATGATT GGAATTGTCT ATACAGGGTT TAGTGGTGCT AATGTTTTTG	17400
	GTGTACCAAT TGAACGGTT ATCGGCGATT TAGTAGGTTG GCGCTATACA TTTCTATTCT	17460
50	TAATTATTGT GAGTATTATT GTTGGCTTCT TGATGATGAT CTATTTACCG AAGGATCAGG	17520
	AAATACAACG AGGCCCTGTG AATCATGAGA CACCATCTCA TGAAAATCAT GTTACTTCGA	17580

55

CAAACTCAGT GACATTCGTC TTTATAAATC CACTTATTTT ATCTAATGGT CATGATATGT 17700
 CATTTCGTTTC ATTAGCACTT CTAGTAAATG GAATCGCTGG CGTTATTGGA ACATCATTAG 17760
 5 GTGGTATATT CTCCGATAAA ATTACAAGTA AGCGTTGGTT AATGATTTC TTTTCTATTT 17820
 TTATCGTCAT GATGTIACCT ATGAATTTAA TCTTACCTGG TTCAGGTCTA TTGTTAGCAG 17880
 GACTATTTAT TTGGAATATC ATGCAATGGA GTACTAATCC AGCAGTGCAA AGCGGTGTGA 17940
 10 TTCAACATGT TGAAGGCGAC ACAAGCCAAG TAATGAGTTG GAACATGTCT AGTTTAAACG 18000
 CTGGTATTGG TGTGGAGGC ATTATTGGAG GCTTGGTCAT GACACATGTT TCTGTTCAAG 18060
 CTATCACATA TACGAGTGCC ATCATTGGCG CATTAGGATT AATCGTTGTT TTCACATTGA 18120
 15 AAAATAATCA TTATGCTAAA ACATTTAAAT CATCATAATT CTCATATGAm AAGCACGCCT 18180
 GCTATCAAAT TCAGGTGTGC TTTTITAGAT GCGATAACGT TATTGATATG TGCATAATA 18240
 20 GCGACGTTCA TTATGATACA TCGGCCAAGG CATTTTACCG CTTTITAGCAA AATTAGCTAA 18300
 ATCATTTTGC ATTTGTGCGAC TTAATAATTT AAGGTGAGCA GTTGTGGaT ATgAT 18355

(2) INFORMATION FOR SEQ ID NO: 68:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1192 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

35 CGCAAAGAAG TACAAAAAAT GTTTTACAA GAAGGTATTA AAACACCTCA ACCAATTATG 60
 ACTGCTTATA ATCATAGTGA AAACGgTGTT TAGTAGTTTA TAATACATGG AGGTCATATT 120
 TAATGGCGTC AAAATATGGA ATAAATGATA TAGTAGAAAT GAAAAACAA CATGCGTGTG 180
 40 GAACAAACCG TTTTAAGATT ATTAGAATGG GTGCAGACAT AAGAATTAAA TGTGAAAATT 240
 GTCAAAGAAG TATTATGATT CCACGTCAA CGTTTGATAA AAACTTAAA AAAATCATCG 300
 AATCTCATGA TGATACACAA AGATAGGAGA ATGATTAATG GCTTTAACAG CAGGTATCGT 360
 45 TGGATTGCCA AACGTTGGTA AATCAACATT ATTTAATGCA ATAACAAAAG CAGGTGCTTT 420
 AGCAGCGAAC TATCCATTCG CTACGATTGA TCCTAATGTA GGGATAGTAG AAGTGCCAGA 480
 TGCTAGATTA CTAAATTAG AAGAAATGGT TCAACCTAAA AAGACATTGC CGACTACATT 540
 50 TGAATTTACA GATATCGCTG GTATTGTGAA AGGTGCTTCA AAGGGAGAAG GGTTAGGTAA 600
 TAAATTCTTA TCACATATTA GAGAAGTAGA TGCATTTGT CAGGTCGTTT GTGCATTTGA 660

55

TAATATGGAA TTAGTACTAG CGGACTTAGA ATCTGTTGAG AAACGTTTGC CTAGAATTGA 780
 AAAATTAGCA CGTCAAAAAG ATAAGACTGC TGAAATGGAA GTACGTATTT TAACAACTAT 840
 5 TAAAGAAGCT TTAGAAAATG GTAAACCCGC TCGTAGTATT GACTTTAATG AAGAAGATCA 900
 AAAATGGGTG AATCAAGCGC AATTACTGAC TTCTAAAAA ATGCTTTATA TCGCTAATGT 960
 TGGTGAAGAT GAAATTGGTG ATGATGATAA TGATAAAGTA AAAGCGATTG GTGAATATGC 1020
 10 AGCGCAAGAA GACTCTGAAG TGATTGTTAT TAGTGCAAAA ATTGAAGAAG AAATTGCTAC 1080
 ATTAGATGAT GAAGATAAAG AAATGTTCTT AGAAGATTTA GGTATCGaAG AACCAGGATT 1140
 15 AGATCgrTTA ATTAGGAmCA ctTATGAATT ATTAGGnTTA TCCACCATAA TT 1192

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7494 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

AATATAGCTG CAATAGCATC TCGTTTCATT TGTATAATCA ATTCCGGTTT AAATATCAGT 60
 GTGAACGTAA GCACGACACA GATTAAAAAT AACACTGCCG GAATGAGTCG TTTCAATCGT 120
 30 CGCTtCCAAA ACTCTAGCAA ATCGATTTTT TGCGTCCGAT AATACTCACT TATCAACAAA 180
 CTTGTTATTA AATAACCTGA AATAACGAAG AATGTATCTA CTCCTAAAAA GCCCCCACTT 240
 AACCATTGTG CATTCAAGTG ATAAATAATG ATTCTTATAA CTGCGAATGC CCTCAATCCA 300
 35 TCTAATCCAG GTAAGTATCG CGGGGAATAC ATTTTTTCTA AACGTTTAAA GTCTTTTGTA 360
 TCCAIGTTAA TAAACGCCCC ATTTATTTTT CTCTATTTTG TAGTATATCA CAATATTTTT 420
 GAAAAATAAA TATTGCACTG aTTTTCATT ATTGATTAA CCCTTAATTA AGATAGTTTT 480
 40 AAATTTTTTA TTAAGTAGAA AACAATTATT ACAGTTGATT TCATTACTGC AAACCACATA 540
 TAAATTTGTC GATTTTACTA CATAACATAG ATTATCATAG ATTCTTGAAT TTTTAGCAAA 600
 45 ATAACGTGTA TTTTCATTAT ATTTTACAA AAAAAGGTTT GTTTTATATT TTATGCATCT 660
 TACTGTAACA GAATCATTAA GATATGCTAT TCGAATATAC TTTTCAAAA TTTATATAAT 720
 GAATAAATTA ACATGTATTG AAAAAAAGC GAAATGCAGC CTATCCTCTA ATGTAAACCA 780
 50 AACGATATAT CTCGTCAGAC TTTATATTTA AACGCTATGT GTCACTTTTA AAATGAATAT 840
 TACTAAGATT GTCATATCAA TTATTATTGC ATCGAATTAA TCTTTTAAAT TTCTGTAATA 900

55

ACGGAAGTCA TTATTAGAAT AAAAATACTG TGCACATAA AATTTATCAA TTGTTCTTAA 1020
 ATAAATACCA TCGATATTTT GTTCTTTACA TGTCATTATA ACTTTATCTA AAAGTTTTTT 1080
 5 ACCTATTTTT AAATTCCTAT AACCTTTATC AACAAACATT TTTTAAAGTG CAGACATATT 1140
 ATTATCTAGT CTAATCAAAC CTATAGTACC AACAAATATT TGATGATTGT TTATTGCAAG 1200
 CCAAAATGCC CTCCATTATT CAAATAGTTA TGTTGATGT TCTCCAAATC AGGTTGATCA 1260
 10 TCTCTATCAA TTTTATATA AATTCATTTT TTTGAATCGA TAAAATAAAC TCGATTAGCT 1320
 CTTCTTATA AGACCTATTA TATTCAATTA TGTTTATAGC CATTTTATC TCCTTTTTCA 1380
 TTTAATTTAA TTATAAAATG TCGGTTTGTG TTGTATCTAG TGTACTCAGT ACAGCCTCAA 1440
 15 ATGAAGTTTC ATTCCACTTG GCACTTAATA AAGACAAGTA TTTTAGCAGT AATACAATAA 1500
 AGTCCAATAA ATTTCCCTAA CTTCAATATC CACTTTTAA AAAATGTATT TTTAATTAAT 1560
 20 AAAAAAATC TCCCCAATTT CTATGGGAAG AGCTATATAT TTAATGTCTA AACATTACTT 1620
 TTATTTATTA TGAAGGAATT AGAATCCCCA AGCACCTAAA CCTGTGCTT TGTATGCTTT 1680
 AACAGCTGCG TTGATTTGTT GGTCAACAGT GTTTGTTGGA CCCCACCTG GCATAGTTTG 1740
 25 GAATAAACCT GAAGCACCTG ATGGGTTGTA AGCATTTACT TGACCATTG ATTACAGAGC 1800
 GATGATTGCA GCCCATGTAG AAGCTGAAAC ACCAGTACGT TGAGCCATGA TTTGAGCTGC 1860
 TGATGAACCA GTAGCACCTG CAGTATTACC ATTGCTTAAT CTCACTGAAC TTGAAGTAGT 1920
 30 TGAAGTGCTG TAGTTATGGT AAGTTGGAGC TGAACAGCT TCAACGTTTG AGTTACTTGA 1980
 TTGTGCATTG TAGCTTACTG ATTGTACATT TGAACCTTGG TTGTATGAAG TAGTGTAGTC 2040
 TGCACCTGCA ACGTTTGAGA AACCAGCAGT TTGACCATTA GCTGCTTCAT AGCTCCATGA 2100
 35 CCATGTAGTA CCATTTGAAG TGAAGTTATA TTGAAACCA TCTTTTACAA AGTGGATGTC 2160
 ATATGCACCA TCTTTGATTG GAGCTGCATT TAATTGATCT TGGTGATTAT GCGCTAAGTC 2220
 AACTAAGTGT GCTTGATCAA CGTTTACTTC AGCAGCGTGT GCTTGATGTC CTGTACCTGC 2280
 40 TCGTAACCT GTTACACCTA ATGCCACTGC TAATGATGAT GCCATAATTG TCTTTTTCAT 2340
 AGTAAAAAAT CCTCCAGTAA TAATTGTnAG TTTATGTTTT TAGTAATTAT AtTTTGaATT 2400
 45 TGAATGTCGT AGTgCAAGTT TAAATTGTCT TTTATTTCTT TCaACGGTAC TCACTATATC 2460
 ACaAaAAACC AGCCAGTAAA TTACACTTTC TTTACAAAC ATTACAATAT CAAGTGTTAT 2520
 TTGtAATGTT GAAATATGGC TGTTTTATAC TGTAATGTGA AATATGTGCC CTTTAGAATC 2580
 50 CAATCAACCC TTGAAATAGT CTTTAACACA TAAGATTTTT ACTATATTTA GCTCAACTAT 2640
 TACAGCTTTC GTAATATTAC AGATTGTATT TTTGTTACAT AGCTGTAATA TATCTGACAT 2700

55

	TACACATGTA TTGATTGCTA TTATTGTTGT ATATTCAAAG TTTTAAAACA CACATCTTTT	2820
	GTGAATTGTC TTATCTTTTA TTAGCGCAA TAACTGCAG CTCAATTATA TTGTTCAACT	2880
5	TCATTCTCGC AATTCACAAT AACATTAAAT AATTTTGGT CTCATATTTT CAAAAACAT	2940
	ACTGTTATTA TCCCATGAAT TTAAAAATAT CATTAGTATA TAAACGAAAC ACTTTACGAT	3000
	AAATGATATC TGCAAGCCAA GCTGTTACAA ATGGTACAAC AAAGAACGCT ACTACAATTA	3060
10	GTAAGACACT CAACCAAGCA GAATCAACCT CCATAAATTT AAATGCATTA ATCGGTCCTA	3120
	CCATTCTAT AAAACCAAAT CCAGCTGACT CTTTCGTTCC ATGAATACCT ACTAATGCTG	3180
15	ATACCAACC TGATACAATG GCTGTCGTTA ATATTGGTAA CATAAGAATT GGATATTTCA	3240
	CCATATTAGG TATCATCATT TTAACGCCTC CAAAGAAGAC GGATAACGGC ACCCCTAAAC	3300
	GATTCACTTT ACTTGTACCA ATTATCAATA CTGCTTCAGT CGCGGAGATA CCAATTGACG	3360
20	CTGATCCAGC TGCTAAACCT GTAATACCTA TCGCAAAGGC AATGGCCACA GTTGATAGTG	3420
	GCGAAATAAT AATAAGACTA AATACCATTG AAATCAAAT ACTCATGACA ATCGGTTGTA	3480
	ATTCTGTAAC ACCATTAAAC ATATTACCGA TGGCTGTTGT AATCATTTTC GTATACGGCA	3540
25	ATATTAAAC ACCAATTGCA CCTGAAATAC CGCCAACAAC TGTGGAAT ACAATCAATG	3600
	CCATACTACC TACGCGATGT TGAATAAGTA AAATGAATAA CACTGCAATC GCTGCTGTAA	3660
	TCATTGTATT AATTAAATCA CCAATACCCG TAATCATCCA AGCACCATT TTAAGTGGC	3720
30	CTGCACCGCT TCCTACATAT GCTGCATTG CCACAACAGC AATTGCTAAT GGCGATAGGT	3780
	CAAATTTTCAT GGCAACCAAT GCACCAATCA AAGCAGGTAC TGTAATTTGA ATTGCAACGA	3840
35	CAACGCCTAA TAACGTTTTA AAAATCGGAT GATAATCCAT AAAGTATTTA AAAATTTCTC	3900
	CAAGTATCGC ATTAGGAAC TAAACCCGCA CAATACCTAT GGCGACACCT GATAAACTC	3960
	TAAATATAAA ATCTTTGGGT GTAATTGTTT TAATTGATGT CATAATATCA TCCTTCCATT	4020
40	TATGTATATA CATCTGTATG CAAATAATAA AGAGCCTTAA GTTATAAGCT GCCACTAGCT	4080
	TAAATTCTAA GATGTGCATG CCGATGTTGT TATATTTAGG CTAGCAGTAT CATCTATAAC	4140
	TCAAGACTAT GAAAAATAGT ATATCACAAA ATTCTGAATT TTTAGATAAA TAAATTGGCA	4200
45	ATTTTTCAAA CATATTGTTA CAATACACTT TTATTTTATC TTCATTTTAA AAATCCATTA	4260
	ATACAATAGA AGAAAGACAT TCAAATGCTT ACCAAAAAGG TACATTATTT GTTAGGAGCG	4320
	TATCAGCACT TACATATCAT CAACACAATT GACAATATAA TAGAAGATAC TGATAATAAG	4380
50	TGTTAAACA ACAGATGTTA GGTAGTGAAC AAATGATGGA AAGTAAATCC ATAGATCCAA	4440
	GAATCGTTAG AACCAACAA TTGCTTGTCG ATGCTTTTCT TAAATTTCT AGAGAAAAGA	4500

55

TTTACGCTCA TTTGCTGAT AAAGAAGACC TCCTAGACTA CACATTATCT GTAACCATTT 4620
 TAAAAGACTT GAATGATAAT TTGAGCATTT CTAATGTCAT TAATGAAAAG GTTCTGCGTA 4680
 5 ATATTTTCAT TTCAATTGCG AGTTATATCA AAGATGCTGC AAAGTCTTGC GAATTAAATA 4740
 GTGAAGCATT TTGCAACAAA GCACATCAAC GTATTAATAA TGAATTAGAA GATATTTTTG 4800
 10 CGATTATGTT AGAAAACAGC TATCCGAGC ATCAACGAGA TATCATTGTA AATAGTGC GA 4860
 GTTTTTTAGC AGCTGGTATC TCAGGCTTAG CATTACATTG GTTTAACACG AGTCAAGAGA 4920
 CAGCCGATGT GTTTATCGAT CGCAACCTTC CATTTTTAAT TCATCATATA GCACATTTTT 4980
 15 AATAAACTT GGTATTTAGT CATGCATCTT GAAATCACTA TGTGACTTAG GTTCATACTT 5040
 GTACACACAA TAAAATTAA CGTATTACGA TTGATTAGCC GTGTCTAGGA CATAAATCAA 5100
 CGTCCTATAC TCTACAATGT CATATTAGCA GTCGTAACT GAATGAAAAT AAGCTTGTCA 5160
 20 TTAAACATA TAGATTTTAG TGACAAGCAT TTTTGT TTTT GCGTACTTAA ACAACACTTC 5220
 AGGCAATATG TTGTTTAGGC AACAAATGAT ATGTGCGTGT TTATTGGCAA ACGTACGACA 5280
 TAGTAGTATA GTATGTCTAA ACAACATATG TTGCATAGTT GATATGCGTT GTTTAAATAC 5340
 25 TAAGATAGGA GGGATTGACG TGAGCGAGAC AGATGAACCT CAGGGGTTTG AACGCACGCA 5400
 TAATATATTA AATATTAATC AGAGTAGTCT GGGTGTAGTG ACATACATTA CAAATAAATT 5460
 AAAGTCGACG TTGAAGCAAC ACATAATAAT TGCTCGTGGT AAAAAGCGAA TCGACTATCG 5520
 30 ACTGTCGTAT AACTTTTACA TACGTATTAT GATAATGTAG AAATCAAGAA AATCGACTGT 5580
 GAATATACCT ATGCTATGCC CATTGCAATT TTAATAAGAC ACACGATGTC ATTCGACAAT 5640
 GCTCATTCTT TTGCTCAGTT ACGTCATCCT GTCTTATAAA ACAACATTGC AGACATGTAT 5700
 35 ATCAAACGAC ACTTCAATAA CATCACTTTG CCcATCGTAC TACTAGTAAA ATCGTGTCTC 5760
 AAATcCCTTA TTTTAATTCC AAAAAcCTGC TGGTCAAAG ACCGAGAAAC TAAAAACATT 5820
 40 ACTTAATGTG TTGATAAATT ACCATATAAA AATAATCTCA AAATATATCA ACACTTGATT 5880
 CTAAGGAGGA TATGACAATA TGAAAATTTT AGATAGAATT AATGAACTTG CAAATAAAGA 5940
 AAAAGTACAA CCACTTACTG TAGCTGAAAA ACAAGAACAA CATGCATTGC GTCAAGAcTA 6000
 45 CTTAAGcATG ATCCGAGGAC AAGTATTAAC AACATTTTCC ACAATAAAAG TGGTTGATCC 6060
 AATCGGTcAG GATGTCACAC CAGATAAAGT TTATGATCTT CGCCAACAAT ACGGTTATAT 6120
 TCaAAATTAA cATTTGCTCA CGAGGTATTG CACTTAAGGT GCCAACTGAC CTCATAACA 6180
 50 AAGCCCATAC TGATTGAAGA CACTAATGTG tCsaCCATGG TGCACATTAC GCTTCATCTC 6240
 TGTATGGGCT TTTTATTTAT TCTTTTGAGA ATTTcATTTT AGCAGACCAA AAAATTAAAA 6300

55

TGAACGACTG TGCCACCCGC TTCTTTCAC TATTTCACCA ACTGGTCAAC TTCTTCATTT 6420
 GTGTTACACAC CTAGAGAAAT CATCACTTCA TTTGGTTCAG TATTAAGGCT TTGCTGACTT 6480
 5 ACATTTTGAA AATGCTTGTr TTCTATTAAA ATTACGGkTG tTTGACCTAT tTGAATGCCG 6540
 ACCATTTTAT CTAACATTG TGGGTTTCTA TTTATTTTAA ATCCTAACGC TTTATAAAAC 6600
 10 TGTGCGCTCT TTTCTAAATC TTGCACATGC AAATTAAACC ACATTGATTG AATCATGATT 6660
 GCACCCCAT TACTACTTAT TATAGTTTTG GACTTTAAGC CAATCACTTA ATGATAATCT 6720
 TGTGATTTT ATTTAGCCA TTAATTCAAA GTCTACTTCA TAACCTTTTT CTTCACCA 6780
 15 TTGCTTTTCT GCAACACCAC TAACAAATTC TCCTTCTATA ACAGTAGATT TACCTGTCAC 6840
 TTCATAAAA ATTGTTGCTG CTTCACTTAA TGTAACCTCA TCGGAACCAA TCTCTATTGA 6900
 TTGATGCGTA AAGCTTTGTG GATGTGCAAA AATATACGAT GCAATTTTAG CTATATCAAT 6960
 20 AGAAGAAATC ATTGTGAATT TTATATTCGG ATTAATAAAT TCTGGTAATG TAATACGTTT 7020
 ATCTTCGACT TTAGCAATGC GTAAAAAATT ATCCATAAAG AATGATGGTT TGATAACTGT 7080
 TGCATTTATA TTAGATTCCA TTAATCTATT TTCTATTTTT GCTAGTACTT CAAAGTGTGG 7140
 25 GCCAGTTCGA TTTCGATTAA CCCCTCCCGC AGTACTATAC ACAATATGTT GAATATTTTC 7200
 TTGCTCAGCT ATTTCAATTA TCTTCATACC TTGTCTTAAT TCTTCGCTAA CATCATCTTT 7260
 AACGATTGGC TGAATACTGT ATAAGCCATA CTTACCTTTC ATCGCTGATT GCAAACCTAAC 7320
 30 ATTATCACTC AGATCACCTT CArCGATTGA TAAATGCGGA TGTCTATGT CTGAAAGTTT 7380
 ACGATTATnC TTATTCTAG TTAATGCACT TACATACCAT CCATCCTCTA ACAACTGTTT 7440
 TACAACTGCA TTACCTTGCT TCCCTGTTGC GCCTATTACH AAAATATCTT TCAT 7494
 35

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11802 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

AATTTATTTT GCCGTCCCAC CCCAACTTGC ATTGTCTGTA GAAATTGGGA ATCCAATTTT 60
 TCTTTGTTGG GGGCCGCCCC CAACTCGCAT TGCCTGTAGA ATTTCTTTTC GAAATTCTCT 120
 50 GTGTTGGGGC CCCTGACTAG AATTGAAAAA AGCTTATTAC AAGCGCATTT TCGTTCAGTC 180
 AATTACTGCC AATATAACTT CGTAGATCAT AGAACATTGA TTTATTTCCC AGCCTATTCT 240

	AGCAAAGGTA ATAATGATAT TAATAATGTA CAAAAAATAT AAATCAAATC GACATCCTTA	360
	TAAACATCA GAACCACTAA AAACAAAAAA GCACAAAATA AAATTAAATT TAAATATAAC	420
5	GACCACTTTT CAAAAAATC TCtTTTCaTa TTTCCACCCC TAATTTTAAT AAGCATTATT	480
	TTATATTCTC TTTTAAGTTT ATTATTCAAA AGGAAAACAG AAATATCTTT CaATATTATT	540
10	ATAAACATTT CACTACTTT TAAAAACCAA CAAAAAATA CTTATTTTAA GTAGATGAGC	600
	ATAAGTGAAC ATAGTTCTTT AGTTATAATA ATTAATTCAA CCAAAGTCG ATTTGTTTTT	660
	GCAATTGGTT TTCATTTCTT CTTAAAGATA TTTTCATTAA ATCTGTCAA TCAATAGACG	720
15	CTATATTTTT CACTTATCT CTATATTTAT TTTTAGTACG TCTTTCTAAA TTTCCCCATT	780
	CCTCTTCTTC GTGAGTTAAT AAATGAAGCA TTGCTCGTTC TTGTATATTT TCAATCAATT	840
	TTAAATTCCG TTTTAAATA TGCAAATCAT CAAACAATC TTTCCAACAA TCAACCATAT	900
20	CTCGTTTTAA TTCAATTTCC ACACGCCATA GAAATGTTGA ATCAATTTCA ACATCTGCAT	960
	TATCTTTACG TTCTTGTTTT TATTATAAAT CCGAATAAAC CTATCACTAT TACGCACACC	1020
	AAAATATTTT GTTCTGGTT TTACATTACG TCCATAAAAT ATAGTTTCTT TTACCGACTT	1080
25	ATCTGACAAT GCATAATAGT CATTTAAATC AAATTCAAAA TCAAAAGCCA AATCTAATCT	1140
	CGTAAACTA ACATCGTCCA AATAACTGAT GATATTTTGT TTTAACCAAA GCACTTCATC	1200
30	ATGCGAAAGC TTATTAGGAT TAAATTCAAC GCGCATATAC GTCTATTCCA AAGAGTTGCT	1260
	TTTATTTTGT CATATTCAAT ATAACTTTT TCTTTAAGAG CTTTAGCTTT AAAGTTTGTT	1320
	TGTAAATAT CCCAAGCCG AATTTCAAGG TTAGTACTCA TAAATGTGA AAGTCTCTCT	1380
35	GCGTTAGACA TGCTAAGATT CCCAACAATC GTTATAGCGT CAAAAGACAA TTTTGGAATA	1440
	GCTAGTGACA TCCTATGTCG ATTTAACCGG CTATTACCGG ATATTAGAGT ATCCAGTTTT	1500
	ACAAATGGAT GAAACGAAAT TCAAAACACT AAAAAATATG TTCCACTAAC AGCAAAAAAA	1560
40	TACCATTATG TTCCTACTAA AAAACyAAAA ATACTGGAGA ACAAATGTCA GGATATAACT	1620
	TAGGATACTA TGTAATAAAA ATTTACAATA AAAAAACAGG AAAACAAATT TCAAGTAAAA	1680
	GmATACCCAT ACAAAGAGGA TAAATAAAA AACCTCGAAC TGaAATGATG ATCTTTTCAG	1740
45	CTCGAGGTTT AAATATTGGT GCCTTATTTA TATAGATTCG TTATATTATA TTCTCTATTT	1800
	TCATTAACmT AATCCTTAAA GAGTTTTAAA TTAATACCTG CTAGATGATT CAAAAATGTT	1860
50	TCATCAACTT TTAAATAATT CAATAATTTT TGTGGTGTCA GTAAATnTCT ATCAAAATAC	1920
	AACTTTAATA AACTATTCTT TTTGACAGGA CGTGACATTT CAATCACGTC GTCTAAAGAT	1980
55	AATACTTTCT CGCTTTAnAC AAAnACAAAA ACTTACCCGA TTAAATCAA GTAAGTTTTA	2040

	TATTTGATAA AAAATCAATA AGTAATTGTG CGCCTTCAAC TTGAATATCT TTTACAACTG	2160
	GCGCGTCGAT ATACATATCA TACTGACCAC CGCCTACTGC ACGATAATTA TTTACACAAA	2220
5	TTGTATATGT CTGCTTTAAA TCAACTGCGT GACCTTGAAT CATCATATTG CTCACACGTT	2280
	GTCCCTTTGG TCTTCCAACA TGAATGGTAT AACTTACGCC ACCATATATA TCATAATTAA	2340
10	AGTGTGTGTG TTTGGGTTCA AGGAAGTCTG CGCTCACACT AACTTCATCA TTTTTCACGT	2400
	CAAAATATTC TGCTGATCGT TCAATGGCTT CTTTAAAGTT GGCACCACTT ACAGCTAAAA	2460
	CTTTAAATGT APTTGAAAAT GGGTAATTGT TAATAACATC TCGCATCGTC ACGACTTGCT	2520
15	TGAAACCACT AGCAGAATCA AACAAAGCTG TACAGGCAAC ATCTGCGTCA CTTTTTTCTA	2580
	ATAAAGCGTA ATTCATAAAA TTTGTAAAAG GATGCGGTGC CACACGTGCC TCAAATGCAT	2640
	GATTAATCGT CATATCATAT GGCAATGTAG TAATTCGTA ATCTAACCAG TCCTCTAACT	2700
20	GCTTTCGTAA ATGTTGGTCA TCTTCATCAA TAGTAAATGT GGAATCATCT ATAACAGGAA	2760
	GTAATTCACA TGATTCAACG GATAGATTTT CATATTCATC AGTACTCAAG ACTACTCTGC	2820
25	CTACAGTTGT ACCTCTCGTA CCAGGTTGAA TCACAGCCGT TTGCTTAAAC CTTTCAGCAA	2880
	TTTGTCGATG TTGGTGACCC GTAATAAAGA TATCTATATC TTTAGAAAAC GCTTCTAACA	2940
	TGGCATATCC TTCATTTTCA CCCGTTAATA CTTCGGTCGG CGTACCACTT TCTAAATCCT	3000
30	TTTCAAATCC ACCATGGTAA CAAACCACAA TGATATCTGC ATGTCGCTTC ATTTCAGGTA	3060
	AGTATTGTTG AAGTATTTCA AAAGCACTAT GAAACGTAT GncnTGAATA TGCTCTGGTT	3120
	GTTCCCAATG GGAATAAAT TGTGTCGTTA AACCTATCAC ACCAACAGTT TGATCTCCAA	3180
35	CCTGAAAATA CTTACACCCG TTATCAGTCA ATGTACTATC ATTTTCATAT ATATTAGCGC	3240
	ACAAACTGG ATAATTGAGT CTGCGTAAAG TGTCTTTTAA GTATGGTAAT CCATAATTAA	3300
	ATTCATGATT ACCAAGCGTA CCAAAGTCGA ATGCCATTG ATTATAAAAA TCAACTAAAG	3360
40	GCTGGCTACT GCCGCTATGC GCGATTAAGT AATTACAAAA TGGTGACCCT TGCAAAAAAT	3420
	CACCATTATC TATTTTAAAA CTTTGGTCAT ACTGCCTTCT GTsTTGTTCT ATAACATGAT	3480
45	TCGCTAGTAA CAATCCCAT GGTGATATT GATTTCCTACT CGTAAAATCT GTTGGGAAAA	3540
	TATAACCATG TACGTCACTC ACGACATAAA ATGCTATGTT TGACATCCTC ACTCACTCCT	3600
	TCAATCACAA ACATCTTTCT TATTTCTATT ATATATTTAT TTGAAGTCTG TTGTAATCAA	3660
50	GGTTTTGTCA CCGAGTTTTA AACGAATCTT TGAACCTTCC ATACTTTCAA GTACTTTAGC	3720
	ATTGACCTTA ATTGTGACAT TTCCGTTTTT ATCTGCTTTA ACTGTTGGCA AAGTACTGTA	3780
55	ACCTGGTGGG TTATAATCGT TATCTTTACT TGAAAATTGT CCGATTGAC GTCCGCCTTC	3840

	TATTGTCATT TCAAATGGCT CATTACAGA AACATTTTGC GGGATATCAA ATGTTACTTT	3960
	TTCGTTCTGA TTTGGTGGTG TATGATCATC TGGTGTGTTT GGCTGAGGAT CTGCGCCTTT	4020
5	TTGCTGCCA TAACTACCTG CTTTAAATGT TGTGGATCA TACCATTAT AACCACCTCG	4080
	CGGTTGTGAC CATGGCTCTT TTTCAGGCTC AGTTGAACGC TCTGGTCGTT CAAAATCAAG	4140
10	CAACTTAGTC TTTGTATCTA ATGTTAGGCT ACTCGCCTTA AGTGATTTC CATCATTATC	4200
	TTTAGACATC CAAGCCGTTA TATTATTTAA TAGCTTACCG TTGTCTTGTT CTTTAAACC	4260
	ATCATATGTT TTCTTCTTTT CTCCATTATC TTCTCTTACA TATTTGGGCG AACTATCTTC	4320
15	CACAAGTGAT GAATCACCGA TAAATGCTGC TTTACCTTTT CCAACTTTAG AAATTGCTAC	4380
	ATAGGGGCGT TCTGCTTTAC CGCCCCATT ATAAATACCT TGATCTACAG CATGTGACCA	4440
	TTTACTTTTC GCTGGCAATT GTTCTGGTGT ATACACAATA CCTTTTGCTT TCTCTGGATT	4500
20	AGTAATTGCT AATGTCGATC CGGCATGCAT AGAGACAGAT TTCACACCTT CAGTAATACC	4560
	GAACTTTCT TTTGAAGAAA CAATATTGCT CGTATTTAAA TCACCTAGTG CATTATATCG	4620
	AAAACGTACG CCAAAGTTTG TAGATAACCA ATCTGAACTT TTCACACCTT GCATTGCAGT	4680
25	AGAACTTTT TCTTCTGCAT TCATACCTT CGACATATCT TCATATGCTC CACGTCGATA	4740
	ACCATTCAAT GCCTCCGATG AATCAATACG ATTTAAATTT CGGTCAGCAT TGTAATGATC	4800
30	TGAAATAAAG ACAACATTGC CACCTTGTT CACATATTTA ACAATTGCTG CCTGTTCTGA	4860
	TTCTTTGAAA GGAATGTTAG CCTCAGGAAT TACAAATATT TTGGAACCTT TCAAACCTGC	4920
	TTCTGTTATG TTCGAATGAC CATCAATAGC TTTAACGTCA TAACCTTGTT TTTGTATTGA	4980
35	ATCCGCATAA TCTGAAAATG CACCATCACT AACCCTTCT GCAGCACCAG CTGTTTGACC	5040
	ATGAGAACGA TCGAATAATA CCGTTCGCTG TTGCTTTGTA GGTGCGATT CATGCGTTAT	5100
	AGCTAAAGAT TGCGGTAAAG CACTTAATGA TACCGTTGCA ACAATTGCAG AGACAGTTAA	5160
40	TGACTTATAT ATTTTCTTCA TTTTGTGAGG CTCCTTTTAA AATAAATTTG TTCTTGAATT	5220
	ATAGGATAAA AATTCGTTGC ATATGAGCAA TTTAACGAAA AATTTACAAA ATCTTATCAA	5280
45	ACTCTTAAAG AAAGTTATTA AAATTCATTT TTATAAATA CTTTTTAACA TTAAATGTG	5340
	GTACGCTATA AGTGTAATTT CATTGCATAC ATATTACAG ATTAAGAATG TGAAGGGGAC	5400
	AGTTATCAAA TGAAAAATTT TAAGTGTTTA TTTGTATTAA TGTTAGCAGT CATTGTTTTT	5460
50	GCAGCAGCAT GTGGAACTC AAGTTCTTTA GATAATCAAA AGAACGCTAG TAATGATTCTG	5520
	GATTCTAAAT CAGGAGGATA CAAACCTAAA GAATTAACCG TTCAATTTGT ACCTTCGCAA	5580
55	AATGCTGGAA CATTAGAAGC TAAAGCAAAA CCATTAGAAA AATTACTATC TAAAGAATTA	5640

	TCTAAAAAAG TTGATGTTGG TTTCTTACCA CCAACGGCAT ACACATTAGC ACATGATCAA	5760
	AAAGCAGCTG ATTTATTATT ACAAGCACAA CGTTTCGGTG TAAAAGAAGA TGGTTCAGCA	5820
5	AGTAAAGAAC TTGTAGATAG TTATAAATCA GAAATTCTTG TAAAAAAGA CTCAAAAATT	5880
	AAAAGCTTGA AAGATTTAAA AGGTAAGAAA ATTGCCTTAC AAGATGTAAC ATCAACTGCT	5940
10	GGATATACAT TCCCACTTGC GATGTTAAAA AACGAAGCAG GTATTAATGC AACTAAAGAT	6000
	ATGAAAATTG TGAATGTAA AGGTCATGAC CAAGCAGTTA TCTCATTATT AAATGGAGAT	6060
	GTAGATGCTG CGGCTGTATT TAACGATGCA CGTAATACTG TGAAAAAAGA CCAACCAAAT	6120
15	GTATTTAAAG ACACACGAAT TTAAAAATTA ACACAAGCTA TTCCGAATGA CACAATTTCT	6180
	GTAAGACCAG ATATGGATAA AGATTTTCAA GAAAAATTGA AAAAAGCTTT TATAGACATT	6240
	GCTAAATCAA AAGAAGGTCA CAAAATTATT AGCGAAGTTT ATTCACATGA AGGATACACA	6300
20	GAAACGAAAG ATTCAAATTT CGACATTGTA AGAGAGTACG AAAAATTAGT TAAAGATATG	6360
	AAATAATCAT TATTTAACAA ATGAATCATT AGCGAATTTG GTATTAAAAG CTTTCGTTCA	6420
25	ATAGATATAT TCTAGATTAA TATTGAAAAG CTAGGCGCTA AACTGAAACA GATATAGAAA	6480
	GGTGTGCTG TACATTTGAA ACCATTTGTA CACAGAAACC CAATGTCTAT GATATTTTCA	6540
	TTTACCTTGG CTTTTCTTTA TTAAAGAAAG GTGTCAAACA TGAGTCAAAT CGAATTTAAA	6600
30	AACGTCAGTA AAGTCTATCC TAACGGTCAT GTAGGCTTGA AAAATATTAA CTTAAATATT	6660
	GAAAAAGGTG AATTTGCGT TATTGTGCGA CTATCTGGTG CTGGGAAATC CACGTTATTA	6720
	AGATCTGTAA ATCGTTTGCA TGATATCAGC TCAGGTGAAA TTTTCATCCA AGGTAAATCA	6780
35	ATCACTAAAG CCCATGGTAA AGCATTATTA GAAATGCGCC GAAATATAGG TATGATTTTC	6840
	CAACATTTTA ATTTAGTTAA ACGGTCAAGT GTATTACGAA ATGTACTAAG TGGACGTGTA	6900
	GGTTATCACC CTAATTGGAA AATGGTATTA GGTTTATTCC CAAAAGAAGA CAAAATTAAG	6960
40	GCAATGGATG CACTAGAACG CGTCAATATC TTAGATAAAT ATAATCAACG CTCTGATGAA	7020
	TTATCAGGTG GCCAACAACA ACGTATATCT ATTGCACGTG CGCTATGCCA AGAATCTGAA	7080
45	ATTATCTTG CAGATGAACC AGTTGCTTCA TTAGACCCAT TAACTACGAA ACAGGTTATG	7140
	GATGATTTAA GAAAAATCAA CCAAGAATTA GGCATCACAA TTTTAATTAA TTTACATTTT	7200
	GTTGACTTGG CAAAAGAATA TGGCACACGC ATCATTGGTT TACGTGATGG TGAAGTTGTC	7260
50	TATGATGGTC CTGCATCTGA AGCAACAGAT GACGTATTTA GTGAAATATA TGGACGTACA	7320
	ATTAAAGAAG ATGAAAAGCT AGGAGTGAAC TAACATGCCT TTAGAAATAC CTACAAAGTA	7380
55	TGACTCCCTT TTAAAGAAAA AGGTTTCTTT AAAACGAGT TTTACCTTCA TGTTAATCAT	7440

	AATACCTCAA ATAGGTGATC TATTCAAACA AATGATTCCA CCTGATTTTCG AGTATTTACA	7560
	ACAAATTACA ACGCCAATGT TAGATACCAT TCGAATGGcT ATCGTAAGTA CAGTATTAGG	7620
5	TAGCATCGTT TCAATACCAA TTGCGTTATT ATGTGCTAGC AATATCGTTC ATCAAAAGTG	7680
	GATTTCAATA CCCTCGCGCT TTATTTTAAA TATAGTTCGT ACTATTCCAG ATTTGTTATT	7740
10	AGCAGCAATC TTTGTGGCTG TATTTGGAAT CGGTCAAATT CCAGGGATAT TAGCACTGTT	7800
	TATTTTAACT ATCTGTATTA TTGGAAAATT ATTATATGAA TCATTGGAAA CGATAGATCC	7860
	AGGTCCAATG GAAGCAATGA CGGCTGTTGG CGCTAATAAA ATAAAATGGA TTGTTTTTCGG	7920
15	TGTTGTACCA CAAGCCATAT CGTCATTTAT GTCATACGTA TTATATGCAT TTGAAGTAAA	7980
	TATACGTGCT TCAGCTGTGC TTGGATTAGT CGGCGCTGGC GGTATTGGAT TGTTTTATGA	8040
	TCAAACACTT GGTTTATTTT AATATCCAAA AACAGCAACG ATTATTTTAT TTACTTTAGT	8100
20	TATCGTCGTC GTCATTGATT ACATCAGTAC GAAAGTGAGG GCACATCTCG CATGACACAG	8160
	GAAATAGCAA AATATAATGT TCACACAAAA GCACACAAAC GAAAATTGAT TAAAAGATGG	8220
	CTTATTGCAA TTGTCGTCTT AGCTATTATC ATCTGGGCAT TTGCAGGTGT ACCAAGTTTA	8280
25	GAACITAAAA GTAAATCATT AGAAATCTTA AAATCCATAT TCAGCGGATT ATTCCATCCT	8340
	GATATCAGCT ATATCTATAT ACCAGATGGC GAAGACTTAT TACGTGGTTT ACTTGAAACC	8400
30	TTTGCGATAG CCGTTGTAGG TACTTTCATC GCCGCAATTA TCTGTATTCC ATTAGCATTT	8460
	CTAGGTGCAA ATAATATGGT AAAGCTACGC CCAGTTTCAG GTGTTAGCAA ATTTATTTTA	8520
	AGTGTTATAC GTGTCTTCCC AGAAATTGTA ATGGCACTTA TATTATCAA AGCTGTTGGC	8580
35	CCAGGTTTAT TTTCAGGTGT ATTAGCTTTA GGTATCCATT CCGTAGtATG CTTGGGAAAC	8640
	TTTTAGCTGA AGATATTGAA GGTCTAGATT TCAGTGCTGT AGAATCATT AAGGCCAGTG	8700
	GTGCGAATAA GATTAAAACA CTCGTATTTG CAGTCATACC ACAAATTATG CCTGCCTTTC	8760
40	TATCACTCAT ACTTTATCGC TTGAACTAA ACTTACGTTT AGCTTCTATA CTGGGGCTAA	8820
	TTGGGGCTGG TGGTATCGGG ACACCACTCA TATTTGCCAT TCAAACACGT TCTTGGGACC	8880
	GTGTAGGTAT TATATTAATC GGTTTAGTAC TAATGGTCGC AATTGTCGAT TTAATTTCCG	8940
45	GTTCAATCCG AAAACGTATT GTTTAACATT AAATCAGGAT ACTCCTAAAT AAGAAGTCCT	9000
	ACCGTCTTAC GTTCTCTAT TATAATAAAA ACAGCAGTGA AGAAACTAT TGTATAGTT	9060
50	AACTTCACTG CTGTTTTTAT AATATCTAAA TTTATTCTAT TTCAATTCCT TTAAATAACT	9120
	TTTACCGAAC TCTGGTAATG TTACGTTGAA ATTATCTGCT ATAGTTGCAC CGATAGAACT	9180
	GAATGTAGTA TCACTTTCTA GTGCATGACC ACCTTTAAAT TTCGGACTGT ACATAATTAC	9240

55

	TGTAATAATT ACTAAATCGT CTTCTTTTAA GTTGCTAAAC AGTTCTGGCA AGCGATCATC	9360
	GAAATCTTTA ATTGCTTG TG CATAACCTGG TTTATCACGA CGATGACCGT ATAATGCATC	9420
5	AAAGTCTACT AAGTTTAAGA AGCTAATACC TGTGAATCT TTCTTAACAA TTTTCATCAA	9480
	TTGATCCATA CCGTCCATGT TACTCTTCGT ACGAACCGCT TCTGTTACAC CTTACCATC	9540
	ATAAATGTCA TTAATTTTAC CGATGGCAAT AACATCATAA CCACCGTCTT TCAAATGATC	9600
10	TAAGACAGTT TTACCAAAG GTTTTAACGC ATAGTCATGT CGATTAGATG TACGTGTAAA	9660
	GTTTCCTGGT TCACCAACAT ATGGACGTGC GATAATACGA CCAATTAAAT ATTTAGGGTC	9720
	TTTTGTCAAC TCACGAACCT TTCACAAAT ATCATATAAC TCTTCTAATG GGATAATGTC	9780
15	TTCATGTGCA GCAATTTGCA ATACTGGGTC TGCACCTGTA TAAACAATTA AGTCACCAGT	9840
	TTTCATTTGG TGCTCGCCCC ACTCATCGAT AATTTGCGTA CCCGATGCCG GTTTGTTAGC	9900
20	AACAACTTTA CGACCTGTCA TTTCTTCAAT TTGTTGAATT AACTCTTCAG GGAATCCATT	9960
	AGGGTATACT TTAAGAGTT GCATAATATT TAATCCCAT AATTCCAGT GACCAGTCAT	10020
	TGTATCTTTA CCAACTGAAG CTTCACTCAA TTTAGTATAG TATGCTTCTG GTTGTTC AAC	10080
25	TGCATTTACT ACTGGTAATT TATCGATGTT CCCTAGACCT AACTTTTCAA GGTTTGGTAA	10140
	AGTTTGATCG AAACCTTCTA AGGTATGTCT TAAAGTATGT GAACCTTCAT CTTTAAATC	10200
	AGCTGCGTCT GCGCCTTCAC CAATACCTAC TGAATCCATT ACGATTAAAT GTACACGATT	10260
30	AAATGGTCTT GTCATAGCTA TCACTCCCAA AATTTATATA TATTAGTAAT CTGAATCTGC	10320
	TTCTAAACCT TGCATAATTT GAACACCTGC GCTCGACCA ATACGTGTCTG CACCTGCTTC	10380
35	AACCATTTTA TTGAAATCTT CTAAATTACG TACGCCACCT GATGCTTTTA CTTCTACATC	10440
	AGCACCTACT GTATCTTTCA TTAATTTAAC GTCTTCTGCA GTCGCACCGC CACCTGCAAA	10500
	ACCTGTTGAA GTTTTAACGA AGTCCGCACC AGCCGCTTTT GTTAATTCAC TCGCTTTTAC	10560
40	AATTTGCTCA TGGTCCAACA ATACCGTCTC AATAATCACT TTTACTGTGT GACCTTTTCG	10620
	AGCTTTAACC ACTGCTTCAA TGTCTTGTG TACATCATCA AAACGTCCAT CTTTTAATGC	10680
	GCCGATGTTG ATGACCATGT CAATTTTCATC TGCACCATTT TGAATTGCAT CTTCTGTTTC	10740
45	AAATGCTTTC GTTGCAGTTG TCGACGCACC TAATGGGAAT CCTATTACCG TACAAACGAG	10800
	CACCTCTGAA TCAGCTAGTC GCTCTGCTGC ATATTTAACA TGTGTTGGAT TCACACATAC	10860
	AGATTTAAAA TTGTATGCTT TCGCTTCATC GATGATTTGA TCGATTTGCG TACGTGTTGA	10920
50	CTCAGGCTTC AATAAAGTGT GATCTATATA TTTCTCAAAT TTCATACTTA CTACTCCTCG	10980
	TGTTATATAA TCTCTTTATT TAATTTTACT ATAAATACGA ATATATCTCG CGAATTTATA	11040
55		

5 ATACTCATTAA AACCTAAAAT AATTAAAATA ATACCGAAAT GTGAACTTAA TGCATCATTG 11160
 CCTGGGAAAT TTAATGCTTT AAAATCGATT AGAGCCGCAG CAATCGCAAT ACCTACAGAT 11220
 ACCGCCACAT TAATAATTAA ATTATAAAAA CCAATAGCCA CACCTGTCAT ATTAAGATCT 11280
 ATTGTTTTTAA TGGCTTCGTT AAGTAAAGGT GCATACATTA AAGCAAAGCT ACCTGCAAAG 11340
 10 AATATCATAG AAATGACGAA GATTGAAATG TGATTACCTA CTGCAAATGC AGGTAAAATC 11400
 AAGCTCAGTG CTATTAAAAT AATTGCTGTG ATAATCGCTT GTTTTGAATT CAGATATTCTG 11460
 CCGATTTTAC CACTTAGTGC ACCAACAATG ACTGCTACTA TATAACCCGG TACTAATAAC 11520
 15 AGTGATGTTG TGTCTAGTTG CAGATGATAA ATTTGCTCCA TTATGAATGG GAACGTAAAA 11580
 ATATAACCCA ATGGGATAGC ATACATTACA AATACTATAA ATAAAAATGA AGCATAACGT 11640
 TTATTTTGA AAAATGATTT ATTTACTAAT GGACGTTGCG CATTTTAAAT ATATAGCGCA 11700
 20 AAAACGATAA TCGCAATTAA GGCACCAATC ATATATAACC AATTAAAGTT CGTAATAAAC 11760
 AGCATGACTG TTGTAGCAGG GGATCCTCTA GAGTCGAnCC TG 11802

(2) INFORMATION FOR SEQ ID NO: 71:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1196 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

35 CTAAAGAAGA TCGAAACAA GATGTTGATA AACAAGTTCA AGCTTTAATT GACGAAATCG 60
 ATCAAAATCC AAATCTAACA GATAAGGAAA AACAAGCACT TAAAGATCGT ATTAATCAAA 120
 TACTTCAACA AGGTCATAAC GACATTAACA ATGCGATGAC AAAAGAAGCA ATTGAACAAG 180
 40 CAAAAGAACG TTTAGCGCAA gCATTGCAAG ACATCAAAGA TTTAGTGAAA GCTAAAGAAG 240
 ATGCGAAAAA TGATATTGAT AAACGTGTAC AAGCTTTAAT TGACGAAATC GATCAAAATC 300
 CAAATCTAAC AGATAAGGAA AAACAAGCAC TTAAAGATCG AATTAATCAA ATACTTCAAC 360
 45 AAGGTCATAA CGACATTAAC AATGCGCTGA CTAAAGAAGA AATTGAGCAG GCAAAAGCAC 420
 AACTTGCACA AGCATTGCAA GACATCAAAG ATTTAGTGAA AGCTAAAGAA GATGCGAAAA 480
 50 ATGCAATAAA AGCCTTAGCT AATGCGAAGc GTGATCAAAT CAATTCAAAT CCAGATTTAA 540
 CACCTGAGCA AAAAGCAAAA GCGCTCAAAG AAATTGACGA AGCTGAAAAA CGAGCACTAC 600
 AAAACGTTGA GAATGCTCAA ACTATAGATC AATTAAATCG AGGATTAAAC TTAGGTTTAG 660

TTGAAGCAAC ACCTGAGCAA ATCCTAGTTA ATGGTGAAC TATTGTACAT CGTGATGACA 780
 TCATTACAGA ACAAGATATT CTTGCACACA TAACTTAAT TGATCAGCTT TCAGCAGAAG 840
 5 TCATCGATAC ACCATCAACT GCAACGATTT CTGATAGCTT AACAGCAAAA GTTGAAGTTA 900
 CATTGCTTGA TGGATCAAAA GTGATTGTTA ATGTTCTGT AAAAGTTGTA GAAAAAGAAT 960
 TGTCAGTAGT CAAACAACAG GCAATTGAAT CAATCGAAAA TCGGGCACAA CAAAAGATTA 1020
 10 ATGAAATCAA TAATAGTGTG ACATTAACAC TGGAACAAAA AGAAGCTGCA ATTGCGAAG 1080
 TTAATAAGCT TAAACAACAA GCAATTGGAT CATGTTAAC AATGGCACCT GGATGTTCCA 1140
 15 TTCAGTTGAA GGAAATTTCA ACAACAAGGA ACAAGCGCCG GATTGGAACA ATTTGA 1196

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1519 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

CAATCGTTTC AACGCTATTA TCTTTAGACA ACAATTGTAA GCGTGTATGT GCAGTTTCTA 60
 AACAGTCTAT AATTCGAGTT CTTAATTCAG CTGGATCATC TTTAAAAATA AAATCCATCG 120
 30 CTGCAACTTT GTAGACAAAT GTTAAATAGG TAAGTTCAC TGTACTCGTA ACGAAAATAA 180
 TGTTACCAAC TGGGTCATGC TTACGAATTT CACTGCCTAA TTTGATACCA TTAATATCAG 240
 35 TTGAAAGTTG AATATCTAAA AAGTAACAGC CTATGTCATT CATATTTTTA GCTTGCTCAA 300
 GCACCTCATA AGGATTATCA GTTGCAGAGG CAATTTCCAT AGGCTTTTCT TCTATCATT 360
 TATAATTTTT AATAATGGTA ACCATGTTTT CTCTTTGTTT TGGATCGTCT TCGCAAATGA 420
 40 AAATTTTCAT ACATTCACAT CCTTATGGCT AGTTGTTAAT AATTTCAACT TTTTGAATAA 480
 AGAAACCATT TTCGATAATT GTATCTAATA AGACATTGTC TGCATTATCA GCAATTTCTT 540
 TTAAAGTTGA TAGACCTAAA CCACGACCTT CACCTTTAGT AGAAAACTT TCTTGAACA 600
 45 ATTCATGAAT GCGTGGTATA TCATCAGCGC ATTTATTCAT AACAATAAAC GTTACTGAAT 660
 TTTCACTTTC AATAAATGCA ACGCGAATGA TAGGGTCATC AATTTCAGTT GATGCCTCAA 720
 TTGCATTATC AAGAATAATA CCAATACTGC GACTTAAATC GATCATATTC AAGTTAATGC 780
 50 TACTTACTTC ATCGGGTATT TCGATACTAA TCGGAATATT CATTTCTTGT GCACGTAAAA 840
 TTTTCGCAGT AATTAAGCCT TTAATTTTAC GTACTTTAAG ATTCTCGATA CCATTTAATT 900

55

GTAGGCCAGG CATGTCATCT TCTCGAATGT ATTCTGAAAG TGTCGTTAAG ATATTGACAT 1020
 AATCATGACG GAACTTGCGC ATTCGTTGT TGATAGCTTC AATCTTCAAT GTATATTCAT 1080
 5 AATAGGTTTC AATTTCTTCT TGATTACGTT TATATTTTCAT CTCTTTAAGG AGAAATTGAG 1140
 AAATAACAAA TGTTAATATA CTAAAAATA TAGTGATACC AATAAAAATA AAAGAATACT 1200
 GCCTTATTAC TTTAGCTTCA TCCGAGTTTA TTTGTGAATA AAAGAAAAAT AATGAAAAAG 1260
 10 TAAGCAGTAA GATAGTCGAA ATAACTATTA AAAATCCTTT GTTTAGTATT AGATATGGTG 1320
 TGCTAATTTT TTTGAGAACT CTATTTATTA TATATGAGAA TAGTATACTA ATAGTCACAT 1380
 AAATAACAAA AAAGCTAGGG AATATTACAA ATATACTATC AGAAATTTTG GTGGATATAT 1440
 15 GCATATATAA CTATATACCT GTAGTTAGCA CnGThATAGG AATAATCnGG CGAGGTCCAT 1500
 AATCCACCAA AATAGAATA 1519

20 (2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5445 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

30 GTAGGAATCT CTTTGTCTTT TTGGGAGGAC ATTTAATATG AATGTATATT TAGCAGAATT 60
 CCTAGGAACT GCAATCTTAA TCCTTTTGG TGTTGGCGTT TGTGCCAATG TCAATTTAA 120
 GAGAAGTGCT GCGAATGGTG CTGATTGGAT TGTCATCACA GCTGGATGGG GATTAGCGGT 180
 35 TACAATGGGT GTGTTTGCTG TCGGTCAATT CTCAGGTGCA CATTTAAACC CAGCGGTGTC 240
 TTTAGCTCTT GCATTAGACG GAAGTTTTGA TTGGTCATTA GTTCCTGGTT ATATTGTTGC 300
 40 TCAAATGTTA GGTGCAATTG TCGGAGCAAC AATTGTATGG TTAATGTACT TGCCACATTG 360
 GAAAGCGACA GAAGAAGCTG GCGCGAAATT AGGTGTTTTT TCTACAGCAC CGGCTATTAA 420
 GAATTACTTT GCCAACTTTT TAAGTGAGAT TATCGGAACA ATGGCATTAA CTTTAGGTAT 480
 45 TTTATTTATC GGTGTAAACA AAATTGCCGA TGGTTTAAAT CCTTTAATTG TCGGAGCATT 540
 AATTGTTGCA ATCGGATTAA GTTTAGGCGG TGCTACTGGT TATGCAATCA ACCCAGCAGG 600
 TGATTTAGGT CCGAGAATTG CACATGCGAT TTTACCAATA GCTGGTAAAG GTGGTTCAAA 660
 50 TTGGTCATAT GCAATCGTTC CTATCTTAGG ACCAATTGCC GGTGGTTTAT TAGGTGCAGT 720
 GGTATACGCT GTATTTTATA AACATACATT TAATATTGGT TGTGCAATTG CrATTGTTGT 780

55

	CGAATCAATT TACTAAAATA AAAAGAAACG TAAATAGCAT AATTTAACAT GTTTGATTCA	900
	TGGATTATGC TATTTTTTCG CCAAAATTTA ACAGATTTTG TACAATGGGT TAGCGATTAT	960
5	TTTTTAATAA AGGAGATACT ACTAATGGAA AAATATATTT TATCTATAGA CCAAGGAACA	1020
	ACAAGCTCAA GAGCGATTTT ATTCAATCAA AAAGGGGAAA TTGCAGGGGT AGCACAACGT	1080
	GAGTTTAAGC AATATTTTCC ACAATCAGGT TGGGTGTAAC ATGATGCAAA TGAAATTTGG	1140
10	ACATCTGTGT TAGCTGTAAT GACGGAAGTA ATTAATGAAA ATGATGTTAG AGCTGATCAA	1200
	ATTGCAGGTA TCGGTATTAC AAACCAACGT GAAACAACGG TTGTTTGGGA CAAaCATACT	1260
	GGCCGCCCAA TTTATCACGC AATTGTTTGG CAATCACGTC AAACACAATC AATTTGTTCA	1320
15	GAATTA AAC AAGGATA TGAACAAACA TTTAGAGATA AGACAGGATT ACTTTTAGAT	1380
	CCGTATTTTG CAGGTACAAA AGTTAAATGG ATTCTAGACA ATGTTGAAGG TGCACGAGAA	1440
20	AAAGCAGAAA ATGGCGATCT ATTATTTGGA ACGATTGATA CTTGGTTAGT ATGGAAATTA	1500
	TCaGgAAAag CtGCGCATAT TACTGATTAT TCaAATGCGA GTCGTACATT AATGTTTAAT	1560
	ATCCATGATT TAGAATGGGA CGATGAGTTA TTAGAACTAT TACAGTACCT AAAAATATGT	1620
25	TGCCAGAAGT TAAAGCTTCG AGTGAAGTAT ATGGTAAGAC AATTGATTAC CACTTCTATG	1680
	GTCAAGAAGT ACCAATCGCT GGAGTAGCTG GTGATCAACA AGCAGCATTa TTTGGACAAG	1740
	CTTGCTTCGA ACGTGGTGAC GTGAAAAACA CATATGGAAC TGGTGGCTTC ATGTTAATGA	1800
30	ATACAGGTGA CAAAGCGGTT AAATCTGAAA GTGGTTTATT AACACAATT GCTTATGGTA	1860
	TTGATGGAAG AGTAAATTAT GCGCTTGAAG GTTCCATCTT TGTTTCGGGT TCAGCAATCC	1920
	AATGGTTACG TGATGGATTA AGAATGATTA ATTCAGCACC ACAATCAGAA AGTTATGCGA	1980
35	CACGAGTTGA CTCTACTGAG GGTGTTTATG TTGTTCCAGC TTTTGTAGGT TTAGGAACAC	2040
	CATAATTGGGA TTCTGAAGCA CGTGGTGCGA TTTTCGGTTT AACACGTGGA ACTGAAAAAG	2100
40	AGCACTTTAT CCGTGCAACT TTAGAATCAC TATGTTACCA AACTCGTGAC GTTATGGAAG	2160
	CAATGTCAA AGACTCTGGT ATTGATGTCC AAAGTTTACG TGTCGATGGT GGTGCAGTTA	2220
	AAAATAACTT TATTATGCAG TTCCAAGCAG ACATTGTTAA TACTTCTGTT GAAAGACCTG	2280
45	AAATTCAAGA AACTACAGCT TTAGGTGCTG CATTTTTGGC AGGTTTAGCA GTTGGATTCT	2340
	GGGAGAGTAA AGATGATATC GCTAAAACT GGAAATTAGA AGAAAAATTC GATCCGAAAA	2400
	TGGATGAAGG CGAAAGAGAA AAATTATATA GAGGTTGGAA AAAAGCTGTT GAAGCAACAC	2460
50	AAGTTTTTAA AACAGAATAA ACTTGTAGAT TAGACTTTTG TATAAACATT GTGATACAAT	2520
	CAATTTAAGT TAATATTTGA ATCGAGAAGC GAGAGATTTG TTCGAACATG TACAATTGAA	2580

55

	GCATTGTCTA CTTTAAAGAG AGAACATATT AAAAAGAATT TAAGAAATGA TGAATATGAT	2700
	TTAGTAATTA TTGGTGGCGG TATTACAGGT GCAGGTATTG CACTAGACGC GAGTGAAAGA	2760
5	GGAATGAAAG TTGCATTAGT TGAATGCAA GACTTTGCAC AAGGAACAAG CTCAAGATCT	2820
	ACAAAATTAG TCCATGGTGG TTTGCGTTAC TTAAACAAT TCCAAATTGG AGTAGTTGCC	2880
	GAACTGGTA AAGAACGTGC GATTGTTTAT GAAAATGGGC CTCATGTTAC GACTCCAGAG	2940
10	TGGATGCTTT TACCAATGCA TAAAGGTGGA ACATTGGTA AATTCTCAAC ATCAATTGGT	3000
	TTAGGAATGT ATGATCGTTT AGCAGGTGTT AAGAAGTCTG AACGTAAAAA AATGTTATCT	3060
15	AAAAAGAAA CTTTAGCTAA AGAACCTATA GTTAAAAAG AAGGTCTAAA AGGCGCGGGT	3120
	TACTATGTTG AATATCGTAC TGACGATGCG CGTTTAACTA TTGAAGTTAT GAAGCGTGCT	3180
	GCTGAAAAAG GCGCAGAAAT TATCAACTAT ACTAAATCTG AACACTTCAC TTATGATAAA	3240
20	AATCAACAAG TAAATGGTGT TAAAGTTATA GATAAATTAA CTAATGAAAA TTATACAATT	3300
	AAGGCTAAAA AAGTGGTTAA TGCAGCAGGT CCATGGGTTG ATGATGTTAG AAGTGGTGAT	3360
	TATGCACGCA ATAATAAAAA ATTACGTTTA ACTAAAGGTG TACATGTTGT TATTGATCAA	3420
25	TCAAAATTCC CATTAGGTCA AGCAGTATAC TTTGATACTG AAAAAGATGG AAGAATGATT	3480
	TTTGCAATTC CACGTGAAGG AAAAGCGTAT GTAGGTACTA CAGATACATT CTATGACAAT	3540
	ATCAAATCTT CACCATTAAAC TACACAAGAA GACAGAGACT ATTTAATCGA TGCGATTAAT	3600
30	TACATGTTCC CTAGTGTTAA TGTTACAGAT GAAGATATTG AATCAACATG GGCAGGAATT	3660
	AGACCATTAA TTTACGAAGA AGGCAAAGAC CCTTCTGAAA TCTCTCGTAA GGATGAAATT	3720
	TGGGAAGGTA AATCAGGTTT ATTAATAATT GCAGGTGGTA AATTAACAGG CTATCGTCAC	3780
35	ATGGCTCAAG ACATTGTTGA TTTAGTATCT AAACGCTTGA AAAAAGACTA CGGTTTAACA	3840
	TTTAGTCCAT GTAATACAAA AGGTCTGGCA ATTTCAAGGTG GCGATGTAGG TGGTAGCAAG	3900
40	AACTTTGATG CGTTTGTAGA GCAAAAAGTA GATGTAGCTA AAGGATTCGG CATTGATGAA	3960
	GATGTTGCAA GACGTTTAGC ATCTAAATAT GGTTCAAATG TTGATGAATT GTTCAACATT	4020
	GCGCAAACAT CTCAATACCA TGATAGCAAG TTACCATTAG AAATTTATGT AGAACTTGTT	4080
45	TATAGTATTC AACAAGAAAT GGTATACAAA CCTAACGATT TCTTAGTTCG TCGTTCTGGT	4140
	AAAATGTATT TCAATATTAA AGATGTATTA GATTATAAAG ATGCTGTCAT CGATATTATG	4200
	GCAGATATGC TTGATTACTC TCCAGCTCAA ATTGAAGCAT ATACTGAAGA AGTTGAGCAA	4260
50	GCAATTAAAG AAGCGCAACA TGGaATAAT CAACCAGCAG TTAAAGAATA AtTAATTTGT	4320
	ACAATCATAA ACTGGTGTCC TGTTTTAAGG GCATCAGTTT TTTTATACGA GATACATTAG	4380

55

GTTATTAAAG GTGTGAGATG ATGACTGAAA AACAAATTTAA ATTAAGTGTG CAAGATAATA 4500
 CGAATATTGA AGTTAAAGTG AATTTTACAG ATGTAGATTG AAAAGGAATT ATTCATATAT 4560
 5 TTCATGGTAT GGCTGAACAT ATGGAACGTT ACGATAAATT AGCACATGCA CTTTCAAAGC 4620
 ATGGCTTCGA TGTGATACGT CATAATCATC GAGGACATGG TATTAATATT GATGAATCAA 4680
 CAAGAGGGCA TTACGATGAT ATGAAACGAG TTATCGGTGA TGCCTTTGAA GTAGCGCAAA 4740
 10 CAGTGAGAGG CAATGTTGAT AAACCATACA TTATAATCGG ACATTCAATG GGATCCGTTA 4800
 TAGCTAGATT GTTTGTAGAA ACATATCCGC AATATGTTGA TGGTCTAATT TTAAGTGGTA 4860
 CTGGTATGTA TTCATTATGG AAAGGTTTAC CAACCGTTAA AGTGTTACAA CTGATTACAA 4920
 15 AAATTTATGG TGCTGAGAAA CGAGTTGAAT GGGTTAACCA GTTAGTATCA AATAGTTTTA 4980
 ATAAAnnAT ACGTCCATTA CGTACACAAA GTGATTGGAT TTCTAGTAAT CCAATTGAAG 5040
 20 TAGATAaCTT TATTAAAGAT CCATATAGTG GaTTTAATGT GTCAAATCAA TTATTATATC 5100
 AAACAGCCTA TTATATGCTA CATAATCAC AATTAAAAAA TATGAAAATG TTAAaTCATG 5160
 CCATGCCTAT ATTATTAGTT TCAGGATATG ACGATCCTTT AGGTGATTAT GGTAAAGGGA 5220
 25 TTTTAAAATT GCGGAATATA TATAGAAACG CTGGCATHAA AAATGTTAAA GTGAATCTTT 5280
 ATCATCATAA ACGTCATGAA GTGTTATTTG AAAAnGATCA TGACnAAATT TGGAAGACT 5340
 TGTTTAAATG GTTGAATCAA TTTTATAAAA AATAAGAAA GTGGAATTAA ATATGAATAA 5400
 30 AAATAAGCCT TTTATTGTAG TAATTGTGGG GCCAACTGCT TGCAG 5445

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2569 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TGGCTTGAAC TACGCCAATA AGTCCCCCTA GTACAAGAAT GAATACCATG ATATCGACCG 60
 45 CTTCTATCGT ACCTTCAACC ATGCTACTTG TTATTTGTTT TGGTCCAGCT GGATGTTGCT 120
 TTAATCTTTC ATAAGTATTC GGAATTGATA CCGGCTTATT AATTGCACCT GATTTAAATT 180
 GTTCAATCTT AATTTTAACC CCCATTTTGT CTAGTTCCTG TTGCGTACCC GGAACCTTTT 240
 50 TCACTTGGTT ATGAGGGTTA ACTATCTTTA GTTCTTGGGA TGAAGGTTCT TAAGAAAGTT 300
 TAGAATATGC ACCAGCAGGA ATAACCCATG TTGCTATAAC TGCAACAACC GTTAAAATGA 360

	TAATTGTATT TTCCACGGTT TCATCTCCTT CGACATTTAA CCTAGCATT TCTACCTTAAA	480
	GATTTTATAA ATATAAATTA AGAAAGTGCA CCCCGCATCA AAATAGAGGC ATTATTTTCA	540
5	GGGGGTGCAC ATAAATAATA AAAATCATGC ATTTGACATA TAGTAATTGA AAAGCGTTTC	600
	AATTCAATTA CTTTTTAATC ACAGTACCTA CTTTACCCTC TAAGGCAGCA TCTAATTCAT	660
10	TTAATGATGT TATAAGCACA CTTCTTTTGG GATTGTTTTT AATAAATGAT ATGGCTGCTT	720
	CAATTTTGG TAACATACTT CCTTTTGCAA ATTGATTTTC GTCTATATAT CGTTTTAATT	780
	CATCAACATT TGTGTTTTTC AAAGGCTGTT GGTTTTCAGT GTTAAAATTA ATATATACAT	840
15	AATCAATTGC TGTTAAAATA ATCAATTGAT CGCATTGAAT ATTAGCACCC AACACGCAC	900
	TTGTTTTATC TTTGTCTATA ACTGCATCAA TACCTTTAAA ACCATCATGT TGCTCTCTAA	960
	TTACTGGTAT ACCTCCACCA CCAGCAGCAA TAACGAGTGT ATCATTTTTA ATAAGTGTTC	1020
20	TAATACTCTC TAATTCAATA ATAGAGATGG GTTGTGGTGA AGGAACAACG CGTCTATATC	1080
	CTCTTCCAGC ATCTTCAACA AATATAAATC CTTTTCTTTT TTGAATTTGT TCAGCTTCTT	1140
	CTTTGTTGTA AAATAACCCA ATTGGTTTTG AAGGATTGTT AAATGCCGGA TCATTTTCAT	1200
25	CAACTTCAAC TTGTGTCACT AGTGTTACCA CTTGTTTATC CATTCCAATA GAATGCAATT	1260
	CATTTTGTA GCTTTCTTGT AATTGATAGC CGATGTAAGC TTGACTCATT GCGCCACATT	1320
30	CAGCAAATGG AAATGCCGGA CCTTGGTTAT GTTCTGCAGC ATAGTTAAGT CCCAAATTAA	1380
	TGCTTCCAAC CTGTGGTCCA TTACCATGAC TAATAACAAT CTCATGTCCT TTTGTnATTA	1440
	AyCCTACTAA TGATTtCGCA GTATTTTTTA CAAGCTCGAG TtGgTyCTTG aGGTGATTTh	1500
35	CCTAAAGCAT TACCACCTAA TGCTACTACT ATTTTCGCCA TCATATTCAC TTCCTTATAT	1560
	CATTTAAAAT TCACCCAATG TAGCAACCAT GaCTGCTTTG ATTGTATGCA TTCTGTTCTC	1620
	AGCTTCTTGG AATACAACTG AAGCTTTACT TTCGAATACT TCATCTGTAA CTTCCATTTC	1680
40	TCGAATACCA TATTTTTTCAA AAATTTGTTG ACCTATTTTC GTATCAGCAT TATGGAAAGA	1740
	TGGTAAGCAA TGCTCAAAA TAACATTTGG ATTACCAGTT TTATCCATTA TTTCTTTATT	1800
	TACTTGATAT GGTTCaATA ATTCAAGTCG TTCTTTCCAT ACTTCATCAG GTTCACCCAT	1860
45	TGATACCCAA ACATCAGTGT AAATTACATC CGAACCTTTT ACaCCTTGGT CaATATCATC	1920
	TGTGATTAAT ATGTTGCCaC CATTTTCaGC GGCAATATTT TTACAGCGAT TTAATAATTC	1980
50	ATCTGTTGGA TTTAATTCTT TTGGACAAAC TAAATGGAAG TTCATACCCA TAATGGCAGC	2040
	ACCTTGCAAT AATGCATTTG CAACGTTATT ACGACCATCT CCAACATATG TAAAGTTAAT	2100
	ATCTGCATAA TCTTTTTTTTA AGACTTCTTT TGCTGTTAAG AAATCAGCAA GAACTTGAGT	2160

55

TTCTACTGTT CTTTGTGAAA AACCACGGTA TTCAATGCCA TCATACATTC CACCAAGCAC 2280
 ACGTGCAGTA TCTTTAGTTG TTTCTTTTTT ACCCATTTGT GATCCAGTTG GGCCTAAATA 2340
 5 AGTTACATTT GCACCTTGAT CATGCGCTGC AACTTCAAAT GCACATCGCG TTCTTGTAGA 2400
 ATCTTTTTTCA AATAACAGTG CAATATTTTT ATTTTTTAAC ATAGGCTTTT CAGTGCCAAT 2460
 ATATTTAGCA CGTTTTAAAT CCTCGGAGAG TGTTAATAAG GTTCTACCTC TTGTCGTGAA 2520
 10 AAGTCTAATA AAGTTAAAAA ACTTCTGTTT CGTAAATTTT TCATTAANA 2569

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 1273 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

CCTGGAACCA TCCaATCGtG CaAATCtTGa AAGaGAATAC GCAACAACAA TTAAATGTAT 60
 25 TGGAACACTA TATTCCAAAT GACCATCCAG CACTCGTTGA ATTAAAAATA TGGGAACGTT 120
 GGTTACATAA ACAAGGTTAC AAAGACATCC ATTTAGATAT TACTGCGCAC CACCTAGATC 180
 CTATTACACA GGGTTATTTA TTCAATGTCA TTTTGCTGAA AATGAATCTC GAGTTTTAAC 240
 30 AGGTGGTTAT TACAAAGGAA GCATCGAAGG GTTTGGATTA GGATTAAACAC TTAAAGTAAG 300
 GGAGTATGCA CAATGTTAAG AATCGCCATA GCCAAAGGAC GTCTAATGGA TAGTTTAATT 360
 35 AACTATTTAG ATGTAATTGA ATATACGACA TTATCAGAAA CATTAAAAAA TAGAGAACGC 420
 CAATTATTAT TAAGTGTAGA TAATATTGAA TGCATTTTAG TAAAAGGAAG TGACGTGCCA 480
 ATCTATGTGG AACAAGGAAT GGCAGACATA GGCATTGTTG GTAGCGACAT ATTAGATGAG 540
 40 CGCCAATATA ATGTTAATAA TTTGTTGAAT ATGCCTTTTG GAGCATGTCA TTTTGCGGTT 600
 GCAGCGAAAC CTGAAACGAC CAATTATCGT AAAATCGCAA CGAGTTATGT TCATACTGCT 660
 GAAACATATT TTAAATCAAA AGGTATTGAT GTCGAATTGA TTAAATTGAA TGGCTCTGTT 720
 45 GAATTGGCCT GTGTTGTAGA TATGGTAGAC GGAATTGTCT ACATCGTTCA AACAGGTACT 780
 ACGCTAAAAG CGAACGGACT GGTGAAAAG CAACATATTA GTGATATCAA TGCAAGATTA 840
 ATAACATAA AAGCAGCTTA TTTTAAAAAA TCACAATTAA TAGAGCAATT TATTCGCTCT 900
 50 TTGGAGGTGT CTATTGCCAA TGCTTAATGC ACAACAATTT TTAAATCAAT TTTCATTAGA 960
 AGCACCATTA GATGAGTCAT TGTATCCaAT TATTCGCGAT ATTTGTCAGG AAGTTAAAGT 1020

55

TTTAGaAATT AGTCATGAmC AAATTAAAGC AGCATTTGAC ACATTAGATG AAAAAACAAA 1140
 ACAAGCATTa CAACAAAGTT ATGAAAGAAT TANAGCATAT CAaGAAaGTA TtaAACAGaC 1200
 5 GaATCAACAG TTAGAAGaAT CAGTGGaGTG tTrTGaAATA TACCATCCmC taGaAAGTGT 1260
 CGGTATTTTAT GTG 1273

(2) INFORMATION FOR SEQ ID NO: 76:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

20 GTTGATAAAT TAAAAATGTT TTTATCAGAT ATTCAAAGTT ACCAACAATA TAGTAAAGAT 60
 CATCCGGTGT ATCAGTTAAT TGATAAATTT TATAATGATC ATTATGTTAT TCAATACTTT 120
 AGTGGACTTA TTGGTGGACG TGGACGACGT GCAAATCTTT ATGGTTTATT TAATAAAGCT 180
 25 ATCGAGTTTG AGAATTCAAG TTTTAGAGGT TTATATCAAT TTATTCGTTT TATCGATGAA 240
 TTGATTGAAA GAGGCAAAGA TTTTGGTGAG GAAAATGTAG TTGGTCCAAA CGATAATGTC 300
 GTTAGAATGA TGACAATTCA TAGTAGTAAA GGTCTAGAGT TTCCATTTGT CATTTATTCT 360
 30 GGATTGTCAA AAGATTTTAA TAAACGTGAT TTGAAACAAC CAGTTATTTT AAATCAGCAA 420
 TTTGGTCTCG GAATGGATTA TTTTGATGTG GATAAAGAAA TGGCATTTC ATCTTTAGCT 480
 TCGGTTGCAT ATAGAGCTGT TGCCGAFAAA GAACTTGTGT CAGAAGAAAT GCGATTAGTC 540
 35 TATGTAGCAT TAACAAGAGC GAAAGAACAA CTTTATTTAA TTGGTAGAGT GAAAATGAT 600
 AAATCATTAC TAGAACTAGA GCAATTGTCT ATTTCTGGTG AGCACATTGC TGTCAATGAA 660
 40 CGATTAACTT CACCAAATCC GTTCCATCTT ATTTATAGTA TTTTATCTAA ACATCAATCT 720
 GCGTCAATTC CAGATGATTT AAAATTGAA AAAGATATAG CACAAATTGA AGATAGTAGT 780
 CGTCCGAATG TAAATATTTT AATTGTGTAC TTTGAAGATG TGTCTACAGA AACCATTTTA 840
 45 GATAATGATG AATATCGTTC GGTTAATCAA TTAGAACTA TGCAAAATGG TAATGAAGAT 900
 GTTAAAGCAC AAATTAAACA CCAACTTGAT TATCGATATC CATATGTAAA TGATACTAAA 960
 AAGCCCTCAA AACAACTCTG TTCTGAATTG AAAAGACAAT ATGAAACAGA AGAAAGTGGC 1020
 50 ACAAGTTACG AACGAGTAAG GCAATATCGT ATCGGTTTTT CAACGTATGA ACGACCTAAA 1080
 TTTCTAAGTG AACAAAGTAA ACGAAAAGCG AATGAAATTG GTACGTTAAT GCATACAGTG 1140

55

GATGGATTAA TCGATAAACA TATTATCGAA GCAGATGCGA AAAAAGATAT CCGTATGGAT 1260
 GAAATAATGA CATTATCAA TAGTGATTAT ATTCGATATT GCTGAAGC 1308

5 (2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1431 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

GATGCCATTn ATnnGTATGC AAGAAGTTGT TCCGGGTTCA GGTGGATTaC CAGTTGGTAC 60
 TGGTGGTAAG ACGTTACTAA TGCTTTCAGG CGGTATAGAC TCACCAAGTTG CTGGGATGGA 120
 20 AGTGATGAGA CGTGGCGTAA CAATTGAAGC GATTCATTTC CATAGTCCAC CATTTACAAG
 TGATCAAGCA AAAGAAAAAG TTATTGAATT GACACGTATT TTAGCTGAAC GTGTTGGACC 240
 AATTAAATTG CATATTGTAC CATTACAGA ATTGCAAAAA CAGGTAAATA AAGTTGTACA 300
 25 TCCAAGATAT ACAATGACTT CAACGAGACG TATGATGATG CGTGTGCTG ATAAATTAGT
 ACATCAAATA GGGGCTTTAG CTATTGTAAA TGGTGAAAAC CTAGGCAGG TAGCCAGTCA 420
 AACACTTCAT AGCATGTATG CAATTAATAA TGTAACCTCT ACTCCTGTAT TACGTCCTTT 480
 30 ATTAACCTAC GATAAGAAG AAATTATTAT TAAATCGAAA GAAATTGGTA CATTTGAAAC
 ATCTATTCAA CCATTGAAG ATTGTTGTAC AATTTTCACC CCTAAAAATC CAGTAACCGA 600
 ACCAAACTTT GATAAGGTAG TCCAATATGA AAGTGCTTTT GATTTTGAAG AGATGATTAA 660
 35 TCGTGCTGTT GAAAATATTG AAACACTTGA AATAACTAGT GATTATAAAA CTATTAAAGA
 ACAGCAAACA AACCAATTAA TAAACGACTT TTTATAAATA AAATCCTAGA GTAAATTTAA 780
 40 ACATAAGGGG ATGTTAAACT ATGGATTGA ACTTAACGAT GATTATAATC ATAATTTTAT
 TTGTTTTTAT CGCGGCGTTT ATAGATTCGG TTGTAGGGGG TGGCGGTTTA ATTTCTACGC 900
 CAGCATTATT AGCAATCGGT CTACCACCAT CTGTGGCTTT AGGTACAAAT AAATTGGCAA 960
 45 GTTCGTTTGG TTCTTTAACT AGTACGATAA AGTTTATAAG GTCCGGTAAA GTGGACTTAT 1020
 ATGTTGTTGC CAAATTATTT GGTTTTGTAT TTTTGGCATC TGCAATGTGC GCATATATTG 1080
 CAACGATGGT TCCGTCACAA ATATTGAAAC CTTTAATCAT CATTGCACTT TCGTCGGTGT 1140
 50 TTATATTCAC ATTACTTAAA AAAGATTGGG GCAATACACG CACGTTTACT CAATTTACAT 1200
 TTAAGAAAGC CATAATATTT GCAGCACTTT TTATATTAAT CGGCTTTTAT GATGGATTG 1260

55

TAAGTGCAGC AGGAAATGCT AAAGTTTTGA ACTTTGCTTC TAATATAGGT GCGCTTGTAT 1380
 TATTTATGGT ATTAGGACAA GTAGATTATG TAATAGGTTT AATTATGGCT A 1431

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4403 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

AAATATTATTT TAAATTCAAT ATTTATTGGT GCATTTATTT TAAACTTATT ATTGCGCTTT 60
 ACCATTATTT TCATGGAAAG ACGTTCTGCC AATTCTATCT GGGCTTGGTT ACTAGTCTTA 120
 GTTTTCTTGC CTTTATTCGG CTTCATTTTA TACTTACTAT TAGGACGACA AATTCAACGT 180
 GACCAAATTT TCAAATTGA TAAGGAAGAT AAAAAAGGAT TAGAGTTAAT CGTTGATGAG 240
 CAATTAGCTG CTTTAAAAAA TGAAAACCTT TCAAATTCCA ATTATCAAAT TGTAATAATT 300
 AAAGAAATGA TTCAAATGTT GTTATATAAT AACGCAGCAT TTTTAACAAC AGACAACGAT 360
 TTAATTTTAT ACACAGACGG CCAAGAAAAA TTTGATGACC TAATACAAGA CATCCGTAAT 420
 GCTACTGATT ATATTCATTT TCAGTACTAT ATTATTCAAA ATGATGAATT AGGTCGTACC 480
 ATTTTAAATG AACTTGGTAA AAAAGCGGAA CAAGGTGTAG AAGTTAAAT TCTTTATGAT 540
 GACATGGGTT CTCGTGGACT GCGTAAAAAA GGCTTACGCC CGTTTCGCAA TAAAGGTGGA 600
 CATGCTGAAG CATTTTTCCC ATCAAAATTA CCTTTAATTA ACTTGCGTAT GAACAATCGA 660
 AACCATCGAA AAATTGTTGT AATAGATGGG CAAATTGGAT ATGTTGGTGG TTTTAATGTT 720
 GGTGATGAGT ACTTAGGTAA ATCAAAAAAA TTCGGCTATT GGCGAGATAC GCATTTACGA 780
 ATTGTCGGGG ATGCAGTGAA TGCATTGCAA TTACGATTTA TTCTAGATTG GAATTCACAA 840
 GCCACACGTG ACCACATCTC CTATGATGAT CGTTATTTCC CAGATGTAAA TTCTGGTGGA 900
 ACAATTGGCG TTCAAATAGC TTCTAGTGGT CCTGACGAAG AATGGGAACA GATTAAATAC 960
 GGCTATTTGA AAATGATTTT ATCTGCTAAA AAATCGATTT ATATTCAATC TCCCTATTTT 1020
 ATACCTGATC AAGCCTTTTT AGATTCTATT AAAATTGCGG CATTAGGTGG TGTTGATGTC 1080
 AATATCATGA TTCCTAATAA ACCTGACCAT CCGTTTGTTC TTTGGGCTAC TTTAAAAAT 1140
 GCAGCATCCT TATTAGATGC CGGTGTTAAA GTATTTCACT ACGACAATGG CTTTTTACAC 1200
 TCAAAAACAC TTGTTATAGA TGATGAAATT GCAAGTGTGG GAACAGCTAA TATGGACCAT 1260

AAATTAAAC AAGCTTTTAT AGATGATTTA GCAGTATCTT CTGAATTAAC AAAAGCACGT 1380
 TATGCTAAGC GAAGTCTTTG GATTAAATTT AAAGAAGGTA TTTACAATT ATTGTCACCT 1440
 5 ATCTTATAAA ATAGAAATAT GAGGAGTGTA aCTTTAATGC AACAATCAGA CGTCATTAGT 1500
 GCTGCCAAAA AATATATGGA ATCTATTCAT CAAAATGATT ATACAGGCCA TGATATTGCG 1560
 CATGTATATC GTGTCACTGC TTTAGCTAAA TCAATCGCTG AAAATGAAGG TGTTAATGAT 1620
 10 ACTTTAGTCA TTGAACTCGC ATGTTTGCTT CATGATACCG TTGACGAAAA AGTTGTAGAT 1680
 GCTAACAAAC AATATGTTGA ATTGAAGTCA TTTTATCTT CTTTATCACT ATCAACCGAA 1740
 15 GATCAAGAGC ACATTTTATT TATTATTAAT AATATGAGCT ATCGCAATGG CAAAATGAT 1800
 CATGTCACCT TATCTTTAGA AGGTCAAATT GTCAGGGATG CAGATCGTCT TGATGCTATA 1860
 GCGCTATAG GTGTTGCACG AACATTTCAA TTTGCAGGAC ACTTTGGTGA ACCTATGTGG 1920
 20 ACAGAACATA TGTCACTAGA TAAGATTAAT GATGATTTAG TTGAACAGTT GCCACCATCT 1980
 GCAATTAAAC ATTTCTTTGA AAAATTACTT AAGTTAGAAT CTTTAATGCA TACAGATACG 2040
 GCGAAGATGA TTGCTAAAGA ACGTCACGAC TTTATGATGA TGTACTTGAA ACAGTTTTTT 2100
 25 ACGGAATGGA ATTGTCACGA CTAGACATTG AAGTTGTAGT ATGATGATGC GATGTAATGG 2160
 CGTGTGTTG TGGAAGCTTG GTGTCATGCC ATGTTACTTT GATGTGTTGT TGTGGGAGCT 2220
 TGGTGACATG TCATGCTACT TTGATGTGCT GGTACCACGA TCGCTCTGA TGTAGTGCTA 2280
 30 TGATGTGGCA TTGCGGTGTT ATGGTGTTAT AGACAGGTTT GCGGTTGATG CCATGTTACT 2340
 TTGATGTGCT GGTACCACGA TCGACTTGA TGTAGTGCTA TGATGTGGCA TTGCGGTGTT 2400
 35 ATGGTGTTAT AGACCGGTTT GATGTTGATG CCATGTTACT TTGATGTGCT GGTGCTACGA 2460
 TCGACTTGA TGTAGTGCTA TGATGTGGCG TTGCGCTGTT ATGGTGTTAT AGCCAGGTTT 2520
 GGTGTTGATG TCATGCCGTT ACGATTCTAT GATATGTTGT TGGGACGTTG CAATGTGTAT 2580
 40 TATGCCGTTG TGACGTTATT ATTTCACT GTTACATGTA TAAGTGAATT GCTGTGGAAA 2640
 TTTGCGACAT ATACTGCTAC ACTGATGAAT CATTGTGTCA AGATGACATT GCGATGAAGA 2700
 ATGACAACCTC TGTTATTAAC CACTTTTTAC ATACTGAAAA CTCGTTAATA TTATTTCAAA 2760
 45 TAAAAACAGC AGTAGGATGA CTTTCACATT TGAAATCATC TTAAGTCTGT TTCTATTTAT 2820
 CACATATTGT ATAATGTGAC ACTAAGTTTC GCTATTGAAG CGAAAAATAA TGTGCGCCCT 2880
 ATAAAGTTAA AATTATCTTC AACTTTTAGG GTGCACATTA TTTGGACTTG CTAAGGTTAT 2940
 50 TTCTTTTTCT TTTTAGACAC AACTTGTGTG TTTTGCCTT TTTTATTGct GCCGCCGTTG 3000
 TGCTCTCTTT CATACGCTTC AATGAAAGGT TGTACTTCTT TTTTAGCGAC TTTTTCATAA 3060

55

CCAAGTGCTG ATGCTGAGCT TAATGAAATC CAGATAATCA TAATTGGTGA AATGACCATC 3180
 ATCATGTAAC CCATTTGACG TTGTTCTGCT GGCATCGTTT TACTTGATAC ATATGCTTGG 3240
 5 ATAAAGTATA AAACACCGGC AATAATTGTA ATCCAAATAT CAGGACGTCC TAAATCGAAC 3300
 CATAAGAAGT GTGGATATTT AAACAAACCA TCTACAAGTT GGTCTTTAAG TACAAAGTAT 3360
 AATCCCATGA TGATTGGTAA TTGGATTAGC ATTGGTAAAC AACCCAACAT ACTCTTAATC 3420
 10 GGGTTCATGT CATACTTTTT ATATACTTGC ATTAATTCTT GGTTCGAGC CATTTTTTCT 3480
 TCTTGTGTAC GCGnCaCGT CACTTTTTCT TGAATTTTTT CAACTTCTGG CTTTGCAACT 3540
 TTCATTTTTT GACGCATCAT ATGACTATTT TTATAGTTTG ACAACATGAA TGGAATAAAA 3600
 15 ATAATACGAA TTACCAATAC AAGGATAATA ATAGCTAAAC CATAATTGTC GTTTAATAAG 3660
 TTATTTCCCA ACCAATCCAA TACATTTTTC ATTGGATCTA CGAATGTATT GTAGAAAAAy 3720
 20 cwCtACGTTT TTCAGGTTTA GAATAGTCAC AACCAGCCAA AAAGACCATA ATACCTAAAA 3780
 ATAATGGTAG TAACGCTTTT TTCTTCATTT TTCCACCTCT ATCATTATAT TCACATAGGA 3840
 TTTATTCTAT CACATTAATG AGTACGTATG AAACAATAAG TGGAAAAATT TAACTAATTA 3900
 25 TTAAAAAAT CTTTGAATCG ATTAACAGTC TTTTCAATAT TTTCACTTTT AGAAATGGCT 3960
 GAAATGACTG AAATTCCATT GGCACCTGCT TCTACAATCG GCGCCACATT ATTAGTATTG 4020
 ATACCGCCAA TAGCTACAAT CGGTAGTTGC GGATTCATTT CTTTAAACGT TGCAATCATT 4080
 30 TCTGGACCTA CTGGTATATG CGCGTCATGC TTCGACGGCG TAGGATAGAT TGGTCCAACA 4140
 CCTATATAAT CmACATGAGT TAAATCAGAT TTTGCATACT CATCTAAATC ACTAATACTA 4200
 AGTCCAATAA TTTTATCAGT GAAATATTGT GCTATCTCTT TGACTTTCGC ATCATCTTGA 4260
 35 CCGACATGTA TACCATCCGC GTTAATTTCT TTTGCCAAGG ATACATCATC ATTAACGATA 4320
 AAAGGCACAT CATATTGATG ACAGAGATGC TGAATTCTT TAGCTAATAC AAGTTTATCG 4380
 40 TTTCCTTTTA AAGCTGATTC ACC 4403

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1808 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

TGGAnCCAAT ATTAGAAATG ATTAAACAT TAACAGGTAT TAATAGTCCT TCAGGAGnCA 60

	TAACAAATAA AGGTGCGTTA TTAATAACAG TGCCAGGCAA AAATGATGAA GTACAACGCT	180
	GTATTACTGC TCATGTTGAT ACTTTAGGTG CaATGGTTAA AGAAATTAAA GAAGATGGTC	240
5	GCTTaGCAAT AGAATTAATT GGAGGATTCA CGTATAACGC GATTGAGGGT GAATATTGCC	300
	AAATTAAAAC TGATGCTGGT CAAATATATA CAGGAACAAT TTGTCTGCAT GAAACAAGTG	360
10	TTCATGTATA TAGAAATAAT CATGAAATAC CTAGAGATCA AAAGCATATG GAAATAAGAA	420
	TTGATGAAGT AACTACATCA GAAGAAGATA CAAAGAGTTT AGGTATTTCA GTAGGTGATT	480
	TTGTTAGCTT TGATCCACGT ACAGTTATCA CGTCATCAGG TTTTATTAAA TCTCGTCATT	540
15	TAGATGATAA AGCTAGCGTA CGgTtGATAC TACAATTACT AAAGAAATTA AAAGAAGAGC	600
	AAATAATATT ACCACATACA ACGCAATTTT ATATTTCTAA TAACGAAGAA ATAGGTTACG	660
	GTGCAAATGC ATCAATTGAT TCGAAAATCA AAGAATATAT TGCAITAGAT ATGGGCGCGT	720
20	TGGGAGACGG TCAAGCATCG GATGAATATA CAGTTTCTAT TTGTGCCAAA GATGCTTCAG	780
	GTCCATATCA TAAGCAATTG AAATCGCACC TAGTTAATCT TTGCAAATA AATAACATTC	840
	CATATAAAGT AGACATATAT CCATATTATG GTTCAGATGC TTCAGCAGCT TTACATGCTG	900
25	GTGCGGATAT CAGACATGGT TTATTTGGCG CTGGCATTGA ATCATCTCAT GCAATGGAAC	960
	GAACACATAT TGATTCTATT AAAGCGACAG AGAAATTACT ATATGCATAT TGCTTATCAC	1020
	CAATTGAGTA AACCAATTAGT GTTGACAAAT GTGaACGACC TATGTAATAT AATGAACTAT	1080
30	AAAAATAATT AGAATTTTCT AAAGAAATAG TAGCAGATAT GAAACGTAGC AAATAGAAAG	1140
	CTAATGGGTG ATGGGAATTA GCACGCCATA TCTTGTAAT TGGACTTTGG AAAACAATTG	1200
35	AATGAGTTTT GAAAGTGAAC ATGAATTATG TTAACAAAGG TGGCACCACG GTAACGCGTC	1260
	CTTACAGGTA TATGCGTTAT GTGGTGTCTT TTTATTTAGA CAAAATGTAG TAGTTAATTA	1320
	AAGGTAGCAA CAGAAAGTTA GTGGATGATG TGAACAAACA CCGAGATTAA TGAAATTGGG	1380
40	TTTTGTCTGC AACAGAAAAA TTATATATAG TAAAGAGTGA ACTATGAATA TTTCGAATAT	1440
	TCGGTTAATT TAGGTGGTAC CACGCGTCAC nTCCTTTATA TTGATAAGGA TGCTGGCGCT	1500
	TTTTTGAAAG GAGCGTATAG AATGGATATA TTTTATAAAA AAATAAAAAGC AAATGTAACG	1560
45	CCCGAAGTTT TAGCACAACT TCATTCCAAG AAGaTCATTT TGGAAAGTAC AAATCAACAA	1620
	CAAACTAAAG GTCGCTATTC AGTTGTTATT TTTGATATTT ATGGCACTTT AACTTTAGAT	1680
	AATGATGTAT TATCAGTAAG TACTTTAAAA GAATCGTATC AAATCACTGA AAGACCGTAC	1740
50	CATTATTTAA CGACTAAaAT AAATGAAGAC TACCATAATA TTCCAAGATG AGGCAACTTA	1800
	AGTCATTA	1808

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1320 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

TGGTCGTCAA	TTTCTTGATT	ATATCTATAA	TCCTCATTTT	CAATATTAGA	GTCTGTAGAA	60
TCATCGATAT	TATTATCATT	CGCATGACTA	GAAGCAGAAT	CATTATTTTT	ATCATTGCTT	120
TCTTCTTTTT	TGAAGTCTTT	ATTTATCAAG	TAAATTTCTT	CATCAAAATC	AGCTTGTTGA	180
GATGTATCAT	CTTTATTTTG	ATTAGAAAAA	TGTGTAGCCT	TTGATCTTTT	TCTTTGCCGT	240
CTTTTCTTAG	ATGTATTCCT	CGTAAATAAT	TCTAATTCAT	CTTTATCTTC	ATTTGATTCT	300
TGTTGATCGT	TCTTCGTTTT	ATCATCCATC	AATACTCACA	CCCTTTAATA	AGATGGTAAA	360
TGGGCACGGA	ATCTTTCAAT	AAATTTCTCT	CCACGCTCTT	CAAAAGTACT	ATATTGATCC	420
CAACTCGCAC	AAGCAGGTGA	CAATAATACA	ACATCATTTG	GTTCTATAAT	ATCTTGTAAT	480
TTATCAACAG	CGTCTTCGAC	ATTGTTGCTT	TCAATGACCG	ATTTCCCTTG	ACTATTACCT	540
AGTTTAGCAA	ACTTAGCTTT	CGTTTGTCGG	AATACAACCA	TCGCGCGAAC	ATTTTCCATA	600
TAAGGAATGA	GTTTCGTCAA	TTCATTCCCT	CGATCCAAAC	CACCACATAA	CCAAATGATT	660
GGTGATTAA	ATGAATTTAA	GGCAAAGTGT	GTTGCTAGCG	TGTTTGTTGC	TTTGGAATCA	720
TTATAATATT	TATTAGTTCT	ATTAGTACCA	ACATATTGCA	ATCTATGCTC	TATTCCTGAA	780
AATGTAGTTA	AACTATCAAT	AATTGCTTTA	ATAGGTACAC	CAGCanaATA	CAAGCAAGCA	840
CAGCTGCTAA	TATATTTCTA	AATTATGTTT	ACCAGGCAAT	ACTAGATCTT	CAGTGTTAAT	900
AATaCGAACA	CCTTTATaAA	CGATAAAACC	ATCTTTaATA	TAAaTACCAT	CaCTTCTTG	960
TTGAGTTGAG	AAATACAATG	TCTTAGCTTT	TAATTCTTCC	GACTCTATCA	CTTGTCTTTG	1020
ATGATAATTA	CAAATCAAAT	AATCCTCTTC	CGTTTGATTT	TTATATATTT	GCTTTTTTAGC	1080
ATTTTGATAG	TTTTCTAAAT	TTTCATGGTA	ATCTAGATGC	GCCGAATAAA	TGTTAGTAAT	1140
TATAGCAATG	TGTGGTTTAT	ACTTTTCGAT	TCCAAGTAAC	TGGAATGACG	ACAACTCTGT	1200
AACTAAATAA	TCTGTAGGCT	TTACTTCTTG	TGCTACTTTA	GATGCAACAT	AACCAATATT	1260
GCCGGATAAT	CTTCCAGTTA	AGCGACTTTT	TTTAAACATA	TCTCCAATTA	GAGAAGTAAC	1320

50

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4280 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

	TTTACACCAA TCAAAAAATC GAACTGATAT AAATAAGTAC AAAGCTTATC TATCAATCCG	60
10	ATTTAGTTAT AAAACAAAAA AAGCCACAGT AATGTGGCTT TTTGTTATAT TCAGTATCAA	120
	AATGGTATCA ATAGCCATTT TCGGAAGTCA AGAATGGCTT AACACGCCG TTTAAAGCTA	180
	TCCAATACTA CCTTCCATTT CGAACTTGAT TAAACGGTTC ATTTCCGACCG CGTATTCCAT	240
15	TGGAAGTTCT TTTGTAAATG GTTCGATGAA TCCATAACA ATCATTCTG TCGCTTCTTC	300
	TTCAGAAATA CCACGACTCA TTAGATAGAA TAATTGTTCT TCAGAAACTT TTGAAACCTT	360
	GGCTTCATGT TCTAATGATA TTTGATCGTT GAATACTTCG TTATATGGAA TTGTATCTGA	420
20	TGTTGATTCTG TTATCTAAGA TTAATGTATC ACATTCAATA TTTGAACGAG CACCTTTTGC	480
	TTTACGTCCA AAATGAACAA TACCGCGATA AATAACTTTA CCACCATTTT TAGAAATAGA	540
	TTTAGAAACA ATTGTAGAAG ATGTATTAGG TGCTTTATGA ATCATTTTAG CACCGGCATC	600
25	TTGAACTTGT CCTTTACCAG CAAATGCAAT AGATAATGTA CTACCTTTTG CACCTTCACC	660
	TAAAAGAACA CAGTTTGGAT ATTTTCATCGT TAACTTAGAA CCTAAGTTAC CATCTACCCA	720
	TTCCATATTT CCGTTTTTCAT AAACAAAAGT ACGTTTTGTA ACTAAATTGT ATACATTGTT	780
30	CGCCCGATTT TGAATCGTAG TATAACGAAC GTGCGCATCT TTATGCACAA TGATTTCAC	840
	AACAGCAGAG TGTAAGAAC TAGTTGTATA AACTGGTGCA GTACAACCTT CTACGTAATG	900
35	TACAGAAGCA CCTTCATCAG CAATGATTAA TGTACGTTCA AATTGACCCA TGTTCCTCAGA	960
	GTTAATACGG AAATAAGCTT GTAGTGGCGT ATCTAGTTTG ATATTTTATG GTACATAAAT	1020
	GAAAGGAACCA CCTGACCATA CTGCTGAGTT TAACGCCGCA AATTGTTTAT CTGCTGCAGG	1080
40	TACTACAGAA GCAAAGTATT TTTTGAATAA TTCTTCATTT TCTGTAAAG CACTATCTGT	1140
	ATCTTTAAAG ATAATACCTT TTTCTTCAAG TTCTTTTCC ATATTATGGT AAACAACCTC	1200
	AGATTCATAT TGAGCAGAAA CACCAGCTAA ATATTTTGT TCAGCTTCAG GAATTCCTAA	1260
45	TTTATCGAAA GTTCTTTTAA TTTCTTCTGG CACTTCATCC CATGAACGTT CAGCTTGTTT	1320
	TGAAGGCTTT ACATAGTAAG TAATGTCATC GAAATTCAAT TCTGATAAGT CGCCACCCCA	1380
	TTGAGGCATT GGCATTTTAT AAAACAATTT TAATGATTTA AGACGGAAAT CTAACATCCA	1440
50	TTCCGGCTCA TTTTTCATGT TAGAAATTC TCTAACGATA TTCTCAGTTA AACCACGTTT	1500
	TGATCTGAAA ATGGACACAT CATCGTCGTG GAATCCATAT TTATAATCCC CAACATCAGG	1560

55

	TTTAATTCAT GATGTAAACC ATATTATAAC AATGACATGA CATCTTATAA AAATTTTAT	1680
	ACTTTTATAT GTCTAATATC AAAATTATCT ATGATTAAAC GCATTCTATT CTTCTTCAGT	1740
5	CGTACCTTCT GCTTTACCTT CTTTAGCAAC AGTACCTTTT TCCAATGCTT TCCAAGCTAA	1800
	TGTGGCACAT TTAATACGAG CTGGGAATTG AGATACACCT TGCAATGCTT CAATATCTCC	1860
	CATTTCTTCT GTAATCACAT AGTCTTCACC AAGCATCATT TTCGTAAATT CTTGGCTCAT	1920
10	TTGCATTGCT TCTCCAAGTG AATGACCTTT AACAGCTTGT GTCATCATCG ATGCATTGCG	1980
	CATTGAAATC GAACAACCTT CACCTTCAAA CTTAGCATCT TTTATAATGC CGTCTTCTAT	2040
	ATCAAATGTT AGTCGTATAC GGTCAACGCA TGTGGGGTTA TTCATATCTA CTGTCATAGA	2100
15	CCCGTTATCT AATACACCTT TATTTCTAGG ATTTTATAA TGATCCATAA TGACAGATCT	2160
	ATATAATTGA TCTAGATTAT TAAAATTCAT AAGAGAAAAA CTCCTTCGTT TGTTC AAGG	2220
20	CATTTATTAA CTGATCAACG TCTTCTTTCG TGTTGTATAT ATAAAACTC GCTCTAGCTG	2280
	TTGAAGACAC ATTTAACCAT TTCATTAAACG GTTGGCGACA ATGATGCCCA GCTCTAACCG	2340
	CTACACCTTC TGTATCTACG GCTGTAGCAA CATCGTGTGG ATGTACATCT TGTAAATTAA	2400
25	ACGTTATTAC ACCTGCACGA CGATCCTTTG GCGGGCCATA AATTTC AATT CCTTCAATTG	2460
	CAGACATTG CTCATAAGCA TATATCGTTA ATTCTGTTC ATATTTATGA ATTGCATCAA	2520
	AACCTATGCG TTCTAAATAG CGAATAGCTT CTGCAAGCCC AATTGCTTGA GCAATTAATG	2580
30	GAGTACCCGC CTCAAATTTA GTAGGTAAAT CAGCCCATGT TGCATCATAC TTACTTACAA	2640
	AATCAATCAT GTCGCCACCG AACTCAATCG GTTCCATTTT TTGTAGTAAC TCACGTTTAC	2700
	CAAATAATAC GCCAATACCT GTTGGTCCAA GCATTTTATG ACCACTAAAA CTATAAAAAT	2760
35	CAGCATTCAAT TTCTTG CATA TCAAGTTTCA TATGTGGTGC TGCTTGC GCCATCAACAC	2820
	TGATTAATTGC ACCATGTTGA TGAGCTATTT CTGCAATGGT TTTAACATCA TTAATTGTAC	2880
40	CGAGCACATT AGATATATGT GCAATAGCAA CGATCTTTGT TTTATCATT AATCGTTTGCT	2940
	TAATATCCTC GATGTTTAAT TCACCGTCAG CTGTCATTGG TATAAATTTT AATGTCGCAT	3000
	TTTTACGCTT TGCTAACTGT TGCCAAGGAA CAATATTGGC ATGATGTTCC ATTT CAGTGA	3060
45	CAACAATTTT ATCGCCCTCT TCAACATTTG CATCACCATA GCTATGTGCT ACAAGGTTAA	3120
	TCGACGCAGT TGTTCGCGT GTAAAAATGA TTTCTTCAAA ATACTTCGCA TTAATAAAAC	3180
	GACGAACGGT TTCACGGGCA TTTTCATAAC CATCAGTTGC CAATGATCCT AATGTATGAA	3240
50	CACCACGATG AACGTTTGAA TTATAACGCT TGTAAGTAATC TTCTAAAACA TTTAACACTT	3300
	GCACAGGCGT TTGACTTGTC GCTGTTGAAT CAAGATATGC TAAACGTTTG CCATTGACTT	3360

55

CTTCAATCAC GACCTTTCTT AAATAAAAAAT CCTAATCATT TAAATACTGA CGTTGTATTA 3480
 GTCTTATACC AATATCGACA GTCTATATCT ATTACAACT TTTATTTTCA AAATATTATT 3540
 5 TAGAAACTTT GCGTTCAATT ACTTCTCTCA ATTGACGTTT AACGTCTTCG ATAGGTAATT 3600
 CACGTACTAC TGGATCTAAG AAACCATGTA TAACAAGACG TTCCGCTTCT CTTTGAGAAA 3660
 TACCACGACT CATTAAATAG TAAAGTTGAT CTGGATCAAC ACGACCTACT GATGCAGCAT 3720
 10 GACCAGCTTG TACATCATCT TCATCAATTA ATAAAATAGG ATTCGCGTCA CCACGAGCAT 3780
 GTTCAGATAA CATTAAATACA CGTGATTCTT GATTAGCAAT TGATTTAGTT CCACCATGCT 3840
 TAATGTAGCC GATACCATTA AATACAGACG ATGCATGTTT TTTATAACA CCATGTTTAA 3900
 15 GGATATAACC ATCTGTTTCT TTACCATATT GTACGATTTT AGATGTTAGA TTAATTTTTT 3960
 GTTCGCCTGT ACCTACAACT ACTGATTTAA GTGAACCTGT TGAACGATCA CCAAATAAAT 4020
 20 TTGTTGTATT ATCAATAATT TGGCTACCCT CATTCAATTA ACCTAGTGCC CAATTAATTG 4080
 AGGCATCCGC TTCAGTAATA CCACGTCGAA TGATATGACC TGTAAGCCTT TTATCCATAT 4140
 AGTCCACTGA GCCATATGTG ATATTGTAAT TTGCACCAGC AATCACTTCA GAAATAATAT 4200
 25 TtAATTGATT TCCTTCACCA GATGCATTG mTAAGTAATT TTCAACATAT GTGACTTCGG 4260
 CGCTTTCTTC AGTAACGATG 4280

(2) INFORMATION FOR SEQ ID NO: 82:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15598 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

40 TCnGACTCGA ACGGTGmAAC TAttCCGTTG TaATTCCgGA GgAAaCAAGG TATGCCCATC 60
 TGCaAAGAAA gaATGsAATG AACTTTTTGG AAATGTAGAA GTGGTAAATA AAGATAAAGG 120
 ATATTACATT CTGAGAAGTA TAAAAGCTTG AAATGAAATG GATATTCTGT TATAGTTATA 180
 45 TAATGTAAAA ATTTATGTTT AATAAGTGTG TACTTTTACG TTAAATAGAT AAGTTAATTA 240
 AGAATAAATA TAGAATCGAA AATGGTGTCA TCATTAGTGT TGCCGTTTTT TTTTGTCTT 300
 TTTATTAATA TGCTTATGGT ATTTAGCTAA AAGCGGATCA CATAATTTTT GAGGGGTGAA 360
 50 TCTGTTTGGC AGGTCAAGTT GTCCAATATG GAAGACATCG TAAACGTAGA AACTACGCGA 420
 GAATTCAGA AGTATTAGAA TTACCAAAT TAATAGAAAT TCAAATAA TCTTACGAGT 480

55

	CTGGTAATTT GTCATTAGAG TTTGTGGATT ACCGTTTAGG AGAACCAAAA TATGATTTAG	600
	AAGAATCTAA AAACCGTGAC GCTACTTATG CTGCACCTCT TCGTGTAATA GTGCGTCTAA	660
5	TCATTAAAGA AACAGGAGAA GTTAAAGAAC AAGAAGTCTT TATGGGTGAT TTCCCATTA	720
	TGACTGATAC AGGTACGTTT GTTATCAATG GTGCAGAACG TGTAATCGTA TCTCAATTAG	780
	TTCGTTTACC ATCCGTTTAT TTCAATGAAA AAATCGACAA AAATGGTCGT GAAAACTATG	840
10	ATGCAACAAT TATTCCAAAC CGTGGTGCAT GGTTAGAATA TGAAACAGAT GCTAAAGATG	900
	TTGTATACGT ACGTATTGAT AGAACACGTA AACTACCAAT AACAGTATTG TTACGTGCAT	960
	TAGGTTTCTC AAGCGACCAA GAAATTGTTG ACCTTTTAGG TGACAATGAA TATTTACGTA	1020
15	ATACTTTAGA GAAAGACGGC ACTGAAAACA CTGAACAAGC GTTATTAGAA ATCTATGAAC	1080
	GTTTACGTCC AGGTGAACCA CCAACTGTTG AAAATGCTAA AAGTCTATTG TATTCACGTT	1140
20	TCITTGATCC AAAACGCTAT GACTTAGCAA GCGTGGGTCG TTATAAACA AACAAAAAT	1200
	TACATTTAAA ACATCGTTTA TTAAATCAA AATTAGCTGA GCCAATTGTA AATACTGAAA	1260
	CTGGTGAAAT TGTAAGTGAA GAAGGTACAG TGCTTGATCG TCGTAAATC GACGAAATCA	1320
25	TGGATGTACT TGAATCAAAT GCAAACAGCG AAGTGTGTTGA ATTGCATGGT AGCGTTATAG	1380
	ACGAGCCAGT AGAAATTCAA TCAATTAAAG TATATGTTCC TAACGATGAT GAAGGTCGTA	1440
	CGACAACCTGT AATTGGTAAT GCTTTCCCTG ACTCAGAAGT TAAATGCATT ACACCAGCAG	1500
30	ATATCATTGC TTCAATGAGT TACTTCTTTA ACTTATTAAG CGGTATTGGA TATACAGATG	1560
	ATATTGACCA TTTAGGTAAC CGTCGTTTAC GTTCTGTAGG TGAATTACTA CAAAACCAAT	1620
	TCCGTATCGG TTTATCAAGA ATGGAAGAG TTGTACGTGA AAGAATGTCA ATTCAAGATA	1680
35	CTGAGTCTAT CACACCTCAA CAATTAATTA ATATTCGACC TGTTATTGCA TCTATTAAAG	1740
	AATCTTTGG TAGCTCTCAA TTATCACAAT TCATGGACCA AGCAAACCCA TTAGCTGAGT	1800
40	TAAAGCATAA ACGTCGTCTA TCAGCATTAG GACCTGGTGG TTTAACACGT GAACGTGCTC	1860
	AAATGGAAGT ACGTGACGTT CACTACTCTC ACTATGGCCG TATGTGTCCA ATTGAAACAC	1920
	CTGAGGGACC AAACATTGGA TTGATTAACT CATTATCAAG TTATGCACGT GTAAATGAAT	1980
45	TCGGCTTTAT TGAAACACCA TATCGTAAAG TTGATTTAGA TACACATGCT ATCACTGATC	2040
	AAATTGACTA TTTAACAGCT GACGAAGAAG ATAGCTATGT TGTAGCACAA GCAAACCTCTA	2100
	AATTAGATGA AAATGGTCGT TTCATGGATG ATGAAGTTGT ATGTCGTTTC CGTGGTAACA	2160
50	ATACAGTTAT GGCTAAAGAA AAAATGGATT ATATGGATGT ATCGCCGAAG CAAGTTGTTT	2220
	CAGCAGCGAC AgcATGTATT CCATTCTTAG AAAATGATGA CTCAAACCGT GCATTGATGG	2280

55

	CAGGTATGGA ACACGTTGCA GCACGTGATT CTGGTGCGGC TATTACAGCT AAGCACAGAG	2400
	GTCGTGTTGA ACATGTTGAA TCTAATGAAA TTCTTGTTCC TCGTCTAGTT GAAGAGAACG	2460
5	GCGTTGAGCA TGAAGGTGAA TTAGATCGCT ATCCATTAGC TAAATTTAAA CGTTCAAAC	2520
	CAGGTACATG TTACAACCAA CGTCCAATCG TTGCAGTTGG AGATGTTGTT GAGTATAACG	2580
10	AGATTTTAGC AGATGGACCA TCTATGGAAT TAGGAGAAAT GGCATTAGGT AGAAACGTAG	2640
	TAGTTGGTTT CATGACTTGG GACGGTTACA ACTATGAGGA TGCCGTTATC ATGAGTGAAA	2700
	GACTTGTGAA AGATGACGTG TATACTTCTA TTCATATTGA AGAGTATGAA TCAGAAGCAC	2760
15	GTGATACTAA GTTAGGACCT GAAGAAATCA CAAGAGATAT TCCTAATGTT TCTGAAAGTG	2820
	CACCTAAGAA CTTAGACGAT CGTGGTATCG TTTATATTGG TGCAGAAGTA AAAGATGGAG	2880
	ATATTTTAGT TGGTAAAGTA ACGCCTAAAG GTGTAACTGA GTTAACTGCC GAAGAAAGAT	2940
20	TGTTACATGC AATCTTTGGT GAAAAAGCAC GTGAAGTTAG AGATACTTCA TTACGTGTAC	3000
	CTCAGGCGC TGGCGGTATC GTTCTTGATG TAAAAGTATT CAATCGTGAA GAAGGCGACG	3060
	ATACATTATC ACCTGGTGTA AACCAATTAG TACGTGTATA TATCGTTCAA AAACGTAAAA	3120
25	TTCAATGTTG TGATAAGATG TGTGGTCGAC ATGGTAACAA AGGTGTCATT TCTAAGATTG	3180
	TTCTGAAGA AGATATGCCT TACTTACCAG ATGGACGTCC GATCGATATC ATGTTAAATC	3240
	CTCTTGGTGT ACCATCTCGT ATGAACATCG GACAAGTATT AGAGCTACAC TTAGGTATGG	3300
30	CTGCTAAAAA TCTTGGTATT CACGTTGCAT CACCAGTATT TGACGGTGCA AACGATGACG	3360
	ATGTATGGTC AACAAATTGAA GAAGCTGGTA TGGCTCGTGA TGGTAAACT GTACTTTATG	3420
35	ATGGACGTAC AGGTGAACCA TTCGATAACC GTATTTTCAGT AGGTGTAATG TACATGTTGA	3480
	AACCTGCGCA CATGGTTGAT GATAAATTAC ATGCGCGTTC AACAGGACCA TATTCACTTG	3540
	CTACACAACA ACCACTTGCC GGTAAAGCGC AATTCGGTGG ACAACGTTTT GGTGAGATGG	3600
40	AGGTATGGGC ACTTGAAGCA TATGGTGCTG CATAACATT ACAAGAAATC TTAACCTACA	3660
	AATCCGATGA TACAGTAGGA CGTGTGAAAA CACACGAGGC TATTGTTAAA GGTGAAAACA	3720
	TCTCTAGACC AAGTGTTCCT GAATCATTCC GAGTATTGAT GAAAGAATTA CAAAGTTTAG	3780
45	GTTTAGATGT AAAAGTTATG GATGAGCAAG ATAATGAAAT CGAAATGACA GACGTTGATG	3840
	ACGATGATGT TGTAGAACGC AAAGTAGATT TACAACAAAA TGATGCTCCT GAAACACAAA	3900
	AAGAAGTTAC TGATTAATAC GCAATTTACA AAACAGGCAA AAAGATACTA AGCTGAATTT	3960
50	TATTGATGAT TCAGTTTAGT ACTTTAAGCC ATTTTAAATA AATGCAAATC AATCAAATAG	4020
	CACAGCTAAT CTAAATTGAA GGAGGTAGGC TCCTTGATTG ATGTAAATAA TTTCCATTAT	4080

55

	AAACCTGAAA CAATCAACTA CCGTACATTA AAACCTGAAA AAGATGGTCT ATTCTGTGAA	4200
	AGAATTTTCG GACCTACAAA AGACTGGGAA TGTAGTTGTG GTAAATACAA ACGTGTTCGC	4260
5	TACAAAGGCA TGGTCTGTGA CAGATGTGGA GTTGAAGTAA CTAAATCTAA AGTACGTCGT	4320
	GAAAGAATGG GTCACATTGA ACTTGCTGCT CCAGTTTCTC ACATTGGTA TTTCAAAGGT	4380
10	ATACCAAGTC GTATGGGATT ATTACTTGAC ATGTCACCAA GAGCATTAGA AGAAGTTATT	4440
	TACTTTGCTT CTTATGTTGT TGTAGATCCA GGTCCAACCTG GTTTAGAAAA GAAAACTTTA	4500
	TTATCTGAAG CTGAATTCAG AGATTATTAT GATAAATACC CAGGTCAATT CGTTGCAAAA	4560
15	ATGGGTGCAG AAGGTATTAA AGATTACTT GAAGAGATTG ATCTTGACGA AGAACTTAA	4620
	TTGTTACGCG ATGAGTTGGA ATCAGCTACT GGTCAAAGAC TTACTCGTGC AATTAAACGT	4680
	TTAGAAGTTG TTGAATCATT CCGTAATTCA GGTAACAAAC CTTTCATGGAT GATTTTAGAT	4740
20	GTACTTCCAA TCATCCACC AGAAATTCGT CCAATGGTTC AATTAGATGG TGGACGATTT	4800
	GCAACAAGTG ACTTAAACGA CTTATACCGT CGTGTAATTA ATCGAAATAA TCGTTTGAAA	4860
	CGTTTATTAG ATTTAGGTGC ACCTGGTATC ATCGTTCAAA ACGAAAAACG TATGTTACAA	4920
25	GAAGCCGTTG ACGCTTTAAT TGATAATGGT CGTCGTGGTC GTCCAGTTAC TGGCCCAGGT	4980
	AACCGTCCAT TAAAATCTTT ATCTCATATG TTAAAAGGTA AACAAGGTCG TTTCCGTCAA	5040
	AACCTACTTG GTAAACGTGT TGAATATTCA GGACGTTTCA TTATTGCAGT AGGTCCAAGC	5100
30	TTGAAAATGT ACCAATGTGG TTTACCAAAA GAAATGGCAC TTGAACTATT TAAACCATT	5160
	GTAATGAAAG AATTAGTTCA ACGTGAAATT GCAACTAACA TTAAAAATGC GAAGAGTAAA	5220
	ATCGAACGTA TGGATGATGA AGTTTGGGAC GTATTGGAAG AAGTAATTAG AGAACATCCT	5280
35	GTATTACTTA ACCGTGCACC AACACTTCAT AGACTTGGTA TTCAAGCATT TGAACCAACT	5340
	TTAGTTGAAG GTCGTGCGAT TCGTCTACAT CCACTTGTA CAACAGCTTA TAACGCTGAC	5400
40	TTTGACGGTG ACCAAATGGC GGTTACGTT CTTTATCAA AAGAGGCACA AGCTGAAGCA	5460
	AGAATGTTGA TGTTAGCAGC ACAAACATC TTGAACCCTA AAGATGGTAA ACCTGTAGTT	5520
	ACACCATCAC AAGATATGGT ACTTGGAAC TATTACCTTA CTTTAGAAAG AAAAGATGCA	5580
45	GTAAATACAG GCGCAATCTT TAATAATACA AATGAAGTAT TAAAAGCATA TGCAAATGGC	5640
	TTTGTACATT TACACACTAG AATTGGTGTA CATGCAAGTT CGTTCAATAA TCCAACATT	5700
	ACTGAAGAAC AAAACAAAAA GATTCTTGCT ACGTCAGTAG GTAAAATTAT ATTCAATGAA	5760
50	ATCATTCAG ATTCATTTGC TTATATTAAT GAACCTACGC AAGAAAACTT AGAAAGAAAG	5820
	ACACCAAACA GATATTTTCAT CGATCCTACA ACTTTAGGTG AAGGTGGATT AAAAGAATAC	5880

55

	GAAGTATTCA ACAGATTTAG CATCACTGAT ACATCAATGA TGTTAGACCG TATGAAAGAC	6000
	TTAGGATTCA AATTCTCATC TAAAGCTGGT ATTACAGTAG GTGTTGCTGA TATCGTAGTA	6060
5	TTACCTGATA AGCAACAAAT ACTTGATGAG CATGAAAAAT TAGTCGACAG AATTACAAAA	6120
	CAATTCAACC GTGGTTTAAT CACTGAAGAA GAAAGATATA ATGCAGTTGT TGAAATTTGG	6180
	ACAGATGCAA AAGATCAAAT TCAAGGTGAA TTGATGCAAT CACTTGATAA AACTAACCCA	6240
10	ATCTTCATGA TGAGTGATTG AGGTGCCCGT GGTAAACGCAT CTAACTTTAC ACAGTTAGCA	6300
	GGTATGCGTG GATTGATGGC CGCACCATCT GGTAAAGATTA TCGAATTACC AATCACATCT	6360
15	TCATTCCGTG AAGGTTTAAC AGTACTTGAA TACTTCATCT CAACTCACGG TGCACGTAAA	6420
	GGTCTTGCCG ATACAGCACT TAAAACAGCT GACTCAGGAT ATCTTACTCG TCGTCTTGTT	6480
	GACGTGGCAC AAGATGTTAT TGTTCTGTGA GAAGACTGTG GTACTGATAG AGGTTTATTA	6540
20	GTTTCTGATA TTAAAGAAGG TACAGAAATG ATTGAACCAT TTATCGAACG TATTGAAGGT	6600
	CGTTATTCTA AAGAAACAAT TCGTCATCCT GAAACTGATG AAATAATCAT TCGTCCTGAT	6660
	GAATTAATTA CACCTGAAAT TGCTAAGAAA ATTACAGATG CTGGTATTGA ACAAATGTAT	6720
25	ATTGCTCAG CATTTACTTG TAACGCACGA CATGGTGTTT GTGAAAAATG TTACGGTAAA	6780
	AACCTTGCTA CTGGTGAAAA AGTTGAAGTT GGTGAAGCAG TTGGTACAAT TGCAGCCCAA	6840
	TCTATCGGTG AACCAGGTAC ACAGCTTACA ATGCGTACAT TCCATACAGG TGGGGTAGCA	6900
30	GGTAGCGATA TCACACAAGG TCTTCCTCGT ATTCAAGAGA TTTTCGAAGC ACGTAACCT	6960
	AAAGGTCAAG CGGTAATTAC GGAAATCGAA GGTGTCGTAG AAGATATTAA ATTAGCAAAA	7020
	GATAGACAAC AAGAAATTGT TGTTAAAGGT GCTAATGAAA CAAGATCATA CCTTGCTTCA	7080
35	GGTACTTCAA GAATTATTGT AGAAATCGGT CAACCAAGTTC AACGTGGTGA AGTATTAACT	7140
	GAAGGTTCTA TTGAACCTAA GAATTACTTA TCTGTGCTG GATTAAACGC GACTGAAAGC	7200
	TACTTATTAA AAGAAGTACA AAAAGTTTAC CGTATGCAAG GTGTAGAAAT CGACGATAAA	7260
40	CACGTTGAGG TTATGGTTCG ACAAATGTTA CGTAAAGTTA GAATTATCGA AGCAGGTGAT	7320
	ACGAAGTTAT TACCAGGTTC ATTAGTTGAT ATTCATAACT TTACAGATGC AAATAGAGAA	7380
45	GCATTTAAAC ACCGTAAGCG TCCTGCAACA GCTAAACCAG TATTACTTGG TATTACTAAA	7440
	GCATCACTTG AAACAGAAAG TTTCTTATCT GCAGCATCAT TCCAAGAAAC AACAAGAGTT	7500
	CTTACAGATG CAGCAATTAA AGGTAAGCGT GATGACTTAT TAGGTCTTAA AGAAAACGTA	7560
50	ATTATTGGTA AGTTAATTCC AGCTGGTACT GGTATGAGAC GTTATAGCGA CGTAAATAC	7620
	GAAAAACAG CTAAACCAAGT TGCAGAAGTT GAATCTCAAA CTGAAGTAAC GGAATAACAA	7680
55		

ATGTTGACGA ATTCTCTTGT TCAATGTAA TATATTAAAG GTTGATGCAA GCAGAACTTT 7800
 GGAGGATAAA TTATTGTCTA AGGAAAAAGT tGCACGCTTT AACAAACAAC ATTTTGTAGT 7860
 5 TGGTCTTAAA GAAACGCTTA AAGCGTTAAA GAAAGATCAA GTTACATCTT TGATTATTGC 7920
 TGAAGACGTT GAAGTATATT TAATGACTCG CGTGTTAAGC CAAATCAATC AGAAAAATAT 7980
 10 ACCTGTATCT TTTTTCAAAA GCAAACATGC TTTGGGTAAA CATGTAGGTA TTAACGTCAA 8040
 TGCACAAATA GTAGCATTGA TTAAATGAGA ATTAGTAAGT GTTTTACTTA CTAAATTTTA 8100
 TTAAACCTAA AAATGAACCA CCTGGATGTG TGGGATTAAA AAGTGAAGAG AGGAGGACAT 8160
 15 ATCAGATGCC AACTATTAAC CAATTAGTAC GTAAACCAAG ACAAAGCAA ATCAAAAAAT 8220
 CAGATTCTCC AGCTTTAAAT AAAGGTTTCA ACAGTAAAAA GAAAAAATTT ACTGACTTAA 8280
 ACTCACCACA AAAACGTGGT GTATGTACTC GTGTAGGTAC AATGACACCT AAAAAACCTA 8340
 20 ACTCAGCGTT ACGTAAATAT GCACGTGTGc gTtTATCAA CAACATCGAA ATTAACGCAT 8400
 ACATCCCTGG TATCGGACAT AACTTACAAG AACACAGTGT TGTACTTGTA CGTGGTGGAC 8460
 GTGTAAAAGA CTTACCAGGT GTGCGTTACC ATATTGTACG TGGAGCACTT GATACTTCAG 8520
 25 GTGTTGACGG ACGTAGACAA GGTGTTTCAT TATACGGAAC TAAGAAACCT AAAAACTAAG 8580
 AATTTAGTTT TTAATTAAAT CTTAACTTA AAATATTTAA TATAAGGAAG GGAGGATTTA 8640
 CATTATGCCT CGTAAAGGAT CAGTACCTAA AAGAGACGTA TTACCAGATC CAATTCATAA 8700
 30 CTCTAAGTTA GTAACATAAT TAATTAACAA AATTATGTTA GATGGTAAAC GTGGAACAGC 8760
 ACAAAGAATT CTTTATTCAG CATTGACCT AGTTGAACAA CGCAGgtTCG TGATGCATTA 8820
 GAAGTATTCG AAGAAGCAAT CAACAACATT ATGCCAGTAT TAGAAGTTAA AGCTCGTCGC 8880
 35 GTAGGTGGTT CTAACATCA AGTACCAGTA GAAGTTCGTC CAGAGCGTCG TACTACTTTA 8940
 GGTTFACGTT GGTAGTTAA CTATGCACGT CTTGCTGGTG AAAAAACGAT GGAAGATCGT 9000
 TTAGCTAACG AAATTTTAGA TGCAGCAAAT AATACAGGTG GTGCCGTTAA GAAACGTGAG 9060
 40 GACTCTACA AAATGGCTGA AGCAAACAAA GCATTTGCTC ACTACCGTTG GTAAGATAAA 9120
 AGCTTTTACC CTGAGTGTGT TCTATATTAA TGAATTTTCA TTAAGCGTTC ATGCTTAGGG 9180
 45 CATCGCCATA TCTATCGTAT TTATTCAGTA ATATAAACTG GAAGGAGAAA AAATACATGG 9240
 CTAGAGAATT TTCATTAGAA AAAACTCGTA ATATCGGTAT CATGGCTCAC ATTGATGCTG 9300
 GTAAACGAC TACGACTGAA CGTATTCTTT ATTACACTGG CCGTATCCAC AAaGknGGTG 9360
 50 AAaCACACGA AGGTGCTTCA CAAATGGACT GGATGGAGCA AGAACAAGAC CGTGGTATTA 9420
 CTATCACATC TGCTGCAACA ACAGCAGCTT GGAAGGTCA CCGTGTAAC ATTATCGATA 9480

55

	CAGTTACAGT	ACTTGATGCA	CAATCAGGTG	TTGAACCTCA	AACTGAAACA	GTTTGGCGTC	9600
	AGGCTACAAC	TTATGGTGTT	CCACGTATCG	TATTTGTAAA	CAAAATGGAC	AAATTAGGTG	9660
5	CTAACTTCGA	ATACTCTGTA	AGTACATTAC	ATGATCGTTT	ACAAGCTAAC	GCTGCTCCAA	9720
	TCCAATTACC	AATTGGTGCG	GAAGACGAAT	TCGAAGCAAT	CATTGACTTA	GTTGAAATGA	9780
	AATGTTTCAA	ATATACAAAT	GATTTAGGTA	CTGAAATTGA	AGAAATTGAA	ATTCCTGAAG	9840
10	ACCACTTAGA	TAGAGCTGAA	GAAGCTCGTG	CTAGCTTAAT	CGAAGCAGTT	GCAGAACTA	9900
	GCGACGAATT	AATGGAAAAA	TATCTTGGTG	ACGAAGAAAT	TTCAGTTTCT	GAATTAAAAG	9960
	AAGCTATCCG	CCAAGCTaCt	AcTAACGTAG	AATTCTACCC	AGTACTTTGT	GGTACAGCTT	10020
15	TCAAAAACAA	AGGTGTTCAA	TTAATGCTTG	ACGCTGTAAT	TGATTACTTA	CCTTCACCAC	10080
	TAGACGTAA	ACCAATTATT	GGTCACCGTG	CTAGCAACCC	TGAAGAAGAA	GTAATCGCGA	10140
20	AAGCAGACGA	TTCAGCTGAA	TTGCTGTCAT	TAGCGTTCAA	AGTTATGACT	GACCCTTATG	10200
	TTGGTAAATT	AACATTCTTC	CGTGTGTATT	CAGGTACAAT	GACATCTGGT	TCATACGTTA	10260
	AGAACTCTAC	TAAAGGTAAA	CGTGAACGTG	TAGGTCGTTT	ATTACAAATG	CACGCTAACT	10320
25	CACGTCAAGA	AATCGATACT	GTATACTCTG	GAGATATCGC	TGCTGCGGTA	GGTCTTAAAG	10380
	ATACAGGTAC	TGGTGATACT	TTATGTGGTG	AGAAAAATGA	CATTATCTTG	GAATCAATGG	10440
	AATTCACAGA	GCCAGTTATT	CACTTATCAG	TAGAGCCAAA	ATCTAAAGCT	GACCAAGATA	10500
30	AAATGACTCA	AGCTTTAGTT	AAATTACAAG	AAGAAGACCC	AACATTCCAT	GCACACACTG	10560
	ACGAAGAAAC	TGGACAAGTT	ATCATCGGTG	GTATGGGTGA	GCTTCACTTA	GACATCTTAG	10620
	TAGACCGTAT	GAAGAAAGAA	TTCAACGTTG	AATGTAACGT	AGGTGCTCCA	ATGGTTTCAT	10680
35	ATCGTGAAAC	ATTCAAATCA	TCTGCACAAG	TTCAAGGTAA	ATTCTCTCGT	CAATCTGGTG	10740
	GTCGTGGTCA	ATACGGTGAT	GTTACATTG	AATTCACACC	AAACGAAACA	GGCGCAGGTT	10800
40	TCGAATTTCGA	AAACGCTATC	GTTGGTGGTG	TAGTTCCTCG	TGAATACATT	CCATCAGTAG	10860
	AAGCTGGTCT	TAAAGATGCT	ATGGAAAATG	GTGTTTTAGC	AGGTTATCCT	TTAATTGATG	10920
	TTAAAGCTAA	ATTATATGAT	GGTTCATACC	ATGATGTCGA	TTCATCTGAA	ATGGCCTTCA	10980
45	AAATTGCTGC	ATCATTAGCA	CTTAAAGAAG	CTGCTAAAAA	ATGTGATCCT	GTAATCTTAG	11040
	AACCAATGAT	GAAAGTAACT	ATTGAAATGC	CTGAAGAGTA	CATGGGTGAT	ATCATGGGTG	11100
	ACGTAACATC	TCGTGCTGGA	CGTGTTGATG	GTATGGAACC	TCGTGGTAAT	GCACAAGTTG	11160
50	TTAATGCTTA	TGTACCACTT	TCAGAAATGT	TCGGTTATGC	AACATCATTA	CGTTCAAACA	11220
	CTCAAGGTCTG	CGGTACTTAC	ACTATGTACT	TCGATCACtA	TGCTGAAGTT	CCaAAATCaA	11280

55

	GCCTAGGTTA AAATACAAGG TGAGCTTAAA TGTAAGCTAT CATCTTTATA GTTTGATTTT	11400
	TTGGGGTGAA TGCATTATAA AAGAATTGTA AAATTCTTTT TGCATCGCTA TAAATAATTT	11460
5	CTCATGATGG TGAGAACTA TCATGAGAGA TAAATTTAAA TATTATTTTT AATTAGAATA	11520
	GGAGAGATTT TATAATGGCA AAAGAAAAAT TCGATCGTTC TAAAGAACAT GCCAATATCG	11580
	GTACTATCGG TCACGTTGAC CATGGTAAAA CAACATTAAC AGCAGCAATC GCTACTGTAT	11640
10	TAGCAAAAAA TGGTGACTCA GTTGACAAAT CATATGACAT GATTGACAAC GCTCCAGAAG	11700
	AAAAAGAACG TGGTATCACA ATCAATACTT CTCACATTGA GTACCAAACG GACAAACGTC	11760
	ACTACGCTCA CGTTGACTGC CCAGGACACG CTGACTACGT TAAAAACATG ATCACTGGTG	11820
15	CTGCTCAAAT GGACGGCGGT ATCTTAGTAG TATCTGCTGC TGACGGTCCA ATGCCACAAA	11880
	CTCGTGAACA CATCTTTTTC TCACGTAACG TTGGTGTACC AGCATTAGTA GTATTCTTAA	11940
20	ACAAAGTTGA CATGGTTGAC GATGAAGAAT TATTAGAATT AGTAGAAATG GAAGTTCGTG	12000
	ACTTATTAAG CGAATATGAC TTCCCAGGTG ACGATGTACC TGTAATCGCT GGTTCAGCAT	12060
	TAAAAGCTTT AGAAGGCGAT GCTCAATACG AAGAAAAAAT CTTAGAATTA ATGGAAGCTG	12120
25	TAGATACTTA CATTCCAACG CCAGAACGTG ATTCTGACAA ACCATTCATG ATGCCAGTTG	12180
	AGGACGTATT CTCAATCACT GGTGCTGGTA CTGTTGCTAC AGGCCGTGTT GAACGTGGTC	12240
	AAATCAAAGT TGGTGAAGAA GTTGAAATCA TCGGTTTACA TGACACATCT AAAACAACG	12300
30	TTACAGGTGT TGAAATGTTT CGTAAATTAT TAGACTACGC TGAAGCTGGT GACAACATTG	12360
	GTGCATTATT ACGTGGTGTT GCTCGTGAAG ACGTACAACG TGGTCAAGTA TTAGCTGCTC	12420
	CTGGTTCAAT TACACCACAT ACTGAATTCA AAGCAGAAGT ATACGTATTA TCAAAAGACG	12480
35	AAGGTGGACG TCACACTCCA TTCTTCTCAA ACTATCGTCC ACAATTCTAT TTCCGTACTA	12540
	CTGAAGTAAC TGGTGTGTTT CACTTACCAG AAGGTACTGA AATGGTAATG CCTGGTGATA	12600
	ACGTTGAAAT GACAGTAGAA TTAATCGCTC CAATCGCGAT TGAAGACGGT ACTCGTTTCT	12660
40	CAATCCGTGA AGGTGGACGT ACTGTAGGAT CAGGCGTTGT TACTGAAATC ATTAAATAAT	12720
	TTCTAATTTT TTAGATTTTA TATAAAAAGA AGATCCCTCA ATCGAGGGGt CTTTTTTTAA	12780
45	TGTGTAAATT TTGTAATGGC TATTCGATTT AGAAGAACAA TAATTGATGA AAGACTGACT	12840
	AATAAAACTT ATAACGATA ATACTGTTTA AATAAAATTG TTGAGTCTTG GACATTGTAA	12900
	AATGCTCCCT TCAAAGTTTT CATTTTTTTca ATGTCTACTT TGAAGGGAGC ATTCATTAG	12960
50	TTTATGTCTC AGATTCATAT CTTTCAATTA ATTTAAATGC TTAATTTGTT TTAAATACTT	13020
	GCTCTAATTC TATGATTTTT AAAAATACAG CTACAGCGTA TTTTAATGAT TTTTCATCAA	13080

55

TCAGAAAGAA TGCACCTGGT CGTACTTTCA AATAATGTGA AAAATCTTCT CCAATCATCA 13200
 TTAAATCTGA TTCATTAAAG CGTACATGTA AGTCATTTGT TGCTTCTTTA ATAAC TTGAT 13260
 5 ATGCTTTCTC GTTATTATGG ACAGGCCAAAT ACCCTTTAAT ATAATTCAAA TCATAGTTAA 13320
 TATCATTGTC TATTGCTAAA CCTGTAGAA GCTTATCCAT TTTGTCCATT ACATGATTCT 13380
 GTATATCTGA ATCGAAAGTT CTAAC TGTAC CTTTACAAA TGCTTGATCA GGAATAACGC 13440
 10 TATCTGTGGT GCCTGCTTGA ATCATTCCAA ATGAAAGTAC AGCTTGTTTA ACTGGATCGA 13500
 TCGTACGTA AATTATTTTT TGTGCACCTA AAATGAACTC TGCCATGATT ACTATTGGGT 13560
 CAATGGTTTC ATGAGGTTTG GCACCATGAC CACCACGACC TTTAAATGTG ACGCTAAATT 13620
 15 CATCTGGAGA GGCCATGATT GCGCCCGCAC GTGAATGAAT AGTTCCAGTA GGATAACCAC 13680
 TCCATAAATG TGTACCGTAA ATTCTATCTA CATTTTCCAG ACATCCAGCA TCTATCATTT 13740
 CTTGAGAACC ACCTGGCATG ATTTCTTCAC CGTACTGGAA TATTAATACA ACATTACCTT 13800
 20 CTAATAAATG TTTATGTTCA TCTAAAATCT CTGCTACAGT AAGTAAAATT GCTGTATGAC 13860
 CATCATGCCC ACACGCATGC ATACATCCTG GATTTTTAGA CTTATAAGGC ACATCGTTTA 13920
 ATTCCTCGAC AGGTAACGCA TCAAAGTCAG CTCTTAATGC AATGGTAGGT CCTGTGCCCCA 13980
 AGCCTTTAAA TGTGGCTTTG ATACCATGTC GGCCGATAGG AGTTTCAATA TCACAAGATA 14040
 ACTGGCTTAA TTGGTTAACA ATATAATCAT GTGTTTGAAA TTCTTCAAAA GATAACTCAG 14100
 30 GATATTGGTG TAAATAACGT CTGAGTTGAA TTGTTTTATT TTCTTTATTA TTTGCTAGTT 14160
 GGAACCAATC TAACACCCTT ATCACTACTT TCTAAAATAA TGTTTATAGT ATAACATTTT 14220
 ATGAAATTAT CGTACTAAAT GATTGCTTTG AGATATTTTA TCTATGAATG ATAAGGCTTT 14280
 35 CAAGTTATGT AGAATTACTG TATGATAAAG GTATTACCAA ACAATACTTA AGGGGGATTA 14340
 TATACTGTGG TTCAATCATT ACATGAGTTT TTAGAGGAAA ATATAAATTA TCTAAAAGAA 14400
 AATGGTTTGT ATAATGAAAT AGATACAATT GAAGGTGCAA ACGGACCAGA AATCAAAATC 14460
 40 AATGGGAAAT CATACTTAA CTTATCTTCA AATAATTATT TAGGACTAGC AACAAATGAA 14520
 GATTTGAAAT CaGctGCAAA AGCAGCTATT GATACACATG GTGTAGGTGC AGGCGCTGTT 14580
 CGTACAATCA ATGGTACATT AGATTTACAC GACGAATTAG AAGAAACACT AGCAAAATTT 14640
 45 AAAGGAACAG AAGCTGCAAT AGCTTATCAA TCAGGATTTA ATTGTAATAT GGCTGCTATT 14700
 TCAGCTGTCA TGAATAAAAA TGATGCTATT TTATCAGATG AGCTTAATCA TGCATCAATT 14760
 ATTGATGGAT GTCGCTTATC TAAAGCTAAA ATTATTCGAG TTAACCATTC AGACATGGAT 14820
 50 GATTTACGTG CGAAAGCAAA AGAAGCAGTT GAATCAGGTC AATACAATAA AGTGATGTAT 14880

55

5 ATTGCAGAAG AATTGGTTT ATTAACCTAT GTTGACGACG CTCATGGTTC AGGTGTTATG 15000
 GGTAAGGCG CTGGTACGGT TAAACATTTT GGTTTACAAG ATAAAATCGA TTTCCAAATA 15060
 10 GGTACGCTTT CTAAAGCAAT TGGTGTCTGT GCGGTTATG TAGCAGGTAC AAAAGAGTTA 15120
 ATAGATTGGT TAAAAGCACA ATCACGACCA TTCTTATTCT CTACATCATT AGCACCTGGG 15180
 GATACCAAAG CAATAACTGA AGCAGTTAAA AAGTTAATGG ATTCAACTGA ATTACATGAT 15240
 15 AAATTATGGA ACAATGCACA ATATTTAAAA AATGGATTGT CAAAATTAGG ATATGATACA 15300
 GGTGAGTCAG AACTCCAAT TACACCACTA ATTATTGGTG ATGAAAAAAC AACTCAAGAA 15360
 TTTAGTAAGC GTTTAAAAGA CGAAGGTGTC TATGTGAAAT CTATCGTTTT CCCAACAGTA 15420
 CCAAGAGGTA CAGGACGTGT AAGAAATATG CCTACAGCTG CACATACAAA AGACATGTTA 15480
 20 GATGAAGCAA TTGCGGCTTA TGAAAAAGTA GGAAGAGAAA TGAAGTTGAT TTAATATTTA 15540
 TTTATTCCCA CGGCAAATAT TGTCGTGGGC TTTTITTAAT GTTTAGTTTA TTAACAGT 15598

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 661 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

30 AAGTAAATCA ACTTACTGGG ATAAGAATAA AGGCGATTAT AGTAACAAGT TGATTTTATT 60
 CGAAAAACAT TTTGAACCGG TTCTGGGTAT CAAGATGCAA CATAGTGGAG GTCATAGCTT 120
 35 TGGCCACACG ATTATTACGA TTGAAAGTCA AGGAGATAAA GCAGTTCATA TGGGTGATAT 180
 ATTCCCAACT ACTGCACATA AAAATCCTCT ATGGGTAACG GCATATGATG ATTATCCTAT 240
 40 GCAATCGATT CGTGAAAAAG AACGCATGAT ACCATATTTT ATTCAGCAAC AATATTGGTT 300
 CTTGTTTTAT CATGATGAAA ACTACTTTGC TGTAATAATC AGCGATAATG GTGAAAACAT 360
 AGATGCATAT ATTTTACGTG AAACATTAGT TGATAATAAC TAAAATAAAG ATGTATTACT 420
 45 AAACAAATTT TCAAAAATAA AAAATGAGC CACATCCAAT CTTACTAATT AGGGTGTGGC 480
 TCATTTTAA GTTTTACgAT CCAAATCAAA TATGGaTAA ATTCgTATTA ACGCTCTACa 540
 ATGtTAATGA CTTCAACAGT ATATGCATCT GCATAAAAT CATAATGAAT ATTTTGACCA 600
 50 TTTTAAATAG TTGTAATTCC ACCTTGATAA ACTAAACGGT ATTTATCAGT TTCAGGATGA 660
 A 661

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5738 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

10 GCAGACGGTA CAGCAGTTAA AGTCGCACCA AaACTGTAGT GAATcTAATC GGTGcATTCT 60
 TTTTAGGATT AGTGTGCGG CTTATATATA TCTTCTTCAA AGTAATTTTC GATAAGCGAA 120
 15 TTAAAGATGA AGAAGATGTA GAGAAAGAAT TAGGATTGCC TGTATTGGGT TCAATTCAAA 180
 AATTTAATTA AGGATGGTTG CTACTTATGT CAAAAAAGGA AAATACGACA ACAACACTAT 240
 TTGTATATGA AAAACCAAAA TCAACAATTA GTGAAAAGTT TCGAGGTATA CGTTCAAACA 300
 20 TCATGTTTTTC AAAAGCAAAT GGTGAAGTAA AGCGCTTATT GGTTACTTCT GAAAAGCCTG 360
 GTGCAGGTAA AAGTACAGTT GTATCGAATG TAGCGATTAC TTATGCACAA GCAGGCTATA 420
 AGACATTAGT TATTGATGGC GATATGCGTA AgcCAACACA AAATATATT TTTAATGAGC 480
 25 AAAATAATAA TGGACTATCA AGCTTAATCA TTGGTCGAAC GACTATGTCA GAAGCAATTA 540
 CGTCGACAGA AATTGAAAAT TTAGATTTGC TAACAGCTGG CCCTGTACCT CCAAATCCAT 600
 CTGAGTTAAT TGGGTCTGAA AGGTTCAAAG AATTAGTTGA TCTGTTTAAT AAACGTTACG 660
 30 ACATTATTAT TGTCGATACA CCGCCAGTTA ATACTGTGAC TGATGCACAA CTATATGCGC 720
 GTGCTATTAA AGATAGTCTG TTAGTAATTG ATAGTGAAAA AAATGATAA AATGAAGTTA 780
 AAAAAGCAAA AGCACTTATG GAAAAAGCAG GCAGTAACAT TCTAGGTGTC ATTTTGAACA 840
 35 AGACAAAGGT CGATAAATCT TCTAGTTATT ATCACTATTA TGGAGATGAA TAAGTATGAT 900
 TGATATTTCAT AACCATATAT TGCCTAATAT CGATGACGGT CCGACAAATG AAACAGAGAT 960
 GATGGATCTT TTAACAACAAG CGACAACACA AGGTGTTACA GAAATCATTG TAACATCACA 1020
 40 TCACTTACAT CCTCGATATA CCACACCTAT AGAAAAAGTG AAATCATGTT TAAACCATAT 1080
 TGAAAGCTTA GAGGAAGTAC AAGCACTAAA TCTAAAGTTT TATTATGGTC AGGAAATAAG 1140
 45 AATTACCGAT CAAATCCTTA ATGATATTGA TCGAAAAGTT ATTAACGGTA TTAATGATTC 1200
 ACGCTATTTA CTAATAGAAT TTCCATCAAA TGAAGTTCCA CACTATACTG ATCAATTATc 1260
 TTTCGAATcA CAGAGTAAAG GCTTTGTACC GATTATTGCA CATCCAGAGC GGAATAAAGC 1320
 50 AATAAGTCAA AACCTTGACA TACTATACGA TTTAATTAAC AAAGGTGCTT TAAGTCAAGT 1380
 GACAACGGcG TCATTAGCGG GTATTTCCGG TAAAAAAATT AGAAAATTAG CAATTCAAAT 1440

	GTTCTTAATG AAAGACTTAT TTAATGATAA GAAATTACGT GATTATTATG AAGATATGAA	1560
	CGGATTIATT AGTAATGCGA AGTTAGTTGT TGATGATAAA AAAATTCCTA AACGAATGCC	1620
5	ACAACAAGAT TATAAACAGA AAAGATGGTT TGGGTTATAA ACAGCAAATG AGGGGTTTTA	1680
	TGGCACATTT ATCTGTGAAA TTGCGGCTTT TAATACTAGC ATTAATCGAT TCACTGATAG	1740
	TGACATTTTC AGTATTCGTA AGTTATTACA TTTTAGAACC GTATTTCAAA ACATATTCTG	1800
10	TCAAATTATT AATATTGGCA GCTATATCAC TATTCATATC GCATCATATT TCaGCATTTA	1860
	TTTTTAATAT GTATCATCGA GCGTGGGAAT ATGCCAGTGT GAGTGAATTG ATTTTAATTG	1920
	TTAAAGCTGT GACGACATCT ATCGTTATTA CGATGGTGGT CGTGACAATT GTTACAGGCA	1980
15	ATAGACCGTT TTTTAGATTG TATTTAATTA CTTGGATGAT GCACTTGATT TTAATAGGTG	2040
	GCTCAAGGTT ATTTTGGCGT ATTTATCGGA AATACCTTGG AGGTAAGTCA TTTAATAAGA	2100
20	AGCCAACITT AGTTGTTGGT GCTGGTCAAG CAGGTTCAAT GCTGATTAGA CAAATGTTGA	2160
	AAAGTGACGA AATGAAACTT GAACCGGTAT TAGCAGTCCA TGATGACGAA CATAAACGCA	2220
	ATATCACAAT TACTGAGGGT GTAAAAGTCC AAGGTAAAAT TGCGGATATT CCAGAACTAG	2280
25	TGAGGAAATA TAAGATTAAA AAAATCATCA TTGCAATTCC AACTATTGGT CAAGAGCGTT	2340
	TGAAAGAAAT TAATAATATT TGCCATATGG ATGGCGTTGA GTTATTGAAA ATGCCAAATA	2400
	TAGAAGACGT CATGTCTGGT GAGTTAGAAG TGAACCAACT TAAAAAAGTT GAAGTAGAAG	2460
30	ATTTACTAGG CAGAGATCCT GTTGAATTAG ATATGGATAT GATATCAAAT GAATTGACGA	2520
	ATAAACTAT TTTAGTTACG GGTGCAGGTG GTTCAATAGG ATCAGAAATT TGTAGACAAG	2580
	TTTGTAATTT CTATCCAGAA CGTATTATTC TACTTGGCCA TGGTGAAAAC AGTATTTATT	2640
35	TAATCAATCG TGAATTGCGA AATCGCTTCG GwAAAAATGT TGATATCGTT CCTATTATAG	2700
	CGGATGTGCA AAATAGAGCG CGTATGTTTG AAATTATGGA AACGTATAAA CCATACGCAG	2760
	TTTATCATGC AGCAGCACAC AAGCACGTGC CGTTAATGGA AGACAACCCT GAAGAAGCAG	2820
40	TACGTAATAA TATTTTAGGT ACGAAAAATA CTGCTGAAGC TGCTAAAAAT GCAGAGGTAA	2880
	AGAAATTCGT TATGATTCTT ACGGATAAAG CCGTTAATCC GCCTAATGTC ATGGGAGCTT	2940
45	CAAAGCGAAT TGCAGAAATG ATTATTCAAA GTTTAAATGA TGAAACGCAT CGAACAAATT	3000
	TTGTTGCAGT GAGATTTGGT AATGTACTTG GATCGAGAGG ATCTGTGATT CCACTTTTCA	3060
	AAAGTCAAAT TGAAGAAGGT GGGCCAGTTA CTGTGACACA TCCTGAAATG ACACGTTACT	3120
50	TTATGACAAT TCCTGAAGCT TCTAGACTAG TTTTGCAGGC AGGGGCATTA GCAGAAGGTG	3180
	GCGAAGTATT TGTGCTAGAT ATGGGAGAAC CAGTGAAAAT TGTAGATTTG GCACGTAATT	3240

55

	CCGGCGAAAA AATGTTTGAA GAGCTTATGA ATAAAGATGA GGTTCATCCT GAACAAGTAT	3360
	TTGAAAAAAT TTATCGTGGC AAAGTACAAC ATATGAAATG TAATGAAGTT GAAGCGATTA	3420
5	TTCAAGACAT CGTCAATGAC TTTAGTAAAG AAAAAATTAT TAACTATGCC AATGGCAAAA	3480
	AGGGAGATAA TTATGTTCTGA TGACAAAATT TTATTAATTA CTGGGGGCAC AGGATCATTC	3540
	GGTAATGCTG TTATGAAACA GTTTTATAGAT TCTAATATTA AAGAAATTCG TATTTTTTCA	3600
10	CGCGATGAGA AAAACAAGA TGACATTCGA AAAAAATATA ATAATTCAA ATTAAAGTTC	3660
	TACATTGGTG ATGTGCGTGA TAGTCAAAGT GTAGAAACAG CAATGCGAGA TGTTGATTAC	3720
	GTATTCCATG CAGCAGCTTT AAAACAAGTG CCGTCATGTG AATTCTTTCC AGTTGAGGCA	3780
15	GTGAAGACAA ATATTATTGG TACAGAAAAT GTCTTACAAA GTGCTATTCA TCAAAATGTT	3840
	AAAAAAGTCA TATGTTTATC TACAGATAAG GCAGCGTATC CTATTAATGC TAGGGGTATT	3900
20	TCAAAAGCAA TGATGGAAAA AGTATTCGTA GCCAAATCAA GAAATATTCG TAGTGAACAA	3960
	ACGCTTATTT GTGTGACAAG ATACGGTAAT GTGATGGCTT CAAGAGGATC AGTAATACCT	4020
	TTGTTTATCG ACAAATCAA AGCTGGAGAA CCTTTAACGA TTACAGATCC TGATATGACA	4080
25	AGATTTTAA TGAGCTTAGA AGATGCGGTA GAACTAGTTG TTCATGCATT TAAGCATGCA	4140
	GAGACAGGAG ATATTATGGT TCAAAAAGCA CCAAGCTCAA CGGTAGGGGA TCTTGCGACC	4200
	GCATTATTAG AATTGTTTGA AGCTGATAAT GCAATTGAAA TCATTGGTAC GCGACATGGA	4260
30	GAGAAAAAAG CAGAAACATT GTTGACGAGA GAAGAATACG CACAATGTGA AGATATGGGT	4320
	GATTATTTTA GAGTGCCGGC AGACTCCAGA GATTTAAATT ATAGTAATTA TGTGAAACC	4380
	GGTAACGAAA AGATTACGCA ATCTTATGAA TATAACTCCG ATAATACACA TATTTTAACG	4440
35	GTGGAAGAGA TAAAGAAAA ACTTTTAACA CTAGAATATG TTAGAAACGA ATTGAATGAT	4500
	TATAAAGCTT CAATGAGATA GGAGAGATTG ACGTTGAATA TTGTAATTAC AGGAGCAAAA	4560
	GGTTTGTAG GAAAAAAGT GAAAGCAGAT TTAAGTTCAA CGACAGATCA TCATATTTTC	4620
40	GAAGTACATC GACAACTAA AGAGGAAGAA TTAGAGTCAG CATTGTTGAA AGCAGACTTT	4680
	GTCGTGCATT TAGCGGGTGT TAATCGACCT GAACATGACA AAGAATTCAG CTTAGGAAAC	4740
45	GTGAGTTATT TAGATCATGT ACTTGATATA TTAAGTAGAA ATACGAAAAA GCCAGCGATA	4800
	TTATTATCGT CTTCAATACA AGCAACACAA GATAATCCTT ATGGTGAGAG TAAGTTGCAA	4860
	GGGGAACAGC TATTAAGAGA GTATGCCGAA GAGTATGGCA ATACGGTTTA TATTATCGC	4920
50	TGGCCAAATT TATTCGGCAA GTGGTGTAAG CCGAATTATA ACTCAGTGAT AGCAACATTT	4980
	TGTTACAAAA TTGCACGTAA CGAAGAGATT CAAGTTAATG ATCGGAATGT TGAACCTAACG	5040

55

5 ATTGAAAATG GTGTACCTAC AGTACCAAAC GTATTTAAAG TGACATTGGG AGAAATTGTA 5160
 GATTTATTAT ACAAGTTCAA ACAGTCACGT CTCGATCGAA CATTGCCGAA ATTAGATAAC 5220
 10 TTGTTTGAAA AAGATTTGTA TAGTACGTAT TTAAGCTATC TACCTAGTAC aGACTTTAGT 5280
 TAYCCCTTAC TTATGAATGT GGATGATAGG GGTTCCTTTA CAGAATTTAT AAAAACACCG 5340
 GATCGTGGTC AAGTTTCTGT AAATATTTCT AAACCAGGTA TTAATAAGG TAATCACTGG 5400
 CATCATACTA AAAACGAAAA ATTTCTAGTC GTATCAGGTA AAGGGGTAAT TCGTTTTAGA 5460
 CATGTTAATG ATGATGAAAT CATTGAATAT TATGTTTCTG GCGACAAAT AGAAGTTGTA 5520
 15 GACATACCAG TAGGATACAC ACATAATATT GAAAATTTAG GCGACACAGA TATGGTAACT 5580
 ATTATGTGGG TGAATGAAAT GTTTGATCCA AATCAGCCAG ATACGTATTT CTTGGAGGTA 5640
 TAGCGCATGG aAAAACTGAA rTTAATGACA ATAGTTGGTA CAAGGCCTGA AATCATTCTG 5700
 20 TTATCATCAA CGATTAAAGC ATGTGATCAA TATtTTAA 5738

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9062 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

ATCATCAACA AGAATGATAT TTTTCCCATC TACTATATCT TTTACCGCAG ATAACCTCAC 60
 TCTCACACCT TGCTCACGTA ATTCTTGAGT TGGTTGAATA AATGTTCTTG CAACATATTG 120
 35 ATTTTTAACT AGTCCCATTT CATATGGCAA ACCTATTTCT TCAGCATAAC CACTCGCAGC 180
 TGATAGCGAT gAATTGGGTA CACCGATGAC CATATCAGCA TTTACAGGGC TTTCTTGGGC 240
 40 TAATTTTTTA CCAGAAGCTT TACGTACTGC ATGGACATTT TTACCAGCTA TTGTTGAGTC 300
 TGGTCTAGCA AAATAAATAT ATTCCATCGC AGAAATTGCA GTTGTCGTAT GATGTGTATA 360
 AGATTTAACT GTAATACCTT TATCGTTAAT CACGACATAT TCACCTGCAT GAATATCTTG 420
 45 AACAAATTCT GCACCTAACA CATCTATTGC ACATGTTTCA CTTGCAAGGA TGTATGTCCC 480
 ATCTTTTCA TTACCTACAA CAAGTGGTCT GATAGCATTT GGATCTACTG CGCCATATAA 540
 CGCATCTTTA GTTAAAATCG CAAATGTAAA ACCGCCTTTA ACTTTTCGCA AACTTTCTTT 600
 50 CAACGCTTCC TCAAAAGTAG GAGCTTTACT TCGACGTATC AAATGCATAA TGACTTCAGT 660
 ATCAGAAGAC GAATGAAGA TAGCACCTTG TTTTCTAAA TTCTGACGCA ATGATTTAGC 720

	CGGTTGAATA TTTTCAATAC CTTTATTACC TGAAGTAGCA TAACGGACGT GACCAATTGC	840
	ATGTTGATAT CCTTTTAATC GTTCCATTTG ATCATCTTTA ATCGCTTCAG TTAGTAAGCC	900
5	TAATCCTCGC TCGCCTTTTA ATTCAATTTG ATCAGAAACA ACTATACCTG cACCTTCTTG	960
	ACCACGATGT TGCAAACTAT GAAGTCCCAT ATAtGTTAGT TGCGCTGctT CaGGATGATT	1020
	CCAAATACCA AACACGCCAC ATTCTTCGTT TAATCCTGAG TAGTTAAACA TTGaGCAATT	1080
10	GCCCCtTCCC ATATTTGTTT AATATCTGAA ACATTTTCAC TAATCTCTGT aTATGGTGTT	1140
	GTTACCTTGx aATTATCACT ATCTGTTAAA AGTCCAATTT CTATTGCATT ATCAATATTT	1200
15	AAAGTTTTAC CTGATTTAAC AGAAACAACA TATCGGCCTT GCGTCTCACT AAACAATTGT	1260
	GCATTTGTTA TATCTATTGA AGATTTTAAT CCTAAACCGT AATGCGCACT TAGTTTAGCT	1320
	AAGGTAATCA GTAAGCCACC TTTACCAACT GTTTGAACAT GTGATAATAG TCCTTCACGA	1380
20	ATAGCGGTCT TGATTGATTC ACCTTTTTCA ACTTCTGAAC TCAAATCTAA TGA CTCAAAT	1440
	TCATGATTAA CTTTGCCATA AATTAACTTT TCAAGTTGAC TACCACCAA GTCGTCCTTA	1500
	GTATCACCGA TTAAATATAA TTTATCTCCA ACTTGAGGTT CAAAATCATT TAAATAATTT	1560
25	ACATTTTCAA TCAAACCTAC CATTCCAACA ACTGGTGTTG GGAAAATAGA AGTACCTTTC	1620
	GTTTCGTTAT ATAAAGATAC ATTACCAGAA ACTACTGGTG TCTTAAGAAT GTCGCATGCT	1680
	TCTGCCATAC CTTTCGTTGA ATCTATCAAC TGTTGATAGA TTTCTTTCTT TTCAGGAGAA	1740
30	CCATAATTTA AACAATCTGT CATTGCTAAT GGTGTTGCAC CCACGGCAAT TAAATTTCGA	1800
	TAAGCTTCAG CTA CTACTACCAT CTTTCCACCT TCATATGGAT TGTTATATAC ATAACGCGCT	1860
	TCACCATCAA TTGTTGAAGC AATTGCCTTA TTTGTGCCTT CCACACGTAC TACCGATGCT	1920
35	TGAAGTCCTG GCTTAATTAT CGTATTGGCA CCAACTTGTT GGTGCTATTG ATCATATAAA	1980
	TAGTGTTTAG ATGCTATAGT CGGATGCTTA AGTAATTTAA AGAAAGTATC TTAAACATCG	2040
	ATGTGTGTAT AATCATTTTT AGAAGTATTA TAATCTTTTT CTTCTCCTTC TAAAATATAT	2100
40	ACAGGTGCTT CATCAGCTAG TGGTTCAACT GGAATGTCAG CATAAACTTC GTCATCATAT	2160
	GTTAAAACAA AACGATTGT ATCTGTAAC TCACTATAA CAGCACTATC CAATTCGTGC	2220
45	TTATCAAATA AATCTAAGAA TTTTGTTCG GTACCTTTTT CAACAACTAG TAACATACGT	2280
	TCTTGAGTTT CTGAAAGCAT CATTCATAA GGAGAAATAC CTGGCTCAG TGTTGGCACT	2340
	TGTTCTAATC TCAAATGTAA CCCACTACCA CCTTTTGCCG CCATTTTCAGA CGATGAAGAT	2400
50	GTTAAACCAG CAGCACC CAT ATCTGAATA CCAACTAATT CATCAAATGT AATTGCTTCA	2460
	AGTGTGCTT CCATTAATTT TTTACCTACA AATGGATCAC CGATTTGTAC AGAAGGTCGT	2520

55

	CGACCAGTTT TCAAACCAAC ATAAATGACC GAATTACCTA CACCTTTTGC TGTGCCTTTT	2640
	TGAATCATGT CGTGATTGaT AACACCAACA CACATTGCAT TAACAAGTGG ATTGCCATCA	2700
5	TAACGTTTCAT CAAATTCGAT TTCACCAGCA GTTGTGGaA TACCAATGCA GTTACCATAA	2760
	CCTCCGATAC CCTTTACAAC ACCTTTAAGT AATCTTTGGT TTTGTTTATT ATCTAATTCT	2820
	CCAAATCTAA GACTGTTTAA CAAATTAATA GGTCTAGCCC CAATAGAGAC AATGTCACGA	2880
10	ATGATTCCAC CAACGCCTGT AGCAGCCCCT TGATATGGTT CAATTGCTGA TGGATGATTG	2940
	TGAGACTCTA CTTTAAATAC TACGGCTTGA TTATCACCTA TATCGACTAC CCCTGCACCT	3000
	TCACCAGGCC CCATAAGCAC ATGGTcACCT GACGTAGGAA ATTGCTTTAA AAACGGTTTA	3060
15	GAATGTTTAT AAGAGCAATG TTCACTCCAC ATAACAGAAA AGATACCTGT TTCTGTAAAG	3120
	TTAGTTGTC TGcCTAAaAT ATCGCAAaCT TTTTCATATT CTTGATCaCT TAATCCCATa	3180
20	TCTTGATATA CTTTTTCAAG TTTAATTTCT TCAACGCTTG GTTCGATAAA TTTAGACATG	3240
	TTGTTCCCTC CAACTTTTTA CCATCGCTTC AAATAATTTC ACACCACTAT CAGTACCTAA	3300
	CAACGTTTCT AAAGCTCTTT CagGATGtGG CATCATGCCA CATAcATTGC CTTTTTCGTT	3360
25	AACAATTCCT GCAATATCAT CATATGAACC GTTCGGATTA TTCACATATT TCAGAATAAT	3420
	TTGATTGTIA GCTTTTAATT GTTGATATAT TTCATCAGTA CAATAATAAT GACCTTCACC	3480
	GTGAGCTACA GGATATATAA CTTTTTCACC TTGTTcATAA AGATTTGTAa ATGCCGTTTG	3540
30	ATTATTCACT ATTCTAACT CTTcATTTCT ACTAATAAAT AAATGTGAAT CGTTATGCAa	3600
	TAATGCACCA GGTAATAAGC CTATTTcAGT TAAaATTTGA AACCcATTAC AAACACCTAA	3660
	TACTGGCTTA CCTTCAGCTG CAAGACGTTT AACTTCCGAA ATAATCGGsG CTACACTAGC	3720
35	CATTGCCCCA GATCTTAAGT AATCCCCGAA TGAAaATCCA CCAGGAATAA GTACGCCATC	3780
	AAATÉCACTT AGTGATGTTT CTCTATAATC TACATATTCC GCTTCAACAC CACTTTTAAT	3840
	AGCAGCATTa AACATGTCTC TATCACAATT CGAACCTGGA AAAACAAGAA CCGCAaATTT	3900
40	CATTTTATGC ATTCTCCTTT TCATCATCTA ACACTTTATA GCTATATTCT TCAATCACTG	3960
	TATTTGCAAA CAATTTTTCA CTTAGAGTTG TAATAATGTT GTGTACCTTT TCATCACTAA	4020
45	CCTCATCCAC TGTCATATAT AATACTTTTC CTACACGAAT ATCATTCACT TGTGCATAAC	4080
	CTAAGTCATG TACAGCTCGA GTAAGCGTTT GTCCTTGCGT ATCTAATACT TGTGGTTGTA	4140
	ATGTGATATG TAGTTCAATT GTTTTCATTA TTTTAAATCC TCCAATTTGT TTAAAAATAT	4200
50	TTGATATGTT TCAATCAGTG ATCCAGTGTT ATTTCTATAT ACATCTTTAT CAAAGTTTGC	4260
	ATTGGTAGCT TTATCCCAAA TTCGACATGT ATCTGGAGAT ATTTCATCCG CTAAcAAAAT	4320

	ATCCATTAAAT TGTTCACAAC CATTATTAAT CTTTAATGCT TTGGATTTTA GTATTTCAAT	4440
	ATCTTCATCT GATGCTATAT TGAGCAATTT AACATGGTCA TCCGTTATCA ACGGATCATT	4500
5	TAACGCATCA TTTTATAGA AAAATCTAC AAGTGGTTCT CTAAAACTT CACCATTTTC	4560
	AAAACCTAAA CGCTTTGTAA TAGATCCACT AGCAATATTA CGAACAACTA CTTCTAATGG	4620
	AATTATTTTC ACAGGCTTAA CTAATTGTTT TGTTCAGAT AATTGTTTAA TAAAGTGAAT	4680
10	TTCTATTCCA TTTTCTTGTA AATATTTAAA TATAATAGAA GTAATTTGAT TATTTAATCG	4740
	CCCCTTACCT GCCATTGTGT CTTTCTTAGC CCCGTTTCCA GCAGTAACTT CATCTTTATA	4800
	TTCAACTCTT AATTCATTTT CTTGATTTGT TGAGAAAATG CGCTTCGCTT TTCCTTCATA	4860
15	TAATAATGTC ATGCTTTAAT TACTCCCCTC AAATTTAGCG TACATATCTT GTTCAGTTTG	4920
	GTTTACATCA TTCGTTAGTA CAGTCATATG CCCCATTTTT CTGCTATCTT TACGCTCAGA	4980
20	CTTACCATAA ATATGTAAGT GCCACTCTGG ATGTTCAATTA AATTCATTTT CCAATAAATC	5040
	TAAATCTTTA CCTAGTAAGT TCATCATGAC TGCTGGCTTT AATAATTCAA TTGAATTTGG	5100
	TAATGATTGT CCGGTAAGT CTAAAATATG AGTATCAAAT TGTGAATAAT CACATGCTTC	5160
25	AATTGAATAA TGTCCGGAAT TGTGAGGCCT TGGTGCTATC TCGTTCACAT ACAATTGGTT	5220
	GTTACTATCT ATAAAAAATT CAACTGTAAA TGTTCCAATG AAATGAATCG ATTGGATAAT	5280
	TTTATTAAT TGTCTTTTCG CCTCAGCTGT TTTATCTATT CTCGCTGGAA CAATTGTTTT	5340
30	GAAAAGTATT TGATTTCTAT GCTCATTTTC TTGTAATGGG AAAAAAGTGA TTTGATTGTT	5400
	GTTTCCTCTT GTAACAGTAA GAGATACTTC TTTCTTGATA TTCAAATATT TTTCAGCTAC	5460
	GCATTCACCTA GTTTCAATTA ATTTAAAACC TTCTTGTAAG TCTTTTTCGT TGTTAATTAA	5520
35	AACTTGACCT TTGCCATCGT AGCCACCAAA TCTAGTTTTT ACAATAAAAG GATATCCTAA	5580
	TGTTTCAATT GCTTTGTCAA TATCTGTAGA TTCTTTTACT GAAATGAACG GGACAACTTT	5640
	GGTACCAGCA CTTTTTAATG TTTCTTTTTC AGTTAAGCGA TCTTGTAATA ACTGTATAGC	5700
40	TTGGTAACCT TGCGGAATAT TGTACTTTTC ACATAATAGT TTTAATTGTT GGGCTGAAAT	5760
	GTTTTCAAAT TCATAAGTAA TCACATCACA TTTTGTCTCT AATTGATTGA GTGCCTTTTC	5820
	ATCGTCATAC TTGGCTTGTA TAAATTCGTG TGCAACGTAT CTACATGGAC AATCTTCAGA	5880
45	AGGATCCAAT ACAACCACTT TATAACCCAT TTTTGTAGCT GATTGTGCCA TCATCTTTCC	5940
	AAGCTGACCA CCACCAATAA TGCCAATAGT CGCACCAAAC TTTAATTTAT TGAAGTTCAT	6000
50	TTTGCATGTC CTCCACTTTT TGAATTAACG AAGATTCATA CTGATTTAGT TTTTCAACTA	6060
	AAGAAGGATT TTGAATACTT AACATTCTTG CTGCAAGTAT ACCTGCGTTT TTAGCACCTG	6120

55

	AAGAATCTAT ACCCTTTAAA CTTTTTGTTC CAATCGGCAC TCCAATAACT GGTAGCGTCG	6240
	TTAATGATGC AACCATACCT GGTAATGTG CCGCACC GCC AGCGCCTGCA ATGATAATGT	6300
5	TTATACCTCT TTCTCTCGCT TCAGAAGCAA ATTGAACCAT CATTTTTGGC GTACGATGTG	6360
	CGGATACTAC TTGTTTTTCG TACGGAATTT CAAAATAATC CAACATGTTA CAACTCTCTT	6420
	GCATAATTTT CCAATCGGAA GAACTGCCCA TAATGACTGC TACTTTCCTT TTGTACACCC	6480
10	TTTCAAAAGT TTGAATTGTG AATTACTTTA GTTGATATAT ATAGATATAG CATAACAAGC	6540
	AATTTCTGCT TTTTCAATCA AAAATCGAAC TTTATTTTGA TTTTTTATTT GAATTTACGT	6600
	CTTTTGCTAT GTAAATTAGT TTTATAAACT AACAAAGTTA GGATATTGAC AATAGGAGGA	6660
15	GAAGTTTTTA TGGTTGCTAA AATTTTAGAT GGTAAACAAA TTGCCAAAGA CTACAGACAG	6720
	GGGTTACAAG ATCAAGTTGA AGCGCTAAAA GAAAAGGGTT TTACACCTAA ATTATCCGTT	6780
20	ATATTAGTTG GTAATGATGG CGCTAGTCAA AGTTATGTTA GATCAAAAAA GAAAGCAGCT	6840
	GAAAAAATTG GTATGATTTG AGAAATCGTA CATTTGGAAG AAACAGCTAC TGAAGAAGAA	6900
	GTATTAAACG AACTAAATAG ACTAAATAAT GATGATTCTG TAAGTGGTAT TTTGGTACAA	6960
25	GTACCATTAC CAAAACAAGT TAGCGAACAG AAAATATTAG AAGCAATCAA TCCTGAAAAA	7020
	GATGTGGACG GTTTTCATCC AATAAATATA GGGAAATTAT ATATCGATGA ACAAACTTTT	7080
	GTACCTTGCA CACCGCTCGG CATCATGGAA ATATTAAAC ATGCTGATAT TGATTTAGAA	7140
30	GGTAAAAATG CAGTTGTAAT TGGACGAAGT CATATTGTG GACAACCACT TTCTAAGTTA	7200
	CTACTTCAAA AAAATGCATC AGTAACAATC TTACATTCTC GTTCAAAAGA TATGGCATCA	7260
	TATTTAAAAG ATGCTGATGT CATTGTGAGT GCAGTTGGTA AGCCTGGTTT AGTAACAAAA	7320
35	GATGTGGTCA AAGAAGGAGC AGTAATTATC GATGTTGGCA ATACGCCAGA TGAAAATGGC	7380
	AAATTAAAAG GTGACGTTGA TTATGATGCG GTTAAAGAAA TTGCTGGAGC TATTACACCA	7440
	GTTCCTGGTG GCGTTGGTCC ATTAACAATT ACTATGGTAT TAAATAATAC TTTGCTTGCA	7500
40	GAAAAAATGC GTCGAGGTAT TGATTCGTAA AGAGCCTGAG ACATAAATCA ATGTTCTATG	7560
	CTCTACAAAG TTATAATGGC AGTAGTTGAC TGAACGAAAA TTCGCTTGTA ACAAGCTTTT	7620
	TTCAATTCTA GTCAACCTTG CCGGGGTGGG ACGACGAAAT AAATTTTACG AAAATATCAT	7680
45	TTCTGTCCCA CTCCCTAATA ACTGAGTTTT AATGAAGTCT TTTAACCCAC ATTAAATATT	7740
	ATTTTGCAAT TGCAATGAAT AACAAGAAAA ATCTGGGACA TTAATCGATC AAATGCTCCC	7800
50	TTCAAAGTAG ACATTGAATA AATGAAGGCT TTGAAGGGAG CATTTCACTT TGTACTTGGC	7860
	TCAACAATTT TATATAGACA GTAGTTAATT GAATGAAAAT AAGCTTGTA CAAGTTTTCA	7920

55

GTTGGGGATG GGCCCCAACA CAGAAGCTGT GACTATGATA AAGTACTACT ACATAGTTAA 8040
 TCATTAGTGG TTCTTTATCA TTTTCGCCTC CCTTTTCTTA TTGTTTGTAT ACACAAAAAT 8100
 5 TTAAGTTCAA ACTGTGGAAT AAAGTTATAT TTGATTTCAA ATTATCCCTA AATTATTAAT 8160
 TKTACAATTG TGGCAGATTT TCAAAATAAT AATTATTTCC TCATTATTTA TAAATTTATA 8220
 TTTAAATTTT ATTCTTTATA GGGTAAGATT AGGACTATAG TATGATGTGT AATAATATA 8280
 10 AATTAAGGTA TAGTAAAGCT AACTCAGAAA TGACTTATCA TTCGGAGGTT ACATTATGAA 8340
 TAAACTATTA CAGTCATTAT CAGCCCTCGG TGTTTCTGCT ACACTAGTAA CACCAAATTT 8400
 AAATGCAGAT GCAACGACGA ATACTACACC ACAAATTAAA GGCGCTAATG ATATCGTTAT 8460
 15 TAAGAAAGGT CAAGATTATA ACCTTCTAAA CGGCATAAGT GCATTTGATA AAGAAGATGG 8520
 AGATTTAACC GATAAAATTA AAGTCGATGG CCAAATTGAT ACATCTAAAT CTGGTAAATA 8580
 20 TCAAATTAAA TATCATGTCA CTGATTGAGA TGGTGCAATT AAAATTCCA CTAGGTATAT 8640
 TGAGGTAAA TAGCCCTCAT CACTATACTG CAAATAAAAT GGTAGCAAAC GAACATGTTT 8700
 TGCTACCATT TTATTTGTTA TTCTAACTTC ATCTGCAACT TTAACCCAAA TATTGTATTT 8760
 25 TTTCTGTATA CCAAAGGACT ACCTATCAAA TTATTAAAC TTAAGTCTC TTTTAAAAA 8820
 AATGTTTGA TTTTGAACAA ACAAATTTCC ACTTTTCATT GTTTAACCAT AAATTACTTT 8880
 TGGCAAATTC CTATTAAAA TGTTGCGCT TCCTTTCAAT CAACTAGCCA TCATTTCAA 8940
 30 TTTATTAGAC AATTCAAAC TTTTTTTATT TTCATTCAAT TAACCTTAA TTGAAAGCTA 9000
 TTCTCAACTT TCCTTTTAAA TATGAAGCAA TTTTTTCAA AACGCTATTA GTCACAAAAT 9060
 GT 9062

(2) INFORMATION FOR SEQ ID NO: 86:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2738 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

AAATATTTTT TCAAACTAT GTGAAATGG ACCATGTCTA AATCATGTAA TAATGCAGyA 60
 CATAATGCCA ACGGTCTmTC TTTATTGTCC CATGCATCAT GACCAATAAA TGAATCATCA 120
 50 ATTAATCGTC TAACTATTTT ATACACACCT AAAGAATGTC CAAAGCGACT ATGTTCTGCT 180
 GTGTGAAAAG ATAGGTACAG TGTTCTAGT TGTCTAATTC GACGTAACCT TTGGAATTCC 240

	TCTTTAAAAA CTTTTCTTC TACTAATTTT AAATCTACAT ATGCGTTAGT CATTATTCCC	360
	CTCCTTTTCG TTTAATATAA TATTTAATTT ACTTAAAATG CTTTGTACAT AAGTGCTAAG	420
5	TCTAACTTTT CGCCATACAT TTCTGGCTCA TAAGAGCGTA AGATTGTAAA ACCTTGCTCT	480
	TTATAGTAAG CTA CTGCTTC TTCAATTTTA TTATCTACTT CTAAGTAAAC ACCTTCAAAT	540
	TTATCTTCAA AACGTGATAA TCCTTCATTT AACAAATGCTG TACCATAACC TGTATGTTGC	600
10	GATTCTGGTT TAACATAATG AGCTGATAAA TATAATTCTT CACCGTAAAT AAAGTTAGCA	660
	AAGCCAACGA TGTCAATACC TTCTTCAACG ACTAAGAATA ATGTTCTTG AAGCTTTTC	720
	TTTAAATGAT GTTCATTATA TGAAGCTCT AACAAGTGAT TAACTGTTGT CGCAGCGTAT	780
15	ATATTTAAGT ATGTATTAAA CCAAGCTTTA GTTGGCAGAT CTCTAATTTG AACACATCT	840
	TTTTCAAGTG CTGTCTTAC CTGGAACATG ACTTCTCCC CTTATTAACA AGTTTTAATA	900
	ACGGCATTAT ACCACAACCT GCTCAATACT TAATAAACAA TGATTGTCTA TTCAATTTAT	960
20	ATATCTATAT TTTCCGTTAA AATTAAAAAT AAAAAATAAC GAAGCAAAAA ATCACTTCGT	1020
	TTAGTATGAG GTATGTCTTA TTGCAATATA CTATTCCACT CAGTTGCACG TGCTAAGGCA	1080
25	TAGTTGTCTT TCATGATGTC ACCAGGCTTT TCAGCAGTTC CAATAATATA ACCATTTAAA	1140
	GTGGCACCTA TAAAGCTAA ACTATATTTT ATTTGCGTAA TTGCTGGTTC GCTTTTATTT	1200
	TTGGACAATC TCCACCAACT AAAATAACTC TAAATCCTT TTCGGCCATT TGTGCCTTAA	1260
30	AATTAGGATA TCGTTTATCT TGTAAATGTT CTGACCAATG TTCGATAAAT GCTTTCAATG	1320
	GTGCTGAAAT GCTATACCAA TACACTGGTG ATGCAAAAAT AATTGTATCA CTAGCCAATA	1380
	TTTTATCTAG AATCGGCAAA TAGTCATCGT CATATGAAGT AATAGTCTCT GCTGTATGTC	1440
35	TCACGTCACG TATCGGTTTA AACTGATGTT GTGTCACGTC AATCCATTGA TACTCTAAAT	1500
	CTTGCAAAGC GAATTTTGTT AATTGTGCAG TATTACCGTT TGGTCTACTC CCACCAAACA	1560
	AAACAGTAAT CATTTTAGCC TAACCTCACT TTGATTAAT AAATATCTGT GTTTTTCGTT	1620
40	ACCTAATTAT ACTATCATAA GCTTTGCCTA CCGAATAGTA AAACGCTTAC AACTTTTATA	1680
	TAAATTTGAC GAAATTTTCT CATGCCTTAT ATAACGTCGT TTGTGATACG GGGCTAATTC	1740
45	ATGATGAAAT TAGATACATA TATCACCATT AAATACAATT CATTTAGTCT TCAATCGGAA	1800
	ACAGTTCATC GATATATTGA ATCTCATCAT CTGATAAAAC GATATCTGCA GCTTTAATAT	1860
	TTTCAACGAC TTGTTCTGCA CGTTTTGCAC CAGGAATAAT CACATCGATA GCTGGTCTCG	1920
50	TTAAATAAAA TGCTAATACA ATGTTGCGAA TTGAAGTTTG ATGTGCTGCA GCTATGCTTT	1980
	CCAAAGCTTT TACGCGACGC ACATTTTCTT CAAATACACC TGGTTTAAAA TCACGACGTG	2040

GCTAATGGGA AATATGGAAT AAATGTGATT TGGTGATCAA CACAATATTG TAATACTGCC 2160
 TCATTTTCGC GATGCAATAA ATTATATTCT AACTGTACAA CATCAACGTA ACCATCTTTA 2220
 5 TTTGCTTCTT TAAGTTGATC TAATGTGAAA TTTGATACAC CAATTGCTTT AATCTTCCCT 2280
 TGTTCCTTAA GCTCTTGTA TGTGCAACT GCTTGATCTT TCGGAGTGTT GTTATCCGGA 2340
 AAATGAATAT AATATAAATC GATATAATCA GTTTGTAGAC GTTTCAAACT ATTCTCAACT 2400
 10 TGTGTGTTTTA AATATTCCGG TTGATTGTTT TGATGTACTT CTTGATTTTC ATCAAATTCA 2460
 TGAGACCCTT TCGTAGCAAT TTTAATTGTC TCTCGCGGAT ATTCTTTAAC AACTTCTCCA 2520
 ACCAATTCTT CTGATCGTTC TGGCCCATAA ATATATGCCG TATCTAATAA ATTAATACCA 2580
 15 TGATTAATGG CTTGACGAAC AACATCTTTT CCTTGTTCTT CATCTAAGTT CGGATATAAA 2640
 TTATGCCCAa CCTAtGCGTT CGTCCCAAGT GCGATTGGAA ACACTTCAAC ATCAGATTTA 2700
 20 CCTAAGTTTA CAAATTGCTn CATTAGACCC AGCnCCTT 2738

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9425 base pairs
 25 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

GATTAGATGA TATTTAACGA AAATTAaGrT GmAATACTtG AATGTArGaa GTCTGATGTC 60
 GAAAATAGCT ATTAAAATAG AGTAGACGTA ATGtAAATGA AAGCACCTAA AATAGAAAAA 120
 35 TTTCAAAAAT AGCGTAATTA TTATAATAAA TAGACTGCCA ATAAAATGCA ATTTTTCACT 180
 TATAAcATTc TTCAAAAAAT AATAGCAAAA TTATGTAAAA AATATCTTGT CATGGCAAGA 240
 TTGGCTGTGC TATAATCTAT CTTGTGCTTA AGAACGGCTC CTTGGTCAAG CGGTTAAGAC 300
 40 ACCGCCCTTT CACGGCGGTA ACACGGGTTC GAGTCCCGTA GGAGTCACCA TTTTITAGGT 360
 CTCGTAGTGT AGCGGTTAAC ACGCCTGCCT GTCACGCAGG AGATCGCGGG TTCGATTCCC 420
 45 GTCGAGACCG TACAAATGCC TATCCAAGAG GATAGGCATT TTTTTCGTT TAATATTATA 480
 TTAATAAAAG ATATATGGAC GAATGATAAT CATATTGATT TATCTGTTCC TCCATTTTCT 540
 TTAATAATGTA TGAACCTCAA GTAACCTAGT GGTTGGATAT GAAAGATAAA CGTAGACAAT 600
 50 AAAATCTTTA TTAGACGTAC AAACATATGC TACTGTCAAC ATATTTCTTC GTTGTGATAT 660
 GCCACCAGTC CTCCATAACA TCAATTGTTA AAGTAACGAA TAACGAATAA TGATATTTAT 720

55

	GACCTCATCA TTGTGTTAAA TATCATTGTC ACAATCCGCC GTGAGAACT AATAAAAAAT	840
	AGTAATATAT AAGTTTATAT TGGAAAATAG AATTAATAGC TTATAAATGG TAAATTATAT	900
5	AATAGGTTAC TATACGTTAT AAGACGGAAA ATGCGCACAA TAACAAAAAT AGTAAGCGAC	960
	ATCCTGTGAT TTTTACACA AACATAAACG ATAAAGAACA AAAAATGATA AAATAATATT	1020
	AATGATTTAA GAAAAGAGGT TTATGCAAAT GGCTAGAAAA GTTGTGTAG TTGATGATGA	1080
10	AAAACCGATT GCTGATATTT TAGAATTTAA CTAAAAAA GAAGGATACG ATGTGTACTG	1140
	TGCATACGAT GGTAAATGATG CAGTCGACTT AATTTATGAA GAAGAACCAG ACATCGTATT	1200
	ACTAGATATC ATGTTACCTG GTCGTGATGG TATGGAAGTA TGTCGTGAAG TGCGCAAAAA	1260
15	ATACGAAATG CCAATAATAA TGCTTACTGC TAAAGATTCA GAAATTGATA AAGTGCTTGG	1320
	TTTAGAACTA GGTGCAGATG ACTATGTAAC GAAACCGTTT AGTACGCGTG AATTAATCGC	1380
20	ACGTGTGAAA GCGAACTTAC GTCGTCATTA CTCACAACCA GCACAAGACA CTGGAAATGT	1440
	AACGAATGAA ATCACAATTA AAGATATTGT GATTTATCCA GACGCATATT CTATTAAAAA	1500
	ACGTGGCGAA GATATTGAAT TAACACATCG TGAATTTGAA TTGTTCCATT ATTTATCAAA	1560
25	ACATATGGGA CAAGTAATGA CACGTGAACA TTTATTACAA ACAGTATGGG GCTATGATTA	1620
	CTTTGGCGAT GTACGTACGG TCGATGTAAC GATTTCGTCG TTACGTGAAA AGATTGAAGA	1680
	TGATCCGTCA CATCCTGAAT ATATTGTGAC GCGTAGAGGC GTTGGATATT TCCTCCAACA	1740
30	ACATGAGTAG AGGTGGAAC GAATGAAGTG GCTAAAACAA CTACAATCCC TTCATACTAA	1800
	ATTTGTAATT GTTTATGTAT TACTGATTAT CATTGGTATG CAAATTATCG GGTATATTT	1860
	TACAAATAAC CTGAAAAAG AGCTGCTTGA TAATTTTAAG AAGAATATTA CGCAGTACGC	1920
35	GAAACAATTA GAAATTAGTA TTGAAAAAGT ATATGACGAA AAGGGCTCCG TAAATGCACA	1980
	AAAAGATATT CAAAATTTAT TAAGTGAGTA TGCCAACCGT CAAGAAATTG GAGAAATTCG	2040
	TTTTATAGAT AAAGACCAAA TTATTATTGC GACGACGAAG CAGTCTAACC GTAGTCTAAT	2100
40	CAATCAAAAA GCGAATGATA GTTCTGTCCA AAAAGCACTA TCACTAGGAC AATCAAACGA	2160
	TCATTTAATT TTAAAAGATT ATGGCGGTGG TAAGGACCGT GTCTGGGTAT ATAATATCCC	2220
45	AGTTAAAGTC GATAAAAAG TAATTGGTAA TATTTATATC GAATCAAAAA TTAATGACGT	2280
	TTATAACCAA TTAAATAATA TAAATCAAAT ATTCATTGTT GGTACAGCTA TTTCATTATT	2340
	AATgCACAGT CATCCTAGGA TTCTTTATAG CGCGAACGAT TACCAAACCA ATCACCATA	2400
50	TGCGTAACCA GACGGTCGAA ATGTCCaGAG GTAACCTATAC GCAACGTGTG AAGATTTATG	2460
	GTAATGATGA AATTGGCGAA TTAGCTTTAG CATTTAATAA CTTGTCTAAA CGTGTACAAG	2520

55

	GTGATGGTAT TATTGCAACA GACCGCCGTG GACGTATTCG TATCGTCAAT GATATGGCAC	2640
	TCAAGATGCT TGGTATGGCG AAAGAAGACA TCATCGGATA TTACATGTTA AGTGTATTAA	2700
5	GTCTTGAAGA TGAATTTAAA CTGGAAGAAA TTCAAGAGAA TAATGATAGT TTCTTATTAG	2760
	ATTTAAATGA AGAAGAAGGT CTAATCGCAC GTGTAACTT TAGTACGATT GTGCAGGAAA	2820
	CAGGATTTGT AACTGGTTAT ATCGCTGTGT TACATGACGT AACTGAACAA CAACAAGTTG	2880
10	AACGTGAGCG TCGTGAATTT GTTGCCAATG TATCACATGA GTTACGTACA CCTTTAACTT	2940
	CTATGAATAG TTACATTGAA GCACTTGAAG AAGGTGCATG GAAAGATGAG GAACTTGCGC	3000
15	CACAATTTTT ATCTGTTACC CGTGAAGAAA CAGAACGAAT GATTGCGACT GTCAATGACT	3060
	TGCTACAGTT ATCTAAAATG GATAATGAGT CTGATCAAAT CAACAAAGAA ATTATCGACT	3120
	TTACATGTT CATTAAATAA ATTATTAATC GACATGAAAT GTCTGCGAAA GATACAACAT	3180
20	TTATTCGAGA TATTCCGAAA AAGACGATTT TCACAGAATT TGATCCTGAT AAAATGACGC	3240
	AAGTATTTGA TAATGTCATT ACAAATGCGA TGAAATATTC TAGAGGCGAT AAACGTGTCG	3300
	AGTTCACGT GAAACAAAAT CCACCTTTATA ATCGAATGAC GATTGCTATT AAAGATAATG	3360
25	GCATTGGTAT TCCTATCAAT AAAGTCGATA AGATATTCGA CCGATTCTAT CGTGTAGATA	3420
	AGGCACGTAC GCGTAAATG GGTGGTACTG GATTAGGACT AGCCATTTTCG AAAGAGATTG	3480
	TGGAAGCGCA CAATGGTCGT ATTTGGGCAA ACAGTGTAGA AGGTCAAGGT ACATCTATCT	3540
30	TTATCACACT TCCATGTGAA GTCATTGAAG ACGGTGATTG GGATGAATAA TAAGGAGCAT	3600
	ATTAATCTG TCATTTTAGC ACTACTCGTC TTGATGAGTG TCGATTGAC ATATATGGTA	3660
	TGGAACTTTT CTCCTGATAT TGCAAATGTC GACAATACAG ATAGTAAGAA GAGTGAAACG	3720
35	TAACCTTTAA CGACACCTAT GACAGCCAAA ATGGATACAA CTATTACGCC ATTTTCAGATT	3780
	ATTCAATCGA AAAATGATCA TCCAGAAGGA ACGATTGCGA CGGTATCTAA TGTGAATAAA	3840
40	CTGACGAAAC CTTTGAAAAA TAAAGAAGTG AAGTCCGTGG AACATGTTTCG TCGTGATCAT	3900
	AACTTGATGA TTCCTGATTT GAACAGTGAT TTTATATTAT TCGATTTTAC GTATGATTTA	3960
	CCGTTATCAA CATATCTTGG TCAAGTACTG AACATGAATG CGAAAGTACC AAATCATTTTC	4020
45	AATTTCAATC GTTTGGTCAT AGATCATGAT GCTGATGATA ATATCGTGCT TTATGCTATA	4080
	AGCAAAGATC GCCACGATTA CGTAAAATTA ACAACTACAA CGAAAAATGA TCATTTTTTA	4140
	GATGCATTAG CAGCAGTGAA AAAAGATATG CAACCATACA CAGATATCAT CACAAACAAA	4200
50	GATACAATTG ATCGTACGAC GCATGTTTTT GCACCAAGTA AACCTGAAAA GTTAAAAACA	4260
	TATCGCATGG TATTTAACAC GATTAGTGTT GAGAAAATGA ATGCTATACT ATTTGACGAT	4320

55

	GCAAACTATA ACGATAAAAA TGAAAAATAT CATTATAAAA ACCTGTCCGA AGATGAAGCG	4440
	AGTTCCAGCA AAATGGAAGA AACGATTCCA GGAACCTTTG ATTTTATTAA TGGTCATGGT	4500
5	GGTTTCTTAA ACGAAGACTT TAGATTGTTT AGTACGAATA ATCAGTCAGG CGAGTTAACA	4560
	TATCaACGTT TCctTAATGG TTATCCAACG TTTAATAAAG AAGGTTCTAA TCAAATTCAA	4620
	GTCACTTGGG GTGAAAAAGG CGTCTTTGAC TATCGTCGTT CGTTATTACG CACCGACGTT	4680
10	GTTTTAAATA GTGAGGATAA TAAATCGTTG CCGAAATTAG AGTCTGTACG TTCAAGCTTA	4740
	GCGAACAATA GTGATATTAA TTTTGAAAAA GTAACAAACA TCGCTATCGG TTACGAAATG	4800
	CAGGATAATT CAGATCATAA TCACATTGAA GTGCAGATTA ACAGTGAAGT CGTACCGCGT	4860
15	TGGTATGTAG AATATGATGG CGAATGGTAT GTTTATAACG ATGGGaGGCT TGaATAAATG	4920
	AACTGGaAAC TGACAAAGAC ACTTTTCATT TTCGTGTTTA TTCTTGTCaA CATCGTGTTA	4980
20	GTATCGATTT ATGTTAATAA AGTCAATCGC TCACACATTA ATGAAGTCGA GAGTAACAAT	5040
	GAAGTTAATT TTCAGCAAGA AGAAATTAAA GTACCGACTA GTATATTGAA TAAATCAGTT	5100
	AAAGGTATAA AATTAGAGCA AATTACAGGG CGATCAAAG ACTTTAGTTC TAAAGCTAAA	5160
25	GGCGATTCCG ATTTGACCAC ATCAGATGGT GGAAAATTAT TGAATGCGAA CATTAGTCAA	5220
	TCGGTAAAGG TCAGTGACAA TAACTTAAAA GATTTGAAAG ATTATGTTAA CAAGCGCGTA	5280
	TTTAAAGGTG CTGAATATCA ATTAAGCGAG ATTAGTTCAG ATTCTGTAAA ATATGAACAA	5340
30	ACGTATGATG ATTTTCCGAT TTTAAATAAC AGTAAAGCGA TGTTAAACTT TAATATAGAA	5400
	GATAACAAAG CGACTAGTTA TAAACAATCA ATGATGGATG ACATTAAGCC CACAGATGGT	5460
	GCAGATAAGA AGCATCAAGT GATTGGTGTG AGAAAAGCAA TCGAGGCATT ATATTATAAT	5520
35	CGTTACTTGA AAAAAGGTGA TGAAGTCATT AATGCTAGAC TCGGTTACTA CTCAGTCGTG	5580
	AATGAAACGA ATGTTCAATT GTTACAACCA AACTGGGAAA TTAAAGTGAA GCATGACGGT	5640
40	AAGGATAAAA CGAATACTTA CTATGTCGAA GCGACAAATA ATAACCCTAA AATTATTAAAT	5700
	CATTAAATATG AATCGTAATA AGCTAGCATT GCAAGCTCAT CATATGTGAG AAGCGGTGCT	5760
	AGCTTTTTTG CTGGTACGGT TTATTATGGC TGATGTTTTT GCGTCTCCAA CGTGCGCATT	5820
45	TATTCATATT TTAAGTAGAA CCGCATTGTA AAATTAGTGT AACTGTTATT TTAAAACTT	5880
	TAGTATTGTG CTAATCATTG TTATAATAAT TAAGAAATTC ATTGCACGTG ATTATCAAAA	5940
	TTTAAATATA AGAAACCGGT CGATGAACATA AAGTTACATA ATAGGAAAGG TATACAAAAC	6000
50	AGCTAATATA CTGATAGTTT CTGTAGGGAA AATCGTATAT TTGCACTGAT GTATATTGCA	6060
	GTCATATAGA GAGATTGACT GTTTAAAGAG AAAGGATGAG CCGCTTGATA CGCATGAGTG	6120

55

	TAGTTGATGT TGGTTTGACT GGAAAGAAAA TGGAGAATT GTTTAGTCAA ATTGACCGTA	6240
	ATATTCAAGA TTTAAATGGT ATTTTAGTAA CCCATGAACA TATTGATCAT ATTAAAGGAT	6300
5	TAGGTGTTTT GGCGCGTAAA TATCAATTGC CAATTTATGC GAATGAAAAA ACTTGGCAGG	6360
	CAATTGAAAA GAAAGATAGT CGCATCCCTA TGGATCAGAA ATTCATTTTT AATCCTTATG	6420
	AAACAAAATC TATTGCAGGT TTCGATGTTG AATCGTTTAA CGTGTCACAT GATGCAATAG	6480
10	ATCCGCAATT TTATATTTTC CATAATAACT ATAAGAAGTT TACGATTTTA ACGGATACGG	6540
	GTTACGTGTC TGATCGTATG AAAGGTATGA TACGTGGCAG CGATGCGTTT ATTTTTGAGA	6600
15	GTAATCATGA CGTCGATATG TTGAGAATGT GTCGTTATCC ATGGAAGACG AAACAACGTA	6660
	TTTTAGGCGA TATGGGTCAT GTATCTAATG AGGATGCGGC TCATGCAATG ACAGACGTA	6720
	TTACAGGTAA CACGAAACGT ATTTACCTAT CGCATTTATC ACAAGACAAT AACATGAAAG	6780
20	ATTTGGCGCG TATGAGTGTT GGCCAAGTAT TGAACGAACA CGATATTGAT ACGGAAAAAG	6840
	AAGTATTGCT ATGTGATACG GATAAAGCTA TTCCAACGCC AATATATACA ATATAAATGA	6900
	GAGTCATCCG ATAAAGTTCC GCATTGCTGT GAGACGACTT TATCGGGTGC TTTTTTATGT	6960
25	TGTTGGTGGG AAATGGCTGT TGTGAGTTG AATCGGCTTG ATTGAAATGT GTAAAATAAT	7020
	TCGATATTAA ATGTAATTTA TAAATAATTT ACATAAAATC AATCATTTTA ATATAAGGAT	7080
	TATGATAATA TATTGGTGTA TGACAGTTAA TGGAGGGAAC GAAATGAAAG CTTTATTACT	7140
30	TAAAACAAGT GTATGGCTCG TTTTGCTTTT TAGTGTAATG GGATTATGGC AAGTCTCGAA	7200
	CGCGGCTGAG CAGCATACAC CAATGAAAGC ACATGCAGTA ACAACGATAG ACAAGCAAC	7260
	AACAGATAAG CAACAAGTAC CGCCAACAAA GGAAGCGGCT CATCATTCTG GCAAAGAAGC	7320
35	GGCAACCAAC GTATCAGCAT CAGCGCAGGG AACAGCTGAT GATACAAACA GCAAAGTAAC	7380
	ATCGAACGCA CCATCTAACA AACCATCTAC AGTAGTTTCA ACAAAGTAA ACGAAACACG	7440
	CGACGTAGAT ACACAACAAG CCTCAACACA AAAACCAACT CACACAGCAA CGTTCAAATT	7500
40	ATCAAATGCT AAAACAGCAT CACTTTCACC ACGAATGTTT GCTGCTAATG CACCACAAAC	7560
	AACAACACAT AAAATATTAC ATACAAATGA TATCCATGGC CGACTAGCCG AAGAAAAAGG	7620
45	GCGTGTCATC GGTATGGCTA AATTAAAAAC AGTAAAAGAA CAAGAAAAGC CTGATTTAAT	7680
	GTTAGACGCA GGAGACGCCT TCCAAGGTTT ACCACTTTCA AACCAGTCTA AAGGTGAAGA	7740
	AATGGCTAAA GCAATGAATG CAGTAGGTTA TGATGCTATG GCAGTCGGTA ACCATGAATT	7800
50	TGACTTTGGA TACGATCAGT TGAAAAAGTT AGAGGGTATG TTAGACTTCC CGATGCTAAG	7860
	TACTAACGTT TATAAAGATG GAAAACGCGC GTTTAAGCCT TCAACGATTG TAACAAAAAA	7920

55

	TGAAGGCATT AAAGGCGTTG AATTTAGAGA TCCATTACAA AGTGTGACAG CGGAAATGAT	8040
	GCGTATTTAT AAAGACGTAG ATACATTGTG TGTTATATCA CATTTAGGAA TTGATCCTTC	8100
5	AACACAAGAA ACATGGCGTG GTGATTACTT AGTGAAACAA TTAAGTCAAA ATCCACAATT	8160
	GAAGAAACGT ATTACAGTTA TTGATGGTCA TTCACATACA GTACTTCAAA ATGGTCAAAT	8220
	TTATAACAAT GATGCATTGG CACAAACAGG TACAGCACTT GCGAATATCG GTAAGATTAC	8280
10	ATTTAATTAT CGCAATGGAG AGGTATCGAA TATTAAACCG TCATTGATTA ATGTTAAAGA	8340
	CGTTGAAAAT GTAACACCGA ACAAAGCATT AGCTGAACAA ATTAATCAAG CTGATCAAAC	8400
	ATTTAGAGCA CAACTGCAG AGGTAATTAT TCCAAACAAT ACCATTGATT TCAAAGGAGA	8460
15	AAGAGATGAC GTTAGAACGC GTGAAACAAA TTTAGGAAAC GCGATTGCAG ATGCTATGGA	8520
	AGCGTATGGC GTTAAGAATT TCTCTAAAAA GACTGACTTT GCCGTGACAA ATGGTGGAGG	8580
20	TATTCGTGCC TCTATCGCAA AAGGTAAGGT GACACGCTAT GATTTAATCT CAGTATTACC	8640
	ATTTGGAAAT ACGATTGCGC AAATTGATGT AAAAGGTTCA GACGTCTGGA CGGCTTTCGA	8700
	ACATAGTTTA GCGCGACCAA CAACACAAAA GGACGGTAAG ACAGTGTTAA CAGCGAATGG	8760
25	CGGTTTACTA CATATCTCTG ATTCAATCCG TGTTTACTAT GATATAAATA AACCGTCTGG	8820
	CAAACGAATT AATGCTATTC AAATTTTAAA TAAAGAGACA GGTAAGTTTG AAAATATTGA	8880
	TTTAAAACGT GTATATCAG TAACGATGAA TGACTTCACA GCATCAGGTG GCGACGGATA	8940
30	TAGTATGTTT GGTGGTCCTA GAGAAGAAGG TATTTTATTA GATCAAGTAC TAGCAAGTTA	9000
	TTTAAAAACA GCTAACTTAG CTAAGTATGA TACGACAGAA CCACAACGTA TGTTATTAGG	9060
	TAAACCAGCA GTAAGTGAAC AACCAGCTAA AGGACAACAA GGTAGCAAAG GTAGTAAGTC	9120
35	TGGTAAAGAT ACACAACCAA TTGGTGACGA CAAAGTGATG GATCCAGCGA AAAAACCAGC	9180
	TCCAGGTAAA GTTGTATTGT TgtAGCGCAT AGAGGAACTG TTAGTAGCGG TACAGAAGGT	9240
40	TCTGGTCGCA CAATAGAAGG AGCTACTGTA TCAAGCAAGA GTGGGAAACA ATTGGCTAGA	9300
	ATGTCAGTGC CTAAAGGTAG CGCGCATGAG AAACAGTTAT TTCATAATCA ACAGTCATTG	9360
	ACGTAGCTAA GTAATGATAA ATAATCATAA ATAAAATTAC AGATATTGAC AAAAAATAGT	9420
45	AAATA	9425

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3886 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

	AGTTGTAATG TCACATTTCC AGAGTCTGAA ATTATCTTTA TCACGTTACA TTTACTAGGC	60
5	TCTAAAATGA CTGAACATAC AGCATCTTCA ATTACCTTTG AATACCATGA TTTATCGCAA	120
	AATATACATG AATTGATCAC TTGTGTTAGC CAAGAATTAG GCATTGATAT GTCAAAAGAC	180
	AACAAGTTAC ATACCAGTCT GATCACACAT ATCAAACCAG CTATACATCG TATTAAATAC	240
10	GATATGCTAC AACCTAATCC TTTGAGGCAA GAAGTTATGC GTCGCTATCC TCAAATCATT	300
	GAAGCCGTTA GCAAGCATAT TAGTCCAATT GAACAAGATG CTGCTATTCTG CTTCAACGAA	360
	GATGAATTAA CATACATTAC AATTCACCTC GCATCAAGTA TAGAGCGTGT TGCAACACAT	420
15	AAACAATCAA TGATTAAGGT TGTCTTACTA TGTGGTTCTG GTATAGGCAC GTCACAACCTT	480
	TTAAAATCAA AACTAAATCA CCTGTATCCT GaGTTnCACa TTTGGGAtGc CTATTcCATT	540
20	TaTcAATTGG aAGaAAGTCG ATTATTGCAA GATAACATTG ATTATGTCAT TTCAACAGTA	600
	CCTTGTGAAA TATCAGCTGT ACCAGTTATT CATGTCGATC CATTATATCAA TCAACAATCT	660
	CGTCAAAAAT TGAATCAAAT TATCAATGAC TCAAGAGAAC AACGAGTCAT GAAAATGGCA	720
25	ACTGATGGCA AGTCACTCGC AGATTTATTG CCTGAACATC GCATCATTAT AAATAAACAA	780
	CCATTATCAA TTGAATCCGC AATTGCAGTG GCTGTGCAAC CTTTAATCAA TGATGGCATT	840
	GTCTATTCAA ATTATACAGC TGCAATTTTA AAACAATTTG AACCAATTCGG GTCATATATG	900
30	GTCATTAGTC CACATATTGC ACTTATTCAC GCTGGTACTG ATTATGTACA GAATGGTGTA	960
	GGTTTCGCAC TAACATATTT CACTGAAGGG ATTATCTTTG GTAGTAAAGC TAACGATCCC	1020
	GTTACCTTG TAATTACATT AGCAACGGAC CACCCCAATG CACATTTAA GGCATTGGGA	1080
35	CAGTTAAGCG AATGCTTAAG CAACGACTTA TATCGACAAG ATTTCTTAGA TGGGAATATT	1140
	TTTAAATTA AACACACAT TGCTTTAACT ATGACAAAGG AGGCTTAATA ACGTGTCAAT	1200
	AGACATTTTG TCAACAACAC GCATCATTGT AAAAGAACAA GTAAATGATT GGAAGTGAAGC	1260
40	TATAACTATA GCTTCTCAGC CATTACTACA AGAACAAATT ATTGAACAAG GCTATGTTCA	1320
	AGCAATGATT GATAGCGTTA ATGAACCTGG ACCTTATATC GTTATCGCAC CTGAAATTGC	1380
45	AATTGCACAT GCAAGACCGA ACAATGACGT ACATCAAGTT GGTTTAAGTC TATTAAAGTT	1440
	GAATCAACAT GTGGCATTTT GTGATGAAGA TCACTACGCA TCTCTCATTT TTGTATTGAG	1500
	TGCCATCGAC AATCATTAC ACTTATCTGT ATTACAAAT TTAGCAACCG TACTGGGCGA	1560
50	TAACCAAACA GTCCAGCAAC TATTAACTGC AACAAATGCA CAAGACATTA AAAACATTTT	1620
	AAAGGAGCAT GATTAAATATG AAAATTTTAG TAGTATGTGG CCACGGTTTA GGAAGTAGTT	1680

55

	AAGTTGAACA TAGTGACATT ATGACAGCAA GTCCAGAGAT GGCTGACTTG TTTATTTGTG	1800
	GTAGAGATTT AGCTGAAAAT GCCGAACGTC TAGGGGATGT CTTAGTTCTT GATAATATTT	1860
5	TAGATAAAGC TGAATTACAA CAAAAGCTCT CAGAAAAATT ACAACAACCTT AACATGATTT	1920
	AAAGGAGGTA CGACCTATGC AAGCAATCCT TAATTTTATA GTCGATATTT TAAGTCAACC	1980
10	AGCCATTCTT GTTGCACTGA TTGCCTTTAT AGGTTTAATC GTTCAGAAAA AACCTGCCGC	2040
	AACGATCACT TCAGGAACCA TTA AACGAT ATTAGGCTTC TTAATTTTAA GTGCAGGTGC	2100
	TGATGTCGTC GTTCGATCTC TTGAACCATT CGGCAAAATA TTCCAACACG CATTGGTGT	2160
15	GCAAGGTATC GTACCTAACA ACGAAGCTAT CGTCTCACTA GCCTTAAAG ATTTTGAAC	2220
	AACAGCTGCA CTCATCATGG TCTGTGGCAT GATTGTAAAT ATTTTAATTG CCCGCTTCAC	2280
	TAATTTAAAA TATATCTTTT TAACAGGTCA TCATACATTT TACATGGCTG CGTTTTTAGC	2340
20	AATCATTTTA ACAGTCAGTC ATATTAAAGG CTGGCTAACG ATTGTTATCG GCGCACTCGT	2400
	ATTAGGATTA ATCATGGCAG TATTACCTGC ATTACTCCAA CCTACGATGC GAAAAATTAC	2460
	AGGGAATGAC CAAGTAGCTT TAGGTCATTT TGGCTCAATC AGTTACTTTG CCGCAGTGCT	2520
25	GTAGGTCAAT TATTCAAAGG TAAGTCTAAA TCAACGGAAG AGATTAAATT TCCAAAAGGC	2580
	TTAAGTTTCT TACGAGAAAG TACAATTAGT ATCTCGATTA CGATGGCATT ACTTTACTTC	2640
	ATCGCATGCT TATTTGCGGG CGTTAGTTAT GTACACGAAT CTATTAGTGA TGGTCAAAAC	2700
30	TTTATTGTCT TTTCATTAAT TCAAGGTGTG ACATTTGCTG CTGGTGTATT TATTATTTTA	2760
	ACGGGCGTTC GTTTAATCTT AGCTGAAATC GTCCCAGCAT TTAAAGGAAT TTCTGAAAAG	2820
35	CTTGTAACAA ATTCTAAACC TGCATTAGAC TGCCCTATTG TGTTCCCTTA TGCACAAAAT	2880
	GCAGTATTAA TTGGATTCTT TGTGAGCTTT ATTACAGGTG TCATCGGTAT GTTTATCTTA	2940
	TTCTTATTTG GTGGCGTCGT CATTTTACCT GCGGTAGTTG CACACTTCTT CTTAGGTGCA	3000
40	ACGGCTGCTG TATTCGGTAA TGCAAGAGGC GGTATTAAAG GTGCTATTGc TGGCGCCGCT	3060
	CTAAATGGTA TCCTAATCAC GTTTTACCA TTATTATTCT TGCCATTTTT AGGCGAATTA	3120
	GGTGGTGCTG CAACAACATT CTCAGATACA GACTTTTTTAG CTGTCGGTAT CGTGTTCCGT	3180
45	AACGCAGTAA AATATATGGG ATTATTTGGT GCGATTCTAT TTATTATTAT CGTAGGTGCG	3240
	ACAACAATTT TATTAAAAGG CCGTCAAAAA GAACAGCAAT AGTGTTAACG TAGAAATATA	3300
	AAACACCGTC ACATATTGAG TGAATGCCCC TTTtATCAAG AGGAAAGCCA CTTACTTATG	3360
50	GACGGTGTTT TGTATTATAT TAAATGATAC TTAGCCATAC TATCGACAGC TGCTAAAATT	3420
	GCTTCTTCTT GTGTCGCAAT CGGTTCCCAA CCAAGTAATG TTTTtgCACG TTCGTTACTT	3480

55

CCTAGACTCA AAATAAAGTC TGGTAATTTT TTAGTAGAAA CTTTTTGAGC TATTTTCAGGT 3600
 CTCCTTTCTT TAATTAATTT TGCAATTTCC AACAAATTAA TTTGTCCATC AGCCGTCGCA 3660
 5 ATAAATCGCT TGCCATTAGC TTGTTCAATTT GTCATTGCCA AAATGTGCAG TTCAGCTACG 3720
 TCTCTCACAT CAACAACATT TAACGGAATT TGCGGTACAC GTTTCATTGA ACCATTCAAT 3780
 AAATTTTCTA ATAAATGAAA GCTTCCTGAA ACGTGTGCAT CTAATGATGG CCCAAAAATT 3840
 10 GCAACTGGAT TGATTGTGGC AAATTCTACT GTTGTATTTT CATTCT 3886

(2) INFORMATION FOR SEQ ID NO: 89:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4879 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

25 GTCATCTATC AAAAATTTGG TATACAGACC GACAATTATT AATTAATAAT TTAATTTCCC 60
 AGGCAATACC AGTGATTAAA TATCCACAAA TACAACATAA AGAACAACCA TTAGAATCTA 120
 TTTCACAACT TATATTGTCT AAGATGACAT CTAATCAATA GTGTTTAAAT TTCTCAGTGG 180
 CTGTGAATGA GGTTTAAAAG TACTATAAAA CGTAAACTTT GATACTTTAA AATACGCAAA 240
 30 AAACGGTAAA CCCTAATTCA TATTATAGAG TTTACCGTTT TATTTTTTAA CTTGCATCAT 300
 AGTTATATTA ACATTATTGT TGGTAGTTTG GATCAGTAAC CATTGCTTGT CCAGTATAAT 360
 CAACCGTTAC AATTGAATAT TTTCCaTTTG CATTTGGGTC TTTAAACTA AACACATACT 420
 35 TATAGTTGCC ATTATGTTCT TCAATAGAAT AATCATTATA CACTTTATTA TTACTACCAA 480
 ATTTTATTGC TTCATTATTA GCCGCATTTA AAGCTGTTTG GAAATTTGGC AATTGCTGTA 540
 AAGCTTGATT TTTATTTCCA TTAAACGGAT AAATTTGACG TGCAACCGGC GCGGCATTTT 600
 40 GACCATAATA TGGTGCAACG TAACTTGATT TTTGATTATT ATTCGCTTGG TTATTACTTG 660
 ATTGTTTATT ATTTGTTTGG TTTTGGTCAT TGTTTGTTGC ATTTGAATTA GATTGTTGCT 720
 45 GGTATCGTT TGCATATTA TCTTTATTAT CTTTGTTTAC GTCTTTACTA TCATCTTTAT 780
 TATCTTTCTT ATCTTTAGAT GAATCATTTG TTTTTTATC TTGTTGTTCA GTTTTCGCTT 840
 TATCATCTTT TTCTTTATTA CCGTCTTTTT GTTGGTCACT ATCTTGACCA CATGCAGCTA 900
 50 AAAATAATGA TAATGCTAGT AACCTGTAA CTAATCTTTT CACATATATC TCCTCCTATA 960
 ATTCGATATT CATTGAATAA TCTTGAAATA CATATCTACC ATGTGTATCT TTTCATGGCT 1020

55

	TAAGGTTCTT TTTATTATAC CCTAATTTTT GTTCATTATT ATTAAATTTT TGTGAATTTT	1140
	ATG=TTKCTA TAAATTTAAT TATTTTACTT TAACAATTCA TTACGCATTT AGCATTTCAA	1200
5	GGTATACACA ATATTTATTA CTATGATTTC ATTTTATCTG CTGCAAAAAC AATCATTATA	1260
	ACTCTTTTTC CATAATTAAA TCTGTATCCG TTACATCACC TGTGTGAAAA TGATGTTCAC	1320
	CAACCACTTT AAATCCATGA CGTTTATAAA ATGCTTGAGC ACGAGGATTA TGCTCCCAAA	1380
10	CTCCTAGCCA AATTTTATGT TTATTATGTT CTTGAGCAAT TTTTTCGGCC AATTCCTATCA	1440
	ATTGTGAACC TCTTCCGCCA CCTTGAAAGT CTTTCAAAAA ATATATGCGC TGCACTTCTA	1500
	AATAGGTCTC CCCCATTCTC TCAGTTTGAG CACTATTAAT ATTCATCTTT ATATAACCAA	1560
15	CATTCGCACC ATCTTCTTGA TAAAAATAAT GAAATGAATC TACATGGTTA ATCTCTTGTC	1620
	TAAATTTCTC TACAGTATAA TTGTCTTTAA AAAATTGATC AAAATCTTTG TCATCATAGT	1680
20	AAGAACCAAA CGTGTCTATA AATGTTCTAG TTGCTAATTC AACTAATTCA CTAGCATTTT	1740
	GTTCTGAAAT TTCTTTGATT ATCCCAGCCA TATAAATCCT CCAATAAACA GTGATCGAAT	1800
	CAAAATATTA CTTATGTTAT TTTTCAGCCA AACTATTTA AAAATACATT AACACAAATC	1860
25	AATTACAAAT TGTATTGATT GTGTGTAACA TCAATAAATG ATACATTTAT TCCAGTAAAA	1920
	TGGCCGTATT TTCAAAGAG AAAAAGAGAG GATGTATCGT TGTGATAGAA ACATTTAAAG	1980
	CGTTTGTAAT TGATAAAGAT GAGAGTGGTA AAGTGACACC AACTTTCAA CAATTATCGC	2040
30	CTACTGATTT ACCTAAAGGA GATGTGCTGA TTAAAGTACA TTACTCTGGT ATAAATTATA	2100
	AAGATGCTTT AGCGACTCAA GATCATAATG CAGTCGTAAA ATCGTATCCT ATGATTCCAG	2160
	GAATAGATTT AGCTGGAACA ATTGTTGAAT CCGAAGCACC AGGCTTTGAA AAAGGAGAAC	2220
35	AAGTAATTGT AACGAGTTAT GACCTAGGTG TCAGCCATTA TGGCGGTTTT AGTGAATATG	2280
	CGCGTGTAAA ATCAGAAATG ATTATCAAGC TTCCTGATAC TTAAACATTA GAAGAATCAA	2340
40	TGATATATGG CACAGCTGGT TATACTGCCG GTTTAGCAAT TGAAAGACTT GAAAAAGTTG	2400
	GAATGAATAT TGAAGATGGT CCTGTACTCG TTCGCGGTGC TTCAGGTGGT GTCGGTACTT	2460
	TAGCAGTACT CATGCTTAAT GAACTTGGTT ATAAAGTTAT CGCAAGTACA GGTAAACAAG	2520
45	ATGTTAGCGA TCAATTACTT GAACTTGGTG CCAAAGAAGT TATCGATCGA CTTCTGTGTG	2580
	AAGATGATCA TAAAAAGCCA CTCGCATCAT CAACTTGGCA AGCTTGTGTA GACCCTGTTG	2640
	GTGGCGAAGG TATTAAATTAT GTTACAAAGC GTTTAAATCA TAGTGGGTCA ATTACAGTTA	2700
50	TTGGTATGAC TGCCGGTAAT ACTTATACTA ATTCTGTATT CCCTCACATT TTAAGAGGTG	2760
	TAAACATTTT AGGAATTGAC TCGGTATTTA CTGCTATGAA ATTAAGACAG CGCGTTTGGC	2820

55

	TTGATGAACT	TCCAGAACAA	CTTAACAAAG	TAATTAAACA	TGAAAATAAA	GGGCGCATTG	2940
	TTATCGATTT	CGGTGTAGAT	AAATAGTATT	CATGAAAAAG	ACATCCCGTT	ATGCGAGATG	3000
5	TCTTTTTTAA	TTTAGTATTT	GATATACATA	CCGCTGAAT	CTGGTTCGGT	AGGTATAAAT	3060
	CCAAATTTTG	TATATAATTT	ATCCGCTGGG	TAGTCTGCAA	TCAGACTAAC	GTATGTACTC	3120
	TCAACAGCCA	CACCTTTAAT	ATATTGCATA	ATATGCTCCA	TAATTAGACT	GCCGTAACCT	3180
10	TGACCTTGGT	AACCTTTCAA	AACTGCAATA	TCAACAATTT	GAAAAACAGT	TCCGCCATCG	3240
	CCAATCACTC	TACCCATACC	AATTAACCGA	TCTTTATCAT	ACAAGGTTAC	TGTAAATAAG	3300
	GCATTAGGTA	ATCCTTTTTC	AGCTGTTGCG	GCGTCTTTGG	ACTCATACCT	GCGTTAATCC	3360
15	TTAATGCGCA	ATAATCCTCG	CAAGTCGGAA	TATCATATGT	CACTTTAAAC	ATTATTTACC	3420
	CCACTTTTCA	TCACACAATA	TATCAACCTA	GTATAAATGT	TTATTTACAA	TAGTCTTATT	3480
20	CGCTTCTTTA	AACACTTCAT	GATGACTTGA	AACATAACCC	TCTGCATTCT	CATCTGGTTG	3540
	GATATATGTT	TTAGCAAGGT	TCGCTGCATT	TGCACCATCA	CTAAATGCAC	TTGCAATTAG	3600
	ATGTGATTTT	GCATCATGAT	AAACAATATC	TCCACACGCA	TAGATACCAG	GTATACTAGT	3660
25	TGTCGTATTA	CCAAATCCTT	TAACACGACA	ATCATCATGC	ATATCTAGCT	TTGAAGATGT	3720
	TtCACTCAAT	AATGTATTAC	AACGATCAAA	CCCATGACTA	ATAATGACAT	CGTCAAATTT	3780
	AACTGTATGC	CTATCGCCAC	TTTCAACATG	TTCCAAAACA	ACTTCACCTA	TATGCGTTTC	3840
30	ATCATCATTG	CCGACCAAGT	ATTTAATACG	TGTTTTTGGG	CATAGTTTCA	CATTTAAATC	3900
	TGTCACCAAC	GTTTTTCATG	CTTCATGACC	ACTTACATCT	TCTTTTCGAT	AAACAACTGT	3960
	CACGCTTTTA	GCAATCTTGG	CAATATCATG	CGCCCAATCT	AATGCTGTAT	TTCTCCACC	4020
35	TGATATTAAT	ACATCTTTAT	CTTTGAAACG	TCTGTAACCT	TGTACAACAT	AATGTAAATT	4080
	AGTTtATTGA	TATCTCTCTA	CACCTTTAAC	ATCTAATTGT	TTTGGATTAA	TAATACCCGC	4140
	ACCAATTGCA	ATGATAACTG	CTTTCGATGT	ATATATTTCT	CCCGCTTCTG	TTTCAACTTC	4200
40	GAAATGACGT	TCTGCCTTTT	TCCTAATATC	TACCACACGT	TCATTCAAAT	GAACCTCCGG	4260
	TTTAAATAT	AATCCTTGCT	TAATTGTATC	TTTTAAAATT	TCATGACAAG	GTTTTGGCGC	4320
45	AATGCCGCCA	ATATCCCAA	TAATTTTTTC	AGGGTAAATT	CTCATCTTAC	CCCCTAATTC	4380
	AGATTGAACA	TCTATCAATC	TTACAGACAT	ATCTCGCAAT	CCAGCATAAA	AGCTTGATA	4440
	CAAACCAGAC	GGACCGCCAC	CAATGATTGT	AACATCTTTC	ATTATGTGCC	TCCTATGACT	4500
50	CTCTATATTC	ATTTCTTTCA	TTAACGTGCT	CAAATTGATA	ATTATTATCA	TTTAAAGCCA	4560
	TTATACTATT	AATATTTATA	TTGTTAAAT	AAATCGCATA	GTTAGCCATG	AATTATCAAT	4620

55

GAAAGATGTG TATATTTTTT AGTTC TAGTT ATATTATTTT TTAAAAGACT CATCACGTGG 4740
 TTCTTTAAGA ATTGCTTGTC TTAAAAGGAA AAATAGCAAC AATAAACCTG CAAGCATACC 4800
 5 TGTGTGCCCCA ATACCTGCAA AGCCTGcNAA TGCTTCTGGA GAGTATGATT TACCA GTGAC 4860
 TTGGAAGAAT CCTTTTGTC 4879

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1560 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

20 ATAATGTCTT AGaTTGATTG GGAGTTTTTT TAATTTTTTT GAAATTAAAT TAATCTGTAs 60
 YTAATAAAAA ATTTGAATAA CTGACACaYT TTTTGTATCA TAGCTaYATA CTTTGTGAAT 120
 TAATTCACAT TATAATAAGA GTGAAGATAA GAGTATTATA AATnATCTTT AAATAAATAT 180
 25 ATGTGAAGTA AAAATTACAC GTTAGCATAT CGATTATGgT CATTTckTTT AACATATTAA 240
 CTgGGGaACG TTAAAAGTTA ACGGkTGATA TcYAAcLAAA AACAAGGTCA CAGTAGTATG 300
 TTTTAATCTG GCGTCTATTA CAAATAAAAA TTACATCTAT AATTATTcGT TTTCTTTTTT 360
 30 GAAAGTAATA GCCAATTAAT ATCATACATA CTGGAGTGAC TATAAGGAGG ACATTATTAT 420
 GAGAGCAGCA GTTGTAAcGA AAGATCACAa AGTAAGTATT GAGGACAAAA AGTTAAGAGC 480
 TTTAAAACCT GGTGAAGCGT TGGTACAAAC GGAATATTGT GCGGTTTGTC ATACCGATTt 540
 35 ACATGTTAAG AATGCTGATT TTGGTGATGT TACAGGCGTT ACTTTAGGTC ATGAAGGTAT 600
 TGGTAAAGTC ATCGAAGTTG CGGAAGATGT AGAATCATTa AAAATTGGAG ACCGTGTGTC 660
 40 TATCGCTTGG ATGTTcGAAA GCTGTGGAAG ATGTGAATAT TGTACAACAG GTCGTGAAAC 720
 ACTTTGCCGT AGTGTGAAAA ATGCTGGTTA TACAGTAGAT GGTGCAATGG CTGAACAAGT 780
 TATTGTTACT GCAGACTATG CTGTGAAAGT ACCTGAAAAA TTAGATCCAG CAGCAGCGTC 840
 45 TTCTATTACA TGCGCAGGTG TGACAACTTA TAAAGCTGTA AAAGTAAGTA ATGTAAAACC 900
 TGGACAATGG TTAGTGTTT TTGGTATAGG TGTTTAGGT AACCTAGCTT TACAATATGC 960
 TAAAAACGTT ATGGGGGCTA AAATTGTTGC CTTcGACATC AATGATGATA AATTAGCATT 1020
 50 CGCGAAAGAA TTAGGTGCTG ATGCTATTAT TAATTCTAAA GATGTTGATC CAGTTGCAGA 1080
 AGTTATGAAA TTAAC TGATA ACAAAGGATT AGATGCAACA GTGGTAACTT CAGTTGCTAA 1140

TTTACCTGTT GATAAAATGA ACTTAGATAT CCCAAGATCA GTGCTTGATG GTATTGAAGT 1260
 AGTAGGTTCA CTTGTTGGTA CAAGACAAGA CTTACGTGAA GCGTTTGAAT TTGCTGCTGA 1320
 5 AAATAAAAGTA ACACCTAAAG TTCAATTAAG AAAATTAGAA GAAATCAATG ATATTTTGA 1380
 AGAAATGGAA AATGGTACTA TAACTGGTAG AATGGTTATT AAATTTTAAA AATATCAACT 1440
 GACTATATAG ATAAAGAAGG TAGTGCTCTG AACACTATCA TTATTAATCA AACCCCGAGG 1500
 10 TTTTCCTGAA AAGATAGTGG AAAATCCCCG TGTTTTTTGG GTTTGAGGnG GTTGThTGTA 1560

(2) INFORMATION FOR SEQ ID NO: 91:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11014 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

GTCCTGTnGC TGCAATGAAT ACGCCTAAAA ATCCAGGGAT GTAATGGATA CTTTGTGGTA 60
 25 GTACTAATGA TAGAAATGAT AAAAATGAAA TCACAAAGGC TACGCTCGCA AAAGCTTGAC 120
 ATGTACGCTT ATCGCCATAA TCTAACCCCTG TACGTATATG TAATAAATAC TGTAATCCGA 180
 TACTTAAATA CATAATTGCC ACGCATAAGA AGAATGGGAA GAATGTCTTT TCAAAGTCCG 240
 30 GATATAGGCT GTTAGATAGG AAGACCATGA TGAACATATT AAACATCATA AACGAGACGT 300
 CTTTGAATGT AACTTGACCA AATCGATTG TAAAAAATGT TTGATGAGAC CACATTAACC 360
 ATAAGAACAA ACTCATGACG ATGTATTGTA AAAATAAATC AGCTGAAATG GAACCGTTTT 420
 35 GTGTTGTAA AATCACATGT GCAATTTTTT GAATGGCATA GACGAAAATT AAATCAAAGA 480
 ACAACTCATG GAATCCTGCA CGCTTTTCAG CTAAATGTTT TGGTGTTAAT GCATTAACCA 540
 TAAAATTTTA ACTCCTTTAA GATGTGTAAT TAATTTACTA AGTATACTAT TTATTTTTC 600
 40 TAGTGAATAG GGCAGATTT GGCAGTAAG TGAAGGAGA GGTGACTGCA AGGTAATTGC 660
 GGAATTAACA ATCATCAGCG ATTAAATATT TGA CTGGAGA CGTCATGGTA ATAAAAAATT 720
 45 GATGAGAAAT TGATGGTGAA ACCAGCTGTG AATAsCGaTG CAATGATrSA TAGaATTTAA 780
 TTAGAGTCAT TACGCGaAAT GATTAATGAT AATTTGTGGT AAATCAAAGC aTAATTTTGT 840
 ACTATAGATG AGGATGATAG AGCATATTTA AGAGGGTGAA ATGTTAAAGT GAAACCGTTT 900
 50 ACGTTTCCGA TTGCCCAAAC AAATTACATC ATTGTATAAT ATGATTTGTT AAATGCATAA 960
 CAAGAATGAA AATGTAACAT ACGTAGCAAT TGGTTTCATA AATTGGATGT TAGTGGCGTA 1020

55

	TGACGAGAGT	CGTATTAGCA	GCAGCATACA	GGACACCTAT	TGGCGTTTTT	GGAGGTGCGT	1140
	TTAAAGACGT	GCCAGCCTAT	GATTTAGGTG	CGACTTTAAT	AGAACATATT	ATTAAAGAGA	1200
5	CGGGTTTGAA	TCCAAGTGAG	ATTGATGAAG	TTATCATCGG	TAACGTACTA	CAAGCAGGAC	1260
	AAGGACAAAA	TCCAGCACGA	ATTGCTGCTA	TGAAAGGTGG	CTTGCCAGAm	ACAGTACCTG	1320
10	CATTTACGGT	GaATAAAGTA	TGTGGTTCTG	GGTTAAAGTC	GATTCAATTA	GCATATCAAT	1380
	CTATTGTGAC	TGGTGAAAAT	GACATCGTGC	TAGCTGGCGG	TATGGAGAAT	ATGTCTCAAT	1440
	CACCAATGCT	TGTCAACAAC	AGTCGCTTTG	GTTTTAAAT	GGGACATCAA	TCAATGGTTG	1500
15	ATAGCATGGT	ATATGATGGT	TTAACAGATG	TATTTAATCA	ATATCATATG	GGTATTACTG	1560
	CTGAAAATTT	AGTAGAGCAA	TATGGTATTT	CAAGAGAAGA	ACAAGATACA	TTTGCTGTAA	1620
	ACTCACAACA	AAAAGCAGTA	CGTGCACAGC	AAAATGGTGA	ATTTGATAGT	GAAATAGTTC	1680
20	CAGTATCGAT	TCCTCAACGT	AAAGGTGAAC	CAATCGTAGT	CACTAAGGAT	GAAGGTGTAC	1740
	GTGAAAATGT	ATCAGTCGAA	AAATTAAGTC	GATTAAGACC	AGCTTTCAAA	AAAGACGGTA	1800
	CAGTTACAGC	AGGTAATGCA	TCAGGAATCA	ATGATGGTGC	TGCGATGATG	TTAGTCATGT	1860
25	CAGAAGACAA	AGCTAAAGAA	TTAAATATCG	AACCATTGGC	AGTGCTTGAT	GGCTTTGGAA	1920
	GTCATGGTGT	AGATCCTTCT	ATTATGGGTA	TTGCACCAGT	TGGCGCTGTA	GAAAAGGCTT	1980
	TGAAACGTAG	TAAAAAAGAA	TTAAGCGATA	TTGATGTATT	TGAATTAAAT	GAAGCATTTG	2040
30	CAGCACAATC	ATTAGCTGTT	GATCgTGAAT	TAAATTACC	TCCTGAAAAG	GTGAATGTTA	2100
	AAGGTGGCGC	TATTGCATTA	GGACATCCTA	TTGGTGCATC	TGGTGCTAGA	GTATTAGTGA	2160
35	CATTATTGCA	TCAACTGAAT	GATGAAGTTG	AACTGGTTT	AACATCATTG	TGTATTGGTG	2220
	GCGGTcAAC	TATCGCTGCA	GTTGTATCAA	AGTATAAATA	ATAAGAAAAC	AGGTTATCAC	2280
	AACAgtATTA	ATtACATGTT	GGCATAACCT	GTTTTTATTT	GTTTATGGAT	TTATTGGGTA	2340
40	ATATTAGTCA	TTTGATGGTT	TAATTGCAAA	TGCTCTAACA	GGGAACCCAG	GTGCATCTTT	2400
	TGGTTTAGGG	CTGATAGCGT	AAATGATGGC	GCCACGAGTT	GGTAATTGAT	CTAAATTAGT	2460
	TAATAACTCG	ACTTGGTATT	TATCCTGACC	AAGAATATAA	CGTTCGCCAA	CTAAATCACC	2520
45	ATTTTTTACA	ACGTCCACAG	ATGCATCGGT	ATCGAATGTT	TCATGACCAA	CAGCTTCAAC	2580
	ACGACGTTCT	TCAATTAAGT	ACTTCAAAGC	ATCTAATCCC	CAACCCGGTG	CATGTTGTTG	2640
	TCCGTTTCGA	TCTTTGTTTT	CAAACTTTTC	AATATTAGGC	CAACGTTTTG	ACCAATCGGT	2700
50	ACGAAGTGCA	ACAAAAGTGC	CAGGTTCAAT	AGTACCATGC	TCTTTTTCCC	ATGCTTCTAT	2760
	ATGCGCACGT	GTTACGATGA	AATCATGTGT	GTTGCTACT	TCTGTTGAAA	AGTCTAATAC	2820

55

	AAAGTGAATT GGTGCATCAA TGTGAGTACC ATATTGCGTT ACAATATTCC AACGTTGCAC	2940
	ATAGAAACCA TGATCTTTAA CCGTGAATAA AGTTGAACT TCGCCTTTTT CAAACTCACT	3000
5	AAAACGTGGT ATTTCCGGAT CAAATGTATG CGTTAAATCA ACCCAAGTTG CTTGTTTTAA	3060
	AGTATTTAAT TGTGCCATA AAGGATATTG TGTCAATAAA TCACCCGTTT TTAGTTTATT	3120
	ATATGATAAA TGCTGCGATT ATTCTTGGCG TTTAGCTTTA ACAGCATTCA CAAGCACAGT	3180
10	CAATGCATCT TTAACCTCTT CTTCTTTTCG CGTTTTTAAA CCACAGTCAG GGTTTACCCA	3240
	GAATAATGAG CGGTGCGATT GTTGTAGTGA ACGATTGATT GCTGTAGTAA TTTCTTCTTT	3300
	TGTTGGAATA CGTGGACTAT GAATATCATA TACACCTAGA CCAATACCTA AATCATAATT	3360
15	AATATCTTCA AAGTCTTTAA TTAAATCACC ATGGCTACGA GATGTTTCAA TTGAAATAAC	3420
	ATCAGCATCT AAGTCATGAA TAGCATGAAT GATTTGACCG AATTGAGAAT AACACATATG	3480
20	TGTATGGATT TGAGTTTCAT CACGAACTGA AGACGTTGCA AGTTTAAATG ATAAAACAGC	3540
	ATCTTTAAGA TATTGTTTCGT GATATTCAGA GCGTAATGGT AAGCCTTCAC GTAATGCAGG	3600
	TTCGTCAACT TGGATAACTT TGATTCCTGC AGCTTCAAGT GCTAATACTT CTTCGTTGAT	3660
25	TGCTAAAGCA ATTTGATCTT GAACGACTTT ACGTGGTAAA TCAACACGTT CAAATGACCA	3720
	GTTTAGAATT GTTACAGGTC CAGTTAACAT ACCTTTAACT GGTTTATCTG TTAAGCTTTG	3780
	TGCATAAACT GTTTCATCAA CAGTTAAAGG CGCTGTCCAT TTTACATCAC CATAAATGAT	3840
30	TGGTGGTTTT ACGGCACGTG AACCATATGA TTGCACCCAA CCGAATTTAG TTACTAAGAA	3900
	ACCTTGTAAT TTTTCTCCGA AGAATTCAAC CATGTCATTA CGTTCAAATT CACCGTGAAC	3960
	TAATACATCT AAGCCAATGT CTTCTTGAAT TTTAATCCAT CGAGCAATTT CATTTTTTAA	4020
35	GAATGTTTCA TATGCTTCGT CTGTAATGCG TTTGTTCTTC CAATCTGCAC GGTATTTTCG	4080
	AACTTCTCGG CTtTGTGGGA ATGATCCAAT AGTTGTTGTT GGTAAATCCG GTAAGTTCAA	4140
	ACGTTTTTGT TGTGTTTCAA TACGTTGCGC GAATGGTGAT TGTCTTGAAG TACGCACGCT	4200
40	TTCGAAATCA TAATCTAAGT TTTTGAATGA TTGATTTTGG AAACGCTCAT AACGTGCTTT	4260
	TAATTTATCA TATTTAACAC TATCGTTTTG ATTAAATAGG CGACGCAATG CATCTAATTC	4320
45	GTCTAATTTT TCAGTTGCAA AGCTTAAGCC TTCGCCAACA CTTGTATCTA ATGTTTCATC	4380
	ATCTAAAGAT ACTGGAACAT GTAATAATGA AGATGATGGT TGAATGACAA GTTCATTAGT	4440
	GTGTGCTAAC AATTTATCGA TTAAGACTTT TTTAGCTTCA ATGTCACTTG CCCATACATT	4500
50	ACGACCATCA ATAATTCCAG CGTATAATGT TTTTGATTTA TCAAAATCTC CAGCTTCAAT	4560
	TTGTTTAAGG TTATAGCCAT TATCATGGAC AAAGTCTAAA CCTATACCAC CAACAGGTAA	4620
55		

	AACACCAGCT TTTTCGAAAT AGTCATAAGC TTCACGTGTA ATATTTTCAT AGCTTTCGCT	4740
	GTCGCTGTGA ACTAAGATTG GCTCATCAAC TTGAATGTAC TCAGCACCTG CATCAATTAA	4800
5	TGATTCAAAC ACTTCTTTAT AAAGTGGTAA TAACGTTTAA ACTTTTCTT CAAAAGTTTG	4860
	GTGACCGCCT TTTGATAATT TAACAAAAGT AATCGGACCA ACAATGACAG GGTGAGCGTT	4920
	AACGTTTAAA GATTGGGCAT ATTTAAAGCG ATCTAATAAT ACATTGCGAC TCACTTTAGG	4980
10	CTCAACATTG TCCCATTGAG GTACGATGTA ATGATAGTTA GTGTTAAACC ATTTTATAAG	5040
	TGCACTTGCA ACATGGTCTT TATTACCGCG AGCAATATCA AATAATAAAT CATCATCAAT	5100
15	AGTTCTTCCT TGGAAACGTT CAGGGATGAT GTTGAATAAT AATGACGTAT CTAATATATG	5160
	GTCATATAAA GAGAAATCAC CAACTGGGAT GCTATCTAAG TGATAGTACT TTTGLAATAA	5220
	TAAATTTyCT TTATGTAGAT CAGTTAATGT TTGATCTAAT TCTTCTTTAG AAATCTTCTT	5280
20	TGCCCAATAA CTTTCGATGG CTTTTTCCA TTCTCTTTTT CTACCTAATC TTGGGAATCC	5340
	TAAGTTTGAT GTTTTAATTG TTGTCATAAT ATTGCCTCCT TGTGAGCAGT AATAGATTTT	5400
	GAGTATGCTG CAAGTTCTAA TGAATCTTCG ACATTTTGAA ACGGTGTGAT AATGTATAAA	5460
25	CCATTAAAAT ATTCAATGAC AGTATCGATT AAATCCTTTG AAAGCTTAAG ACTTAGTTCT	5520
	CGTGTTTTGG CTTTATCATC TTTAACTGCT TCAAATTGTT GTAAAATTC ATCTGACATC	5580
	TTGATTCTG GCACCTTCATT ATGCAAAAAG AGTGCCTTTT TGTAACTTGC GATAGGCATA	5640
30	ATGCCTATGA AAAATGGTTT GTTCAAGTGC TTAGTGGCAT GGTAAATTC AATGATTTTC	5700
	TCTTTGCTGT ACACGGGTTG TGTATAAAA TAAGACATTC CGCTTTCTAT CTTTTTCTCT	5760
35	AATCTTTTGA CGGCACCATA TAATTTACGA ACATTAGGGT TAAAGGCGCC AgcGATGTTG	5820
	AAGTGTGTAC GTTTCTTCAG CGCATCACCG TCAGTGTTAA TACCTTGATT AAATCTTAGA	5880
	GCGAGTTTCA TTAATCCTTT AGAATTAACA TCATAGACAT TGGTTGCACC TGGTAAGTGA	5940
40	CCAACTTTTG AAGGATCACC AGTTATGGCT AATATTCGT TAACGCCAAT GAGCGATAAT	6000
	CCAAGTAAAT GGGACTGCAA GCCGATTAAG TTTCGGTCTC GACATGTAAT ATGTACGAGT	6060
	GGTTCAATAT TGTAATATTG CTTAATTAAG CTAGCAGCAG CAATATTGCT AATTCTGACA	6120
45	GTTGCCAATG AATTATCTGC GAGTGTACC GCATCTACAT TAGCTTTATC AAGTTTAGCG	6180
	ATATTTTCAA AAAATCTATC CGTGTCTAAA TGTTTCGGTG TATCCAATTC GATAATAACG	6240
	GTTGGACGTT CTTGAACCTT AGATGTTAAT GATTGTCTAA CTTTATTTTG AGATGGATTG	6300
50	AAAAGTGCTT TCGTTGGTAT CGGAATCACT TTTTGTGTCAT TAACAGGTTT AAGTGTCTGA	6360
	ATAGATTCTT TAATAAATTT GATGTGCTCT GCGTGTGTAC CACAGCAACC ACCAATTAAA	6420

55

	TACTTAAATT	CACTATTTTC	AATATCTAAT	AAGCTGGCAT	TTGGATAACA	AGATAAGAAT	6540
	GCGTGCTCTG	GTAATTCAAT	ATGTGTGAAA	GACTCTTGCA	TATGGTGCCG	GCCATGATGA	6600
5	CAATTGAGTC	CCACGATGTT	TGCACCACAT	TGAACGAGTT	GTTTTAATCC	TTCATTGATT	6660
	GCCTGACCAT	TAACTAAGTA	ATTTGTGTTT	GAAGCGGTTA	ATTGAGCAAT	GATTGGAATG	6720
	TCGTATTTCT	TTCTCGTTCT	TGAAATGACA	TTTGTTAACT	CTTCTAGGTC	GTAATACGTT	6780
10	TCGAAAAGTA	GCGCGTCAAC	GCCTTCTTCA	ATTAAGGTGT	CTATTTGAAT	TTCAGTATGA	6840
	TAAAGAATAG	TTTGTAAGCT	GATATCCTCT	TGTTTGATAC	CTCTAAACCC	ACCAACTGTG	6900
	CCTAATATAT	ACGTATCTTT	ATTTGCTGCT	TTTTTTCGCA	TGCGAACGGC	GGCTTGATGT	6960
15	ATTGCTTTAA	CTTTATCTTC	AAGACCGAAT	CGTTTTAACT	TTTCAAAAT	TGCACCATAA	7020
	GTATTGGTTT	GAATGACATC	AGCACCGGCT	TCAATATATG	AACGATGGAT	GCGTTCAACT	7080
20	TTATCTGGAT	GGCTAAGATT	ATATGCTTCT	GGACAGGTGT	CTAATCCTTC	AGAGTATAAA	7140
	ATGGTTCCTA	TAGCGCCATC	AGCTACTAAA	ACATTATCTT	TCAATTGTGT	GAGGAATTGA	7200
	CTCATTGAAT	GCCTCCTTTA	ATGCGTATTT	GATGTCTGCA	ATGAGTTCAT	CAGGATCTTC	7260
25	GAGACCAACA	CTTAATCGGA	ATAGACCGAA	AGTGATACCA	CGTTCTTGTC	TCACTTCTTC	7320
	AGGTAGTGCA	GCGTGAGACA	TTGTTGCTGG	ATGTGAAAGG	ATCGTTTCAA	CACCGCCCAG	7380
	ACTCACTGAA	ACGAGTGGTA	ATGTCAGTGC	ATCGACAAAT	TGTTGTGCTT	TAGACTCATC	7440
30	AGCTAAACGA	AAGCCAATAA	CGGCACCGCC	ATTTTTAGCT	TGTTCTAAAT	GAGCAGTAGT	7500
	GAGTCCCGGA	TAATAAACTT	CTGAAATTTT	ATCTTGCTTT	ATTAAAAATG	ACACGATTTT	7560
	TTGAGCGTTT	TCGACAGATT	GTTTAAATCT	GATTGGAATA	GTTTTTAAAT	GTTTAGCAAG	7620
35	TGTCCAGCTA	TCCTGAGCAG	ATAACATATT	GCCTGTACCA	TTTTGTATTA	AATAAAGAGC	7680
	GTCCTAATT	GCCTCATTAT	TAGTTATGAC	AGCACCAGCA	ATTAAATCGC	TATGTCCACT	7740
	TAAAAATTTT	GTAGCACTAT	GAATGACAAT	ATCAGCGCCA	AGTAATAAAG	GTGATTGACC	7800
40	TAACGGTGTC	ATAAATGTAT	TGTCCACAGC	TACCAGTAGT	TCATGCTTTT	CGGCTATTTT	7860
	AGAAACAGCT	TTGATATCAG	TAATTTTAAA	ACAGGGATTG	GATGGTGTTT	CGATATAAAT	7920
45	TAATTTTGTG	TTTGATTGAA	TGGCACCCCT	GATTTGTTTC	AGCTTTGTAG	TATCTACGGT	7980
	TGTAAATTCA	ATATTAAATC	GATTCAAAAT	TTGCTCAGTG	AGGCGAAAAG	TACCGCCATA	8040
	TACATCATCG	GGTAAGATGA	CATGATCACC	AGATTTGAAA	GTCAAAAGTA	CTGCTGAAAT	8100
50	AGCAGCAATA	CCTGATGCAA	AAGCAAAAGC	GAATTTTCCC	TGTTCTAATC	GTGCTAACTT	8160
	CTCTTCTAAA	AGTTCACGGT	TAGGGTTGCC	CTTCGTGCAT	AATCATATTT	AACATCGCCA	8220
55							

	TCCACACCTC	TACGCCAATC	GAATATCACT	TCTGTCTCTT	TTGAAAGTGT	CATACAATCT	8340
5	CTCCAATCTG	AGCTTTATCT	AATGCTTGGA	TGATATCGCG	TTCGATGTCT	TCATAATTTT	8400
	CAACACCTAG	TGATAAGCGG	ATTAAATACT	CATCAATGCC	ACGTTTATCT	TTTTCAGCAT	8460
	CTGGCATATC	AACATGTGTT	TGGGTGTAAG	GGAAGGTCAC	TAATGTTTCA	GTACCTCCTA	8520
10	AACTTTCTGC	AAAAATGCAA	ATGTCTAAAT	TTTCTAATAA	TTTAGCGACG	CTATAGGCCT	8580
	TGTTAAGTCT	TAAACTAAGC	ATGCCAGTTT	GCCCGCTATA	TAGTACTTCG	TCAATTGCCT	8640
	GAAGTGACTG	ACATTTTTTA	GCAAGTTTTT	TAGCGTTTGA	TTGCGCACGC	TCAATGCGTA	8700
15	AATGCAAAGT	TTTAAGTCCA	CGTAACAACA	AATAACTATC	TATTGGTGAA	AGTGTTCGCG	8760
	CAGTCATGTT	GTGAAATCA	AACAACGTGT	GCGCGAGTGA	TTCATCTTTG	ACGGTTACGA	8820
	CACCTGCTAG	TACATCGTTA	TGTCCGCCAA	TATATTTTCGT	GGCTGAATGT	AAGACTATAT	8880
20	CAGCACCTTC	TGCTAGTGGT	GTTGAAAGAT	AAGGTGTTAA	AAAAGTATTG	TCGATAATTG	8940
	ACAATAAGCC	TTTAGCTTTA	CAAAGTTGAT	AGTATGGCTT	TACATCAATA	GCAATCATTT	9000
	GTGGGTTAGA	TATTGGTTCA	ATGAATAATG	CAACTGTTTT	ATCAGTGATT	TCTTTTTCAA	9060
25	CTTGTTTATA	ATCTGTAAAA	TCAACGTACT	TAAATTTGAT	ATCGTATTGT	TGCTCGTAAA	9120
	ATTCAAATAA	TCTAAATGTG	CCACCATATA	AATCGAATGA	AACTAAAATT	TCATCATGAG	9180
	GTTTAAATAG	ATTACATATT	AATTGAATGG	CTGACATTCC	ACTTGATGTA	GCGAATGATG	9240
30	CAATACCATG	CTCAAGTTTG	GCAAAACAGG	TTTCAAATGT	TGAGCGTGTA	GGATTTTTAG	9300
	TACGTGTATA	ATCAAAACCT	GTCGATTGTC	CTAGTTTTGG	ATGCTTGTAG	GCAGTAGATA	9360
	AATGGATTGG	ATTCGCTATA	GCACCGGTTG	AATCATCGGT	TAATGTGATT	TGGGCTAACT	9420
35	GTGTATCCTT	CATATTAAGA	CCCTCCTATA	AGAAAAAATA	AAAAAAGCTT	CCGTCCTTCG	9480
	TACCCGAATG	AATCGGATAA	AAAGGACGAA	AGCTTATGTT	TCGCGGTACC	ACCTTTATTT	9540
40	GTTATTCCAT	CGCTGAAATA	ACCTTATTCA	GTACGCATTA	AAAGTAAATA	TGCTTACTGA	9600
	ACAATTATCA	CAATTAAAGT	CAGTAAGTAA	GGATATAGTA	ATGTGCTATC	CCATACTTAT	9660
	TAACAAAAAA	TCGTGCGTAA	AGAATCCAGT	ACGCCATTTA	ACATCAATGT	TAATACTGTA	9720
45	TCGCTATAAC	GGGCGAACCC	GTAGACACCT	CATATTGGCA	TCAACACTCC	AAGGCCATTT	9780
	TCAAACACGC	TTTCAAAATC	TTCTCTCAGC	TACTAAAGAC	TCTCTGTATA	AGCAGGGTGT	9840
	GTTTTACTTy	CCTCTTTATT	GTGTTTACGT	TTCATTAAAC	TGTTATAAGA	TATTAATTAG	9900
50	CTTACAGAGT	AAAAAAAGAT	TTGTCAACAA	TTATTCAGAA	AATTTTGATT	TAAAAGTTAA	9960
	TTTGTTTG TG	AAATTGTAAT	TGGTATCTTG	AAGTTGAAAA	ATGAATTATT	TTTTAAATAA	10020

55

TCAAATAAAA AGTGATGTGA GTGAATTGTC AAAAAGTGAA GATCAACGTA TTAATAAAAC 10140
 AAAAGATGAA CAAATTAAGC AAATAGATAT ATCGGATATC AAACCGAATC CGTATCAGCC 10200
 5 CCGAAAAACT TTCGATGAAA ATCATTATAA TGATTTGGCA GATTCAATTA AGCAATATGG 10260
 AATTTTGCAA CCAATTGTGC TTAGAAAAAC AGTTCAAGGT TATTACATTG TAGTTGGTGA 10320
 AAGAAGGTTT AGAGCTTCGA AAATTGCTGG TCTAAAATAC GTATCAGCGA TTATCAAAGA 10380
 10 TTTAACAGAT GAAGATATGA TGGAACTGGC GGTCAATCGA AATTTACAAC GAGAAGACTT 10440
 AAATGCGATT GAAGAAGCTG AAAGTTATCA ACGTTTGATG ACAGATTGTA AAATTACACA 10500
 ACAAGAAGTA GCGAAACGAT TGAGTAAGTC GCGCCCGTAT ATAGCGAATA TGTGAGGTT 10560
 15 ATTACATTG CCGAAAAAGA TTGCTGACAT GGTAAGAT GGGCGACTGA CAAGTGCACA 10620
 TGGACGAACG TTATTGGCAA TTAAGATGA ACAACAAATG CTTAGGTTAG CGAAACGGGT 10680
 TGTTAAAGAA AAGTGGAGTG TCAGATATTT AGAAAACCAT GTTAATGAAT TAAAAAATGT 10740
 20 TTCGTCAAAG TCGAAACAG ACAAGTAGA TATACTAAG CCTAAATTTA TAAAGCAGCA 10800
 AGAACGACAG TTGCGAGAAC AGTATGGTAC CAAAGTAGAT ATATCAATAA AAAAATCGGT 10860
 25 TGGTAAATC TCATTGAGT TTGATTACAA AGAAGATTTT GTGAGAATAA TTGAACAATT 10920
 AAATCGTAGG TATGGTAAAT AGTTACACAA TTTTATATAA TAACTCTTTG TGCAAGTGTA 10980
 AATAAATTGT AATCAGTGAC ATTTGATTCT AGAT 11014

30 (2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6022 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

TCCCCTTATG GAATTCACAA TTCTAGTTTA CATAATATAT ATTATAGGAA GTTATATGTG 60
 TGTAACGCAA AAGGTACCCT ACATCATAAT CATTATCTAA TATCGTCACA TAACTTACTT 120
 45 ATGCTATAAT CATGGTATTA TATTGTTTGG AGTGATTGTA TGAGATTTGT CTTTGATATT 180
 GATGGTACGC TTTGTTTTGA CGGCCGATTA ATTGACCAGA CTATTATTGA TACATTGTTA 240
 CAATTACAAC ATGATGGTCA TGAAGTTATA TTTGCATCAG CACGTCCGAT TCGTGATTG 300
 50 TTGCCAGTTT TACCATCAGT ATTTATCAG CACACATTAA TTGGCGCAAA TGGTGCTATG 360
 ATTTACACAG AATCAAAGAT TTCTGTTATC AAACCAATTC ATACTGATAC ATATCATCAT 420

55

	GCTGCACAAC	TTGACGCTGn	AGAACGCGAT	TTTTGAGCGT	TTAGATCCAC	ATAAGCTGGC	540
	CAGTTGTATT	GATGTTGCAA	ATATCGACAC	GCCAATCAAG	AKTATTTTAT	TAAATATAGA	600
5	CCCGGCACAA	ATTACAACATA	TATTAGACGA	GCTAGATAAA	TACCATCAAG	AATTGGAAAT	660
	GATTCACCAT	TCAAATGAGT	ATAACATTGA	TATAACAGCG	CAAAATATTA	ACAAATATAC	720
	TGCATTACAA	TATATATTTG	ATGCAGATGT	TAAATATATA	GCATTTGGTA	ATGACCACAA	780
10	TGATATTGTC	ATGTTACAAC	ATGCTAGTAG	TGGCTATATT	ATAGGACCAT	CAGAAGCATA	840
	CACACACGCA	ATATTGAAAC	TTGATAAAAT	CAAACACATC	AATAATAATG	CACAAGCTAT	900
	TTGCAAAGTC	TTAAAATCAT	ATAAATAAAA	ACACCCCTAT	CAAATGATAA	TCATTATCAA	960
15	TCGATAGGGG	CTATTTTAAT	AAAATTCGTC	CTCGAACATT	TCTTCCTCTT	CATCTAATCC	1020
	AAATAATTCT	GCCATTTCTC	CATGTTCAAT	TAACATGTTT	AAATATGCAT	CGCGGAGTTC	1080
20	TTCTTCACTC	ATATCATTAA	TCATTTCTTT	AAGACTATCA	ATCCACATAT	TTCTGCGTAA	1140
	TTGATAGTCT	TCTTCAACTT	CGTTTAACAT	CATTATATGT	TTATTTGCTG	CTTCTGGACT	1200
	AGCTGTAAAG	AGTAATGCAA	TCATATGTTT	ACATATCACT	CGTCTTCCAT	CAGCATGAGG	1260
25	ACAATTACAT	ATGGATTTTC	TAGGATGTTT	CATATCAATA	TAACAACGAT	ATACTTTGTT	1320
	GCCACTGCCC	TTTACTTCAG	CCTCATGCTG	CGTTTCTGAA	AATGATTTTA	AGTTAATGAC	1380
	GCATTCACCT	TGATAATAAT	TAAAGCCTCT	TTCTATAGAA	CGAATACTTG	CAATATCAAG	1440
30	TAATCCCATT	AATGaTACTC	CTTTTTTATTA	TTATTTTTTAA	ATAAAGAAaAA	TAAAATAGAT	1500
	AAGTGTCTAG	ATTAAAATAC	TTGATTTATC	TATATTTTAT	AACAAGTCTA	GAATTATCGC	1560
	ATTCTTAAAT	AACTAATATG	AAAATGcTTG	CACTAATTCT	TTTGTATAAG	GGTGTCTATC	1620
35	AACATTAAAT	AATTCCctCTA	TTGCAAAATC	ATCGACTATC	ATGCCATCCT	TAAGAACGAT	1680
	AATTCTATTA	ACTAAGCGTT	GTAACACGGA	TAAATCATGA	GAAATAACGA	TAAAATGATT	1740
	TAAGTTCGTA	ATCGTTTGCG	CTTTTAATAT	ATTGATTACA	TTTTGTTCAG	CTATAACATC	1800
40	TAAATTTGAA	GTTATCTCAT	CACATATTAA	AACGCGAGGC	TGTGCTAATA	ACGAACGCAT	1860
	GACATTAAAT	CTTTGTAAIT	GTCCGCCACT	CACCTCGCTT	GGTAATTTAG	TCAATAATTG	1920
	CGCGTTTAAAC	TCAAAAGTAG	ATAAATGTTG	TAATAATAAT	TGATCCTGAG	CAGTATTATC	1980
45	AGTTAGACCT	CTGTAATAAT	ATAACGCTTC	TTTTAATGAG	GTCTCAATCG	TCCAATCAGG	2040
	GTTAAAGCTA	GTTAAAGGGT	GTTGGAAAAT	CGGTAACACA	GCATTGTCAC	TTAAGTAAAT	2100
50	CTCTCCTTTA	ACAGGTTTAA	ACAAGCCAAG	AACCAATGAA	GCGAGCGTAC	TTTTACCACA	2160
	GCCACTTTTCG	CCTAAAATAC	CAACATTTTC	TCCATCAGGT	ATAGTAATAT	TGATATCTTG	2220

55

	CCCTCTTTAA TTGTGTTCTA TATTTAATTA GACGTTCACT ATACGGATGC AAATGCTCAT	2340
	ACTTGAAATG ATTAATATTA CCTCGTTCAA TGATTTGACC TTCTTTTAAA ACATAAATGT	2400
5	ACTGACAATA TTTCAATACA TGACTTAAGT TATGTGTGAT AATAAATAAT GTTTGACCAT	2460
	GTTCTAATAC AATATGCTGT AATAAATCCA TCACCTGATT ACCGTTCAAA GCATCCAATG	2520
	ATGCAACTGG TTCGTCTGCA ATGATTAATT TAGGCTCCAA CATGAGAACG CTTGCTATGT	2580
10	ATACGCGTTC AAGTTGGCCC CCAGAAAGTT GGAAACTATA TTTATTTAAT ATATCTTTGC	2640
	TTTGTAATTT AACCCACGAC AAAGCCTTAT CAACTTTGGA CAAAGCCTCT TCTTTACTAC	2700
	CTTTATAATG CTTACGATAA ATCGCAGTTA ACTGTTTACC TAATTTAGTA TGGTCGTTAA	2760
15	AACCTTTCTGC ATAATTTTGA GAAATATAGC CAATTGTATG ACCATAATAT TGACTIONC	2820
	TACTAACATT TTCCCCATCA AATTGGTACG AATCATACGT GCAGCTTAAA TCAAATGGTA	2880
	AATATTCAAG TAAAGCTTTA GCAATCAAAC TTTTCCAGC GCCGCTCTCT CCAATCAAGG	2940
20	CATTAATCTG TTGACTAAAA ATTTTCAAAT CAATCCCTTT AATAAGAGAT TTCTCACTAG	3000
	TATTCCTTTAT TGTAAATTT TGTATATCAA TGAGACTCAT CATATTCACC CCGTTGTTTC	3060
25	AGCAATCTAT CTCTTAGTGC ATCACCAGTT AAATTAAAAA TTAAATAGT TATAGCAATG	3120
	ACTGAAGCAG GTGCAATCAA CATAATTGGA TGAGACGAAA TAAATCACG ACCTTGTTGC	3180
	AACATAGCGC CCCaCTCTGG TGTGGCGGT TGTGCACCTA ACCCAATAAA TGATAGTGAA	3240
30	CTTATATATA GAATGATTTT ACCGAAATCA ACGACCATCA AAACGATAAT AGCCGGTATA	3300
	ATTTTAGGTG TTAAATGACG TATTAATATT GTTCTTGTG GTACATGAAA TAATTGTGCC	3360
	ATTTTTATAT AAGGCTTATT CATTTGCTA TTAATATAC TTCTAGTCAA CCTTGTGTAA	3420
35	TTCATCCATT TTATTAATGT AATTGAGATA ACTAAATCC ATAAAGATGG TTGAAAAAA	3480
	CTTGCTAAAG CAATCATGAT GATAAATTCT GGAATACTTA GACCAACATC AATAAACCTT	3540
	AACACTAATC GTTCAATCCA CCCTTTTTTG TATCCGGCAA ATAGACCTAG TGTAACACCT	3600
40	ATGACAACGA TAGCTATTAA TGTAAAACA GTAACAAACA ATGTTGAACG TGCACCGATA	3660
	ATAATTCGGG TAAATAAATC TCTCCATAA TCATCAGTTC CTAATAAATG CAACCAACTA	3720
	ATAGGTTCAA AAGTTTGTGA TAAATTGACT TTGGTTGCAT TTTCACTACT GACAAAGAAT	3780
45	TGCAGTACAA TTACCACAAA AATAAATGCA ACGAATACAA AAAATATCAG GTTATTCTTT	3840
	GAAAATATTT TATGCATGAC GGTCACACT TTCTGATATC AATGGTGTAT TGGTTTTGAT	3900
	TTTTGGATTT CCTAATTGTA AACGCTGCTT CGGATCAAGT AATAACGTTA ATAAATCAGC	3960
50	AATCGTATTG ATAATAACAA CGAAGAAGCC AATAAATAAC ACGCATCCTT GAATAACAGG	4020

55

	ATTTTCAATC ACTACAGTAC CACCTATTAG ACTGCCAAGT GAAATCCCTA GTAATGGGAT	4140
	AATCGGCAAA ATTGTGGTT TTAGTAAATC ATGAATTAAA ATATAACGTT CATTCATACC	4200
5	GCGTAATCTT GATGCTTGTA CGATATTACT TTGCAATAAC ATCAATAAAT TAGAACGCAC	4260
	TAAACGAATG ATGTATGCAC ACATACCTAA AGATAGCGTG ATTACAGGTA ATATAAACTG	4320
	ACTTAGTATA ACGCTATCTA TATTCATTAA ATTTGTGACA ATAAATAATA AAATAATACC	4380
10	GATAAAGAAC GCTGGTAAAC TAATCGATAG TGTTGAGATC ACTCTAATCA CTTTATCCGT	4440
	CCACTTATGA AATCGTTTGG CTGCTATAAT GCCGAGCGGT ATAGATATGC ATAACGACAC	4500
	TACTAATGTT GAAATGATA TGAGTAATGT TATGGGTGCA TAGTTGAATA ATATCTGTGT	4560
15	TACCGGTTCT TTTGATTCAA AACTTTTTCC TAAATTAAAA TGTAATAAAT GATTCATCCA	4620
	ATGCCACCAC TGTACCAATA AAGAATCATT TAATCCCAAT TTATCTTTGG TTGCATTTAT	4680
20	TTGTTCCGTC GACACTTGTG CTACATCAAG ATGTAATATT TTATCAACAG GATTGCCTGG	4740
	TGATAATTTC ATTAATAATGA ATGTAAGTGT AGAAATAACA AATAAAACAA CTATCATTTG	4800
	CATCAGTCTA TACAACATAG ACTTTATTAT GAACATAATA GTCCCCCTCC TTGTGTAAGT	4860
25	TACTAACACT TTCTTTTAC ATGAGAATGG CGCATGTATA TGCAACTTAC ATATTAAGAA	4920
	CTAACGTTCA TTATAGTATT ATCCATAAAG AAATTGAAGT ATATTTAATT TTTTAACAAA	4980
	ATCATTATAA AATATAATAT TTTGAATCAA GTCAACCATG TAAATATAA AAAAGTCAAA	5040
30	ACAAAAACAA CTATAGCACT GTATTCCATC TCTTTGAAA TAATTGTTAC TGCAGTGTA	5100
	CTTAAAAGTC GATGATTTTG TGCATATAGT TGTCGAATAT TATTTTTAT CTTACGGCG	5160
	AAGTTCAGCG CCTCATAGC CGTATTTTC AATTGCTTT TCTAATTAC GCGCTTTTCT	5220
35	TTCTTTACGC CAATTTCTAG TAAATACCA TAATAGAAAA CTAATTAATA AACTCATAAT	5280
	CGCTAAAAAT GCAGCGTATC CTAATAATGG TTGATATTTT ATATCTTGAA AATTGGAAT	5340
	AAAAATGCA AGCACACCTA ATATAACAA TGTAATTACT GCAGATACAA ACCATTTATT	5400
40	TAAACTAAG CAACAGAATA TTGTTAATAA AATCATTATT AATGTTGTGA TCCATAAATA	5460
	ATTAGGCATA TCGAATAATG TCATATTCAT TCTCCTTTTA TTTCATTACT TTCCTTGTAT	5520
	ACATTTTATT ATAAATTTTT AAAAATTAA ACAATAGCAG TCAGTTTCAA GCAATATTCT	5580
45	ATCTACTAAT AGAAAAATCA TTGTTCCCTG CGACATGGAA ATCGTAACAT TATCGTTTAG	5640
	GAGACAAAAT TATGTATAAT GAATGTATTA TACCAAAGGA GTGATTATAT GTCTCAAGGT	5700
50	TTACCTTTAA GAGAAGATGT TCCTGTTTCA GAAACATGGG ATTTAGTAGA CTTATTTAAA	5760
	GATGATCAAC AATATTATGA AAGTATTGAC GCTCTAGTAC AACAAGCAAA TCAATTTTAT	5820

55

GAAAATATTT TAATTGCCTT AGATCGCTTA AGTAATTATG CAGAACTACG TTTAAGTGTA 5940
 GATACTAGTA ATATCGAGGC ACAAGTATTG AGCGCTAAAT TATCTACTAC ATACGGTAAA 6000
 5 ATTGTTAAGC CAATTATCCT TT 6022

(2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 476 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

CCATCAATAA TGTATACATG ATTGGCATCA TATTCCCCTT TAATTAGAGA GCTACGTACA 60
 GTTTGTyTTA TTAAAGTAGA ACTAATAAAT AACCATCTCT TATGTGCACA AACACTTCCC 120
 GCAACAATTG ATTCACTTTT ACCAACCCGT GGCATACCTC TAATGCCAAT CAACTTATGA 180
 CCTTCTTCTT TGAACAATTC AGCTAAAAAG TCTACTAACA AGCCTAAATC TTCACGCTCA 240
 AATCGAAAGG TTTTCTTATC TTTTGCATCT TGCTCAATAT ATCTTCCATG TCTTACTGCA 300
 AGACGGTCTC TTAATTCTGG TTTTTTAAGC TTTGTTATTT CAATTTCAAT TATACCACGA 360
 GCTATTTGCT CAAAACGTTT AACTTTTTTCA AGATTGTCTG TTTTAATTAA AAGGCCTCGT 420
 30 TTACCTTGAT CAACACCATT AATTGTAACA ATACTTATAC CTAACATAAC TAATAA 476

(2) INFORMATION FOR SEQ ID NO: 94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3633 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

AGAAATACAA CGAAGCATAT AAATATAACC GATCTTTTTT CTAATTGAAT ATTAAGTAAG 60
 TGTATGTACT TTCTGGAAGT AGCACCTAGT TGGATTGTCT CTCCTACAAC AGGCCAAAAA 120
 45 TTTTATTTTT TAACTGGCTT AACAGTGTTC AGTTTTTCAT ACTCTTCTCT ACTAATTTTG 180
 GCGCACCTTT TTGGAATGAA CCAATTAATA AATGGAAAAA AGTATACAAG CCAAGTTCTT 240
 ATTACATCGA CCATTAAATA CTCATCATCA TACTTAATAA CTCTGTATTT CGGATTTTTA 300
 50 TTGATAATTT CGGTTTCACA AAGCAATAAT TATCACTTCC TATTAATAAC AAATTCACAC 360

	TTATATGACC TTAAATATAT AACATGAATC TTTTGTCTA TTATTGAAGA CATATTTATA	480
	AAGAAAAATA GCATTGTCAT AATAACCCAA GCAATAAATA CTATAATATT TTGGATAGAT	540
5	AAACTAATCA TTACATCTAA GAACATGATT GATAATCCAC CACAGAAAAA ATAAGAAAAT	600
	AGTACAAAGC AAAGATTCTT GAATGATGGA AAAATCATAA TTTTCCATT GCTACTCCGA	660
10	TCATTATAGA TAGATAACTT TACTTCTGA TTTAAATATA TATAAACAC TAGAATACTT	720
	AATAATAAAA CCGAACAAAT GATAATAACG CAATTTTTTT CTAAATGAGA ATCAGGTATA	780
	TATATTTTAT CTCTAAACAT AGTGCCAAAT AAAAGTATGC TACCTATAGC TGGCCATAAA	840
15	GCTTTaTTTT TAACTGGTTT GACAATATTT AAATTATCAA AATCTTCTCT GCTGATTTGG	900
	ACATATTTTT TTGGTATTAA CCAATTAATA AACGGAAAGA ACAAACATAA CCAGGTGCTT	960
	ACTAAATCAA TCATCAGATA GTCGTTTTTA TATTTAATAA TTCTATATCT GGGATTTTTG	1020
20	TTTACAACCTC TAACCTCGCA AAGCAATATC TCCACTCCG TCTCGTTGGT TTTATATCTA	1080
	ATACACTTTC AGATACTTTA TAAGTGTTTT GTATTTTAGT AACATACTAT TTTCTGTTT	1140
	ATTACTTAAC TTACGAACTA CAATCTAAGT TTAGTAATTT CTATTGCTTT TTAAGTTTGG	1200
25	CATAAACCTT TTTATTACTA ATTGAGCCCA TGCTTATTAG AAAGAAAAAA ATTGTAATAA	1260
	TAATCCACAT AATAAATACC AGTAGATTTT GAGGTTTTAT AGTCATTAGC CATATTAAAA	1320
30	ATAATATAGA ACAACCTCCT AATAATAGAT ATGTGAAAAC TATAAACTT CCATCTTTAA	1380
	AAGTAGGCAC TAATATAACC CTATTTTCAT TATCTAGATT ATCATCATAT ATCTTTAGTT	1440
	TAAGCTTTTT ATTTAAGTAA ATGTAAAATG CTGCAATACC TATAAATCCT ATAAAAATA	1500
35	AAGATATTAA AATCTTATTA TCTAATTGAA CTTCAAACGT ATGTACATAT TTCCGTAAAA	1560
	TAACTACAAA TAAAAACGAA CTACCAGTAA CTGGCCAGAA AATATTATTT TTATTTTGTT	1620
	TATCAACATT TAAATTTTCA AGTTCCTTCT CACTAAGTTT TGCATACCTT TTGGGAATGA	1680
40	ACCAATTAAT AAAAGGAAAA AAGTATACAA GCCAAGTGCT TACTAAATCA ATTAACAAAT	1740
	ACTCATCATT ATATTGAACG ACTTTATATC TCGGATTTTT ATTAATAACC TTAATATTAA	1800
45	AAAGCAAAAC TCACCACGCC CATTTTCATTG GATTTATATG ATTGCTAATA ATATTTTATG	1860
	CTTCACTAAC AGCATTCCCA AACTATCCA TGGATTTTTT TGTAGTTTTT TTAACAACAT	1920
	CTATACTATT ATCGATTTTA TGCCCTACCC AGTCTACTTT ATCTTTTAAT CAAAAATAT	1980
50	TATTTTGATA AATTAAATCT GTTCCTAATG CAAATACTGT ACTCATAGCC AAACCTGCTA	2040
	AAATCACCCA TCCTACTGGA TTACTTCCTA AAACAAAAGT CGCTAATCCA GCTCCAACTG	2100
55	CTGTCCCTGC AGATCCAGCT GCAAGCGTgC ATACCATTAT GCGACAACGC CTCTCCAAAT	2160

CCTTTACCTA GGTATTTTCC GCCTTTTGCA AATTTACTAC CATTTTCTAT AAACACATTA 2280
 CCTGATGTAC GTTTGACTTC CACAAATGAA TTTGGACCTG CTGGGCCTTT CACTCCACCT 2340
 5 GCTGTATTGa TAAATACACC GAATTTACTT GcATTTATAC CGTCTTGCTC TAAAAGTGTT 2400
 GACGTAATAT CTAATCCTAT ATCTCTTTTA ATACTGTCTT TATTGTCATT TATATATTTT 2460
 AATATACTTT TCGGGATATC GTCCTCTGGA TGTCTTTTGG CATATGCCTT TATAACAGCA 2520
 10 AAGTCTGCTT TATTTAAAGT TTCTTCTCT GCTTTATGTT CAATTTTCCC CATAGCAACT 2580
 TTCAAATATT TTTCATGACT TGCTTTGGCC CAATCAAGTT CTTTACCTGA AGGAATATTA 2640
 15 AATTGATTTG TTGAAAAGTT CCAAAAATTC TGCGCTTGGG TAAGTCCTTG TTGGACAATT 2700
 TTTTGAAATT CTTCAACTTC TTAAATATT TCTGGTGATT TTTGATTAAA CTCACGCAAT 2760
 TTGCGTAGCT TCTCTTCTAA TTCATGTTTT TGTGACCTA ATGTTCTGAT TATTTGTTGG 2820
 20 TTCGATGAAA TGGCTTGCTG ATTATCGGAA GCATGCTTTT TCAAATTGTT ATTCAAATTT 2880
 TCATATCGCG TAATTTGTTG ACTTAATGAT CTGATATCTT CTTCAAGCTC TGATTCTTTT 2940
 AAAGATATGC TATCAACCTC ACTCGTATAA CGTGACACAA AATTaTCGCA AGCTTGCTTC 3000
 25 GTTAAATCAC TCAATGTTTT CATACTTGTT GATAATGGAA TTAACACCGT ACTAAAAAAT 3060
 TGCTTAGCTG ACGTATACGC TTTCCCTTTA AGCGCATCAT CATTAAATAA TTGAGTAATT 3120
 GCTTTTTTCCA ACGCATCATA ATTTGAATTC ATTGTTTGAC TCAAATTCCC CACACTTGAA 3180
 30 GCTTGGTTTT GAGATCTGTC TAAATACATG TCAATACTCA TCGGCATGCT CCTTTTTCAA 3240
 AAATATATGA TTTTCAAAC ATTTAAAATC AAATGCTTTT TACATCTACA AAGTTGTAAA 3300
 35 ATTTTAAAAC TCGGCGATGA TTATTTCTTA TGTAAAGGAG TCTAGATGCA GGTAATTTGA 3360
 GATAACATGT CGCCTTTTTT CTTATTTTAG CATATGGATA TAATGGTGTC TTTGTATATT 3420
 CGCAATTAAAT CAATAAAAAT TATCTTTCAA TATTTTAATT TTATTGCGAC AACATCCTTA 3480
 40 ACATTAAATA TATTAATATC TCAAAATATA TTCACTATTA AAATATGTCA TCAGTTGTTA 3540
 AAAGTATTTT CTCATCATGC GAAATATCAA AACGTATCTA AAATACGAAT AAGTTTATAC 3600
 AATCACACAA CATCATCATT CAAAATTTTA TTG 3633

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2365 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

	TGATACGAAt GCATTACAAT TCATATGCAA CATACAATTC CTTCTACAGC AAATGAAGTG	60
	AAACAAATAG TTGATGTGAC ATCTGTAGCA GAAAATGATA CGCATTAGTC ATAAAAATTAA	120
5	ATGGAAATGT CGATGAAGTG TATCAGCAAT TACAGCGATT AATTAAGAAT GCTAATGTCTG	180
	AAGAGAGTGA GAATACTGAC AATATTAATA GTCAAGATAC AAGTTATACA CCTCAAGTAA	240
	AAGTAACAAC ACCAATTTTA GTGAAAGCAC CAATCGCTGG TCGTCGTATT TTACTTAAAG	300
10	AAGTAAGAGA TTCAATTTTT AGAGAGAAAA TGGTAGGTGA AGGCTTAGCA ATCAAAGCTC	360
	ATGAAGAATC CAAAGTAATC GCACCGTTCA ATGGTTTAAT ATCTATGATT GTACCAACTA	420
	AGCATGCAGT TGGTATTCAA TCAGAAGACG GTGTGGACAT AGTCATTCAT ATTGGCGTGA	480
15	ATACAGTTGA CTTGGAAGGT AAAGGGTTCA AGTGCTTTGT AAAGCAAAAT GATCATGTTG	540
	AAGCAGGGCA AACGTTGTTG CAATTGACC AGCAATATAT ACAACAACAA GGCTACAATG	600
20	CTGACGTTAT TGTGCTTATT AGCAACTCTG CCGATTTAGG AAAAGTAGAA CTGACAATGA	660
	ATGAAATCAT TACGACTGAA GATGTTATTT TAAAAATATT TAAAAACTAG GAGTGTGTTG	720
	TAATAATGAC AAAATTACCG CAAAATTTCA TGTGGGGTGG CGCTCTTGCC GCAAATCAAT	780
25	TTGAAGGTGG ATATGATAAA GGTGGTAAAG GGTAAAGTGT AATTGATGTT ATGACGAGTG	840
	GTGCACATGG CAAAGCACGT CAGATTACAG AATCTATAGA TCCCAATCAC TATTATCCAA	900
	ATCATGAAGG TATTGATTTT TATCATCGTT ATAAGGAAGA TATTGCCTTG TTTAAAGAAA	960
30	TGGGATTGAA ATGTTTACGT ACGTCGATTG CGTGGACACG TATCTTTCCG AATGGGGATG	1020
	AAGATGTGCC AAATGAAGAA GGAATCGCCT TTTATGATCG TATCTTTGAT GAATTAATTG	1080
	CACAAGGTAT TGAACCTGTT GTGACGTTAT CACATTTTGA GATGCCACTT CATTTAGCGA	1140
35	AACATTATGG TGGATTTAGA AATAGAGAAG TTGTCGATTA TTTTGTGCAT TTTGCGCGTG	1200
	TTGTATTTGA AAGATATAAA GATAAAGTTA CATATTGGAT GACGTTTAAT GAAATTAATA	1260
40	ATCAGATGGA CACATCAAAT CCTATCTTTT TATGGACGAA TTCTGGGGTA GCATTGACAG	1320
	AAAATGATAA TCCTGAAGAA GTCyTGTATC AAGTAGCACA TCATGAACTT TTAGCCAGTG	1380
	CyTTAGCAGT TCGTCTTGGT AAAGaGATtA ATCCgAaGTT TAAGATTGGr ACmATGATTt	1440
45	CAMaTGTACC CmTTTTATCCa TAwTCGTGTC ATCCGAAAGA TATGATGGAA GCACAAATTG	1500
	CGAATCGCTT ACGTTTCTTT TTCCCGGATG TCCAAGTGAG AGGTTATTAT CCAAGCTATG	1560
	CTAAAAAAT GTTGGCACGA AAAGGATATG ATGTTGGATG GCAAGAAGGG GACGACAGTA	1620
50	TTTTACAGCA GGGCACGGTT GATTATATTG GCTTTAGTTA TTACATGTCT ACGGCTGTAA	1680
	AACATGATGT TGATACTACA GTTGAAAACA ACATCGTCAA CGGTGGTTTG AATCATTCTG	1740

55

5 GATATACATT GAATGTGTTA TATGATCGTT ATCAGTTACC ACTTTTTIATT GTGGAAAATG 1860
 GTTTTGGTGC AGTTGATGAA GTGGTAGATG GACATATTCA TGATGATTAT CGCATTGAAT 1920
 10 ATTTAAAAGC ACATATTACA GCAGCGATAG AAGCAGTTGA TCAAGATGGT GTAGATTTAA 1980
 TCGGTTATAC ACCGTGGGGA ATCATTGATA TTGTTTCATT TACAACCGGT GAAATGAAGA 2040
 AACGCTATGG TTTAATATAT GTTGATCGAG ATAATGATGG TCATGGCAGC ATGGAACGCT 2100
 15 TGAAAAAAGA TTCGTTCTAT TGGTATCAAC AAGTGATAGC ATCAAATGGA GATAAATTAT 2160
 AAAGGTATAT TATAAGTATT TTAGGGTTAG AGCCCGAGAC ATAAATTAAT ATAGTAGGAC 2220
 CTACAGTGTT ATAATGGCGG gCCCCCAACA CAAAGAATTT CGAAAAGAAA TTCTAcAGGT 2280
 aATGCaAGtT GGCGGGGcCC AACACAGAGA AATTCGAAAA GAAATTCTAc AGGTAATGCA 2340
 AGTTGGGGAA GGACAGAAAT AAATT 2365

20 (2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11050 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

30 CTGCGATACG ATTTGTTGAA AGTGGGGAAA ACAAAAAAGT TATCATTACC AATTTAGAGC 60
 AGGCATACGA AGCTTTGATT GGTAATAAAG GTACACACAT TCACATGTAG CACTTTATCA 120
 35 CGCGACAAAA CATTAAATAT GTTCTCCGT TGATTCAAAT GAAAAAGTTG TCTGCTGACA 180
 CTTTGCAAGG TTTGAAGGAG TTTAACTTAT GACAGAAAAC TTTATTTTGG GTAGAAATAA 240
 TAAATTAGAA CATGAACTAA AGGCATTAGC AGATTACATT AATATACCAT ATAGTATATT 300
 40 ACAACCATAT CAAAGTGAAT GTTTTGTGAG ACATTATACG AAAGGCCAAG TTATTTATTT 360
 TTCGCCACAA GAAAGTAGCA ATATTTACTT TTTAATTGAA GGTAACATTA TTAGAGAACA 420
 TTACAATCAA AATGGAGATG TATATCGTTA TTTTAATAAA GAGCAAGTAT TATTTCCAAT 480
 45 CAGTAACTTA TTTCATCCGA AAGAGGTTAA CGAATTGTGT ACAGCATTA CCGATTGTAC 540
 AGTTCTTGGA TTGCCTAGAG AATTGATGGC CTTTTTGTGC AAAGCTAATG ATGATATATT 600
 TTTGACACTT TTTGCATTAA TAAATGATAA TGAGCAGCAA CACATGAAC ATAACATGGC 660
 50 ATTAACAAGT AAATTGCTA AAGATCGAAT TATCAAATTG ATATGCCATC TATGTCAGAC 720
 AGTAGGATAC GATCAAGATG AATTTTATGA AATCAAACAG TTTTAACTA TTCAACTCAT 780

55

	TGAAAACTT GTTGTTAAAG ATCATAAAAA TTGGTTAGTA AGCAAACATT TATTCAATGA	900
	TGTATGTGTT TAATATACAA TGTAAGTGA ATAAGTTGAA CATGAGGTCT AACGTACATT	960
5	TATACGTTAG GCCTTTTTTG CTAGCATGAT GAATAATTGA AAATGTTAGT TAAATTTGAT	1020
	TGTTGAAATT ACAGTAAAT TTAAGGTGAT GAAAAATTGA GAACTTCTAA GTTTTTGAAA	1080
	AGTAAAAAT TTGTAATAGT GTAAAAATAG TATATTGATT TTTGCTAGTT AACAGAAAT	1140
10	TTTAAGTTAT ATAAATAGGA AGAAAAACAA TTTTACGTAA TTTTTTCGA AAAGCAATTG	1200
	ATATAATTCT TATTCATTA TACAATTTAG ACTAATCTAG AAATTGAAAT GGAGTAATAT	1260
	TTTTGAAAAA AAGAATTGAT TATTGTGCGA ATAAGCAGAA TAAGTATTCG ATTAGACGTT	1320
15	TTACAGTAGG TACCACATCA GTAATAGTAG GGGCAACTAT ACTATTTGGG ATAGGCAATC	1380
	ATCAAGCACA AGCTTCAGAA CAATCGAACG ATACAACGCA ATCTTCGAAA AATAATGCAA	1440
20	GTGCAGATTC CGAAAAAAC AATATGATAG AAACACCTCA ATTAAATACA ACGGCTAATG	1500
	ATACATCTGA TATTAGTGCA AACACAAACA GTGCGAATGT AGATAGCACA ACAAACCAA	1560
	TGTCTACACA AACGAGCAAT ACCACTACAA CAGAGCCAGC TTCAACAAAT GAAACACCTC	1620
25	AACCGACGGC AATTAAAAAT CAAGCAACTG CTGCAAAAAT GCAAGATCAA ACTGTTCTCT	1680
	AAGAAGCAA TTCTCAAGTA GATAATAAAA CAACGAATGA TGCTAATAGC ATAGCAACAA	1740
	ACAGTGAGCT TAAAAATTCT CAAACATTAG ATTTACCACA ATCATCACCA CAAACGATTT	1800
30	CCAATGCGCA AGGAACTAGT AAACCAAGTG TTAGAACGAG AGCTGTACGT AGTTTAGCTG	1860
	TTGCTGAACC GGTAGTAAAT GCTGCTGATG CTAAAGGTAC AAATGTAAAT GATAAGTTA	1920
35	CGGCAAGTAA TTTCAAGTTA GAAAAGACTA CATTTGACCC TAATCAAAGT GGTAACACAT	1980
	TTATGGCGGC AAATTTTACA GTGACAGATA AAGTGAAATC AGGGGATTAT TTTACAGCGA	2040
	aGTTACCAGA TAGTTTAACT GGTAATGGAG ACGTGGATTA TTCTAATTCA AATAATACGA	2100
40	TGCCAATTGC AGACATTAAA AGTACGAATG GCGATGTTGT AGCTAAAGCA ACATATGATA	2160
	TCTTGACTAA GACGTATACA TTTGTCTTTA CAGATTATGT AAATAATAAA GAAAATATTA	2220
	ACGGACAATT TTCATTACCT TTATTTACAG ACCGAGCAAA GGCACCTAAA TCAGGAACAT	2280
45	ATGATGCGAA TATTAATATT GCGGATGAAA TGTTTAATAA TAAAATTACT TATAACTATA	2340
	GTTGCGCAAT TGCAGGAATT GATAAACCAA ATGGCGCGAA CATTTCTTCT CAAATTATTG	2400
	GTGTAGATAC AGCTTCAGGT CAAAACACAT ACAAGCAAAC AGTATTTGTT AACCCCTAAGC	2460
50	AACGAGTTT AGGTAATACG TGGGTGTATA TTAAAGGCTA CCAAGATAAA ATCGAAGAAA	2520
	GTAGCGGTAA AGTAAGTGCT ACAGATACAA AACTGAGAAT TTTGAAGTG AATGATACAT	2580

55

	ACCAATTTTAA AAATAGAATC TATTATGAGC ATCCAAATGT AGCTAGTATT AAATTTGGTG	2700
	ATATTACTAA AACATATGTA GTATTAGTAG AAGGGCATT A GACAATACA GGTAAGAACT	2760
5	TAAAACTCA GGTATTCAA GAAAATGTTG ATCCTGTAAC AAATAGAGAC TACAGTATTT	2820
	TCGGTTGGAA TAATGAGAAT GTTGACGTT ATGGTGGTGG AAGTGCTGAT GGTGATTGAG	2880
	CAGTAAATCC GAAAGACCCA ACTCCAGGGC CGCCGGTTGA CCCAGAACCA AGTCCAGACC	2940
10	CAGAACCAGA ACCAACGCCA GATCCAGAAC CAAGTCCAGA CCCAGAACCG GAACCAAGCC	3000
	CAGACCCGGA TCCGGATTG GATTGAGACA GTGACTCAGG CTCAGACAGC GACTCAGGTT	3060
15	CAGATAGCGA CTCAGAATCA GATAGCGATT CGGATTCAGA CAGTGATTCA GATTGAGACA	3120
	GCGACTCAGA ATCAGATAGC GACTCAGAAT CAGATAGTGA GTCAGATTCA GACAGTACT	3180
	CGGACTCAGA CAGTGATTCA GACTCAGATA GCGATTGAGA CTCAGATAGC GATTGAGACT	3240
20	CAGACAGCGA TTCAGATTCA GACAGCGACT CAGATTCAGA CAGCGACTCA GACTCAGATA	3300
	GCGACTCAGA CTCAGACAGC GACTCAGATT CAGATAGCGA TTCAGACTCA GACAGCGACT	3360
	CAGACTCAGA CAGCGACTCA GACTCAGATA GCGACTCAGA TTCAGATAGC GATTGAGACT	3420
25	CAGACAGCGA CTCAGATTCA GATAGCGATT CGGACTCAGA CAGCGATTCA GATTGAGACA	3480
	GCGACTCAGA CTCGGATAGC GATTGAGATT CAGATAGCGA TTCGGATTCA GACAGTGATT	3540
	CAGATTCAGA CAGCGACTCA GACTCGGATA GCGACTCAGA CTCAGACAGC GATTGAGACT	3600
30	CAGATAGCGA CTCAGACTCG GATAGCGACT CGGATTCAGA TAGCGACTCA GACTCAGATA	3660
	GTGACTCCGA TTCAAGAGTT ACACCACCAA ATAATGAACA GAAAGCACCA TCAAATCCTA	3720
35	AAGGTGAAGT AAACCATTCT AATAAGGTAT CAAAACAACA CAAACTGAT GCTTTACCAG	3780
	AAACAGGAGA TAAGAGCGAA AACACAAATG CAACTTTATT TGGTGCAATG ATGGCATTAT	3840
	TAGGATCATT ACTATTGTTT AGAAAACGCA AGCAAGATCA TAAAGAAAAA GCGTAAATAC	3900
40	TTTTTTAGGC CGAATACATT TGTATTCGGT TTTTTGTTG AAAATGATTT TAAAGTGAAT	3960
	TGATTAAGCG TAAATGTTG ATAAAGTAGA ATTAGAAAGG GGTCATGACG TATGGCTTAT	4020
	ATTTCAATTAA ACTATCATT ACCAACAAAT GGTATGCATC AAAATTTGAC AGTCATTTTA	4080
45	CCGGAAGATC AAAGCTTCTT TAATAGCGAT ACAACTGTTA AACCATTAAA AACTTTAATG	4140
	TTGTTACATG GATTATCAAG TGATGAAACG ACATATATGA GATATACAAG CATAGAAAGG	4200
	TATGCGAATG AACACAAATT AGCTGTGATT ATGCCCAATG TGGATCATAG CGCATATGCT	4260
50	AACATGGCAT ATGGTCATAG CTATTATGAT TATATTTTGG AAGTGATGA TTATGTTTAT	4320
	CAAATATTTT CACTTTCCAA AAAGCGTGAT GACAAATTTA TAGCAGGTCA CTCTATGGGA	4380
55		

TTATCTGCTG TGTTTGAAGC GCAAAATTTA ATGGATCTAG AGTGGAATGA TTTTTCAAAA 4500
 GAGGCCATAA TTGGCAATCT TTCAAGTGTT AAAGGAACTG AACATGATCC GTATTACTTG 4560
 5 CTAGACAAAG CTGTAGCTGA AGATAAACAA ATTCCAAAAT TGCTCATTAT GTGTGGTAAA 4620
 CAAGACTTTT TATATCAAGA CAACCTTAGAT TTTATCGATT ATTTATCAGC CATAAATGTT 4680
 CCTTATCAAT TTGAAGATGG ACCAGGAGAT CATGATTATG CATATTGGGA TCAAGCGATT 4740
 10 AAGCGTGCTA TAACATGGAT GGTGAATGAT TAATTATTTC TTGGAAAATA TGTGGCTGCA 4800
 TTAAATACAC AGAGTGAGAG ATACAACTA TTTACGCACG ACTAACATT CTAAGTGTTT 4860
 AAATTATTTT TGTATTAATA TGATTGGCGC AATTGCTGA TACACAAAAA TGTTTCTCGT 4920
 15 GAAACTTAGA TTTAGCTTAT AGTTTTATCA TCATTGTAT GACTTACATT ATAAATTTTA 4980
 TTATAATGAG GTTAACGCTT TGAAAGGAGT CATCATCATG TCGACCAATA AAAACGATTA 5040
 20 TGAGCATATG TTGTTTTATT TTGCATATAA AACCTTTATT ACTACCGCTG ATGAAATTAT 5100
 AGAGAAGTAT GGTATGAGTC GTCAGCATCA TCGTTTTTTG TTTTTTATCA ATAAATTACC 5160
 TGGTATTACT ATTAAATCAT TACTAGAAAT ATTAGAAATT TCTAAmCAAG GATCACATGC 5220
 25 AACACTTCAA AAATTAAAAG AGCAAGGTCT CATTATTGAA AAAGTTTtag AGACTGATCG 5280
 ACGTGTCAA AAATTATATT CGACGGATAA AGGCGATCAA CTCATTGCTG AATTGAACAA 5340
 GGCGCAAGAT GAATTATTGC AAAATATATA TCAACAAGTC GGTTCGGATT GGTATGATGT 5400
 30 GATGGAAGCA TTGGCTAAAG GgCGACCTGG cTTTGATTTT ATTAAGCATT TGAAAGATGA 5460
 AAAAGAAAGC TAGCATCAGA AATGTTAAAA ATCTTCGCAT TCTTAAATTT AAAAAATATG 5520
 TCAAAAAGTG TATAATAAAA ACATATAATT TAATTGAACT CAGTTTCAAC ACATCTTAGA 5580
 35 AAGGAGTTTG AATGATGAAA AAATTAGCAG TTATTTTAAC ATTAGTTGGC GGTTTATACT 5640
 TCGCATTAA AAAATACCAA GAACGTGTTA ACCAAGCACC TAACATTGAG TACTAAATTA 5700
 40 AACCATAAAA AATTCCCGAA CACCTTGTTA TAGTGCTCGG GAATTTTTTT ATGCTTTACT 5760
 TGAATATATC AAATATTATT TTTGCGCTTT CTGTATTTTC GATATTACCA CTAAATGATT 5820
 CTGATCTAGG TCCGTAAGCG TagGTATTAA CATCCTCGCC TGTATGTCCA TCGGAAGTCC 5880
 45 ACCCTGTATA AGATTTATCA TTTACTGGCT TCTGAATAGC GTGTTGTAGG GCTTTTGTIT 5940
 GCGTTTCTAC TTCTGCGGAT TTTTCGTCTT TTTCTTTTTT AAGTAGTCTT TTTAGCTTTT 6000
 TATTCTCTTT TTTAACCTTT TTCATATCAT CTTGTGAAAA TTCAAATCCA TAACCTTCAT 6060
 50 TAATAACTTT TTCAGGTCT TCACCTTTAG CCATTTTTTC TGTATATAT GATCCAGAGT 6120
 GTTTCATAGA TTTAATCGGT TGAGGATTCC ATTCGTATCC TTTATCTTTA CCAATTGTTA 6180

55

	ATTGAATGGC GTCATCGAAT GCTTTTTCAA AACCTTCCAT TTCAGACATA ACGCCTGTAA	6300
	TATCGTTGGA ATGCGCTGAT TTATCTATAG AAGCACCTTC GACCATTAAA AAGAATCCTT	6360
5	TTTTATTGCG CTCAAGCTTA CTAAGTGCAC TTTGTTGCAT ATCAGCTAAT GATGGTTCGT	6420
	CTTTAGAAGC ATCTATTGCA AGTGGCATAT TTTTATCTGC AAACAAACCA AGAACTTTAT	6480
	CTTTATCAGA TTTTGATAAC TCCTTACTGT TCGTGGCAAG GTCGTAACCA TCTTTTTTGA	6540
10	ATTTTTTATC TAAATTGCCA TTACTTTTAC CGAAATATTT AGCGCCGCCG CCTAATAAAA	6600
	CATCAACTTT ATGCTTTCCG TTGATTTTAT CTTTATAAAA TTGTTTAGCG ATTTCTTTTT	6660
	TATCATCTCT AGAAGTCACG TGTGCAGCAT ATGCTGCTGG TGTTCATCT GTTAATTCAG	6720
15	CTGTTGAAAC AAGACCAGTC GACTTACCTT TTTCTTTTGC ACGTTCAAGC ACCGCTTTTA	6780
	CTTTCTGCTT GTTACTGTCA ACACCGATGG CACCATTATA TGTCTTATGA CCAGAACTAA	6840
20	AGGCTGTTCC GCCAGCTGCA GAATCAGTAA TATTCTGTTT TGGGTCATTT GAATATGTAC	6900
	GATTTGTGCC TTTTAAATAT GAATCAAAAG CAGTAGGGGT CATTTCTTTA GCATGCGGAT	6960
	CATTTTTATA ATAACGATAA GCTGTGTAA ATGATGGACC CATGCCATCG CCAACTAAAA	7020
25	AGATAACATT TTTTGGATTT TTAGTATTAC CAACCGCGAA ACTTTCATCT TTAGAACTTT	7080
	TATCGGATTG CGCAATTGCA GGTGTGACAG AACTAAAAAC CGTTGACACG ATAATAAGGT	7140
	TAGCAACTGC AAATTTTGTG GCTTTTTTAA CTGATAACAT AAGACATCCT CCTGAGTATA	7200
30	TGACTATGTC TTCAGTGTA AAGAGGAATT TCGAGCAATT ATGTAGTTTT AGTTAATAAT	7260
	ATGTAAACAG AGTGATTTAG AATAACAAAA ATGAATATA TATGACAATT TGTATAGAA	7320
35	AGCGTTAGAA TAGAAGCGTG TGAAAATATA GAATTAAATA TAATTTGAGG TGGAAAAATG	7380
	ATACTAGTAA TGTTATCTCC ATTATTAATC ATATTCTTTA TAGTGTGTC TATTTTAGAA	7440
	GAGCGTAAAC GTACGAAGAA AAAGCAACTC GAGAAAGAAA AAGCAAATAC ACTAAATCAA	7500
40	AATACAAATG ACACGGAAAG TTCAAATCAA GAGCCGTCAT TGCAGCAGGA TAAAGAACAA	7560
	AAAGATAACA AAGGATAATT CAATTGAAGG AAGAAGATTA TAGATGAAAA TATTAATTGT	7620
	TGAAGATGAT TTTGTTATAG CAGAGAGTTT AGCATCTGAA CTTAAAAAAT GGAATTACGG	7680
45	TGTTATTGTC GTTGAACAAT TTGATGATAT ACTGTCTATC TTAAACCAA ATCAACCTCA	7740
	GCTTGTATTG CTAGATATTA ATTTGCCAAC GTTAAATGGT TTTTATTGGT GTCAAGAAAT	7800
	CCGAAAAACA TCTAATGTGC CAATTATATT TATTAGTTCC CGTATTGATA ATATGGACCA	7860
50	AATTATGGCA ATACAAATGG GGGGAGATGA TTTTATCGAA AAGCCATTTA ACTTGTCATT	7920
	AACGATTGCC AAAATTCAAG CATTATTGAG ACGAACTTAT GACTTGTGAG TAGCTAATGA	7980
55		

	ACAAAACATA CAGCTATCTT TGACTGAATT ACAAATATTA AAGTTATTAT TTCAAAATGA	8100
	AGaTAAATAT GTAAGTAGrA CTGCTTTAAT TGaAAAATGT TGGGaATCAG AAAACtTCAT	8160
5	AGATGATAAC ACATTAGCTG TTAACATGAC GCGCCTGCTG AAAAAATTAA ATACTATTGG	8220
	CGTTAATGAT TTTATCATTa CAAAGAAAAA TGTCGGATAT AAAGTATAGG GTGAATGCAA	8280
	TGACCTTTCT TAAAAGTATT ACTCAGGAAA TAGCAATAGT CATAGTTATT TTTGCTTTGT	8340
10	TTGGCTTAAT GTTTTACCTG TATCATTGTC CATTAGAAGC ATATTTACTA GCACTTGGCG	8400
	TTATTTTATT ATTATTACTC ATATTCATAG GTATTAAATA TTTAAGTTTT GTAAAACTA	8460
15	TAAGCCAACA ACAACAAATT GAAAACCTAG AAAATGCGTT GTATCAGCTT AAAATGAAC	8520
	AAATTGAATA TAAAAATGAT GTAGAGAGCT ACTTTTTAAC ATGGGTACAT CAAATGAAAA	8580
	CACCCATTAC TGCAGCACAA CTGTTACTTG AAAGAGATGA GCCTAATGTT GTTAATCGTG	8640
20	TTCGTCAAGA GGTATTTCaA ATTGaTAACT ATACAAGTTT AGCACTTAGT TATTTAAAGT	8700
	TATTAAATGA AACTTCTGaT ATTTCTGTCA CTAAAATTTc GATTAATAAT ATCATTGCCC	8760
	CAATTATTAT GAAATATTCA ATACAGTTTA TTGATCAAAA AACAAAAATC CATTATGAAC	8820
25	CTTGTCATCA CGAAGTATTA ACTGACGTTA GATGGACCTC TTTAATGATA GAACAATTAA	8880
	TAAATAATGC ACTTAAGTAT GCGAGAGGTA AAGATATATG GATTGAATTT GATGAGCAAT	8940
	CCAATCAATT ACACGTAAAA GATAATGGTA TCGGTATTAG TGAAGCGrAC TTGCCTAAAA	9000
30	TATTTGATAA GGGCTATTCA GGTtATAATG GCCAGCGCCA AAGTAACTCA AGTGGGaTTG	9060
	GTTTATTtAT CGTAAAACAA ATTTCAACAC ACACAAACCA TCCTGTTTCA GTCGTATCTA	9120
35	AACAAAATGA GGGTACAACA TTTACGATTC AATTTCCAGA TGAATAAAAA CTTTCAATAT	9180
	TGTAAGTATA CTAGTAACAT TTTTTTACTA ATTTAAATGT TATTAGTATT TTTTGTtTTT	9240
	AATATAGAAC TAACAAAGAA ATGAGGTGCA TGCCATGTTG CTAGAAGTGn AACATGTAAa	9300
40	AAAGGTTTAT GGTAAAGGTT TGAATGCTAC GACAGCACTT AATCAAATGA ATTTATCAGT	9360
	TGGAGCTGGT GaATTTGTTG CaATTATGGG TGAGTCTGGG tCAGGGAAGT CTACACTACT	9420
	AAATTTAATT GctTCTTTTG ATGGACTAAC TGAAGGTGAC ATTATTGTGG ATGGCGCACA	9480
45	TTTAAATAAT ATGAAAAATA AAAGTAAAGC ATTGTATCGT CaACAAATGG TAGGTTTTGT	9540
	TTTTcAAGAT TTTAATCTTT TACCAACAAT GACGAATAAA GAAAATATAA TGATGCCATT	9600
	AATTTTAGCT GGTGCTAAAC GAAAAGATAT AGAACAAAGG GTACATCAGT TGGCAGTACA	9660
50	ATTACATTTA GAGGGATTCT TAAACAAGTA TCCTTCTGAA ATCTCTGGGG GTCAGAAGCA	9720
	ACGCATTGCC ATTGCACGTG CATTAGTTAC TAAGCCGACG ATTTTACTAG CCGATGAACC	9780

55

TCAATTGGAA CAGACAATTT TAATGGTAAC TCATTCAAAT ATCGATGCGT CTTATGCAGA 9900
 GCGAGTCATT TTTATTAAAG ATGGGCGTCT ATATCATGAA ATATATCGTG GTGAAGAAAG 9960
 5 TCAATTAGCT TTTCAACAAC GAATAACAGA TAGCTTAGCA CTTGTGAATG GAGGAAGTGT 10020
 CAATATATGA AGTTAAGATT GTTATGnACA TAGTGCGACG TCAATTTATT ACGCAGCGAC 10080
 TTGTAATCAT TCCATTCATT TTAGCGGTAA GTGTACTATT CATGATTGAA TATACGCTTG 10140
 10 TGTC AATTGG GTTAAATAGC TACATAAAAC AGAAGAATGA CTTCTAGTA CCATTTATTA 10200
 TCATAGCTAA TTTTTTTATG GCGCTTTTAA CTTTTATTTT TATTTTCTAT GCAAATCACT 10260
 TTATGATGTC ACAAAGACGA AAAGAGTTTA GCATTTTTAT GACATTGGGC ATGACCAAGA 10320
 15 AAAGTATGCG TTTAATTGTA GTGATGGAAA CTATCTTACA ATTTGTGATA ATTTCACTCG 10380
 TTAGTATTGC CGCGGATAC TTAGTTGGTG CGATATTTTT CTTGTTTATA CAGAAAATAA 10440
 20 TGGGCAGTGA AGTTGCGACG TTAAGGTATT ATCCATTTGA CTCTGTAGCG ATGTTTATTA 10500
 CTTTGATTAT CATTGCTGTA TTAATGGGCA TGCTACTTAT ATTCAACTTG TTTAGTATTA 10560
 ATTTTCAACG GCCGATAACT TATCAACATC GTTCCGATTC TAGTGTCATA TCACGATGGT 10620
 25 TGCGTTACGT TTTAATTGTT ATAGGAAGCG CAACTATAT TTAGGTTACT TTATTGCATT 10680
 ACAACAAGAT ACGACGTTTG GTGCCTTTTT TAAATATGG ATTGTCATAG GATTAGTTAT 10740
 TATCGGTACT TATGCATTTT TTGTAGGTAT AAGTGAAATA ATTATTAGTA TATTGCAGCA 10800
 30 GGTATCAAAA GTTTACTATC ATCCACGGTA TTTTTTTGTG GTAGTTGGGA TGCCTGTACG 10860
 TCTTAAATG AATGCAGTCA GTCTTGCAAC AATCACTTTG CTGTGTACAT TTTTGATTGT 10920
 AACGCTCACA ATGACATTAA CAACCTATCG TGATATGAAT CATACCATTA CGAAATTGAT 10980
 35 TACGAATGAT TakGATTTGT CATTTAGCGA CAATTCTAAG TCACAAaTAG AACGTCAACA 11040
 AAC AATTGAG 11050

40 (2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 983 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

50 CGACATAACG AGGCAAGGGT ACATGATACT TTAGCCTCGT TTTTGATATG TATTTTTCTG 60
 AATATAAGGG CAATAGATGG TATTTTATAW TTTTTTTAAG GTAGTGATTA ACATAGATAT 120

TCAAGCGGAA CAGCATTATG CACCAGTATT AACGCATTTT TTAGATCCAA GAGGGCAATA 240
 TATATTGGAA GTGATTTGTG GCAGTTATGA AGATTTAAAC GTATCTTTTT ATGGTGGACC 300
 5 TAATGCTGAA AGAAAAAGAG CAATCATTTT GCCGAACTAT TATGAACCTA AAGAAAGCGA 360
 CTTTGAATTA ACTTTAATGG AAATAGATTA TCCTGAAAAA TTCGTCACTT TAAACATCA 420
 ACATATTTTA GGGACATTAA TGTCTTTAGG TATCGAACGC GAACAAGTTG GAGATATAAT 480
 10 TGTGaATGAA CGAATTCAAT TTGTTTTGAC AAGTAGATTG GAATCATTTA TTATGTTAGA 540
 ATTACAACGT ATTAAAGGCG CATCAGTTAA ACTTTATACT ATTCCAGTAA CAGATATGAT 600
 ACAATCTAAT GAGAATTGGA AAAATGAAAG TGCaCAGTTA GTTCTTTAAG GTTAGATGTT 660
 15 GTTATTAAAG AAATGATACG TAAATCACGT ACGATTGCGA AACAACTAAT CGAAAAAATA 720
 CGTGTAAAG TGAATCACAC TATTGTTGAT TCAGCAGATT TTCAATTACA AGCAAATGAT 780
 20 TTAATATCCA TCCAAGGTTT TGGTAGAGCA CACATTACTG ACTTAGGTGG TAAAACTAAA 840
 AAAGATAAAA CGCACATTAC CTATAGAACA TTATTCAAAT AGTAATGATT TAAGGAGGAT 900
 AACAAATGCC TTTTACACCA AATGAaATTA AGAATAAAGA GTTTTCACGT GTaAAGAATG 960
 25 GTTTTAGAAC CTACTGnAGT TGG 983

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 10322 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

TTTTGCAAAG CTTATTTTAT GTCAAACAGA TAGTCAATGT GAAACAAAGG TTAGTACATA 60
 40 TAATCATCCA GACTTTATGT ATATATCAAC AACTGAGAAT GCAATTAAGA AAGAACAAGT 120
 TGAACAACTT GTGCGTCATA TGAATCAACT TCCTATAGAA AGCACAAATA AAGTGACAT 180
 CATTGAAGAC TTTGAAAAGT TAACTGTTCA AGGGGAAAAC AGTATCTTGA AATTTCTTGA 240
 45 AGAACCACCG GACAATACGA TTGCTATTTT ATTGTCTACA AAACCTGAGC AAATTTTAGA 300
 CACAATCCAT TCAAGGTGTC AGCATGTATA TTTCAAGCCT ATTGATAAAG AAAAGTTTAT 360
 AAATAGATTA GTTGAACAAA ACATGTCTAA GCCAGTAGCT GAAATGATTA GTACTTATAC 420
 50 TACGCAAATA GATAATGCAA TGGCTTTAAA TGAAGAATTT GATTTATTAG CATTAGGAA 480
 ATCAGTTATA CGTTGGTGTG AATTGTTGCT TACTAATAAG CCAATGGCAC TTATAGGTAT 540

55

	GAATGGTTTC TTCGAAGATA TCATACATAC AAAGGTAAAT GTAGAGGATA AACAAATATA	660
	TAGTGATTTA AAAAATGATA TTGATCAATA TGCACAAAAG TTGTCGTTTA ATCAATTAAT	720
5	TTTGATGTTT GATCAACTGA CGGAAGCACA TAAGAAATTG AmTCAAAATG TAAATCCAAC	780
	GCTTGTATTT GAACAAATCG TAATTAAGGG TGTGAGTTAG ATGCCAAATG TAATAGGTGT	840
10	TCAGTTTCAA AAAGCGGGAA AATTAGAATA TTATACACCT AATGATATAC AAGTAGATAT	900
	AGAAGACTGG GTAGTTGTCTG AATCTAAAAG AGGCATAGAG ATAGGTATTG TTA AAAATCC	960
	ATTAATGGAT ATTGCTGAAG AGGATGTTGT GTTACCTCTT AAAAATATTA TTCGCATTGC	1020
15	TGATGACAAA GATATTGATA AATTTAATTG TAATGAACGA GATGCTGAAA ATGCATTAAT	1080
	ACTATGTAAA GACATTGTAA GAGAACAAGG TTTGGACATG CGTTTAGTCA ATTGCGAATA	1140
	TACATTAGAT AAATCGAAAG TTATTTTTAA TTTTACGGCG GATGATCGTA TTGATTTTAG	1200
20	AAAATTAGTA AAAATATTAG CGCAACATTT AAAACACGT ATCGAGTTGA GACAAATTGG	1260
	TGTAAGGGAT GAAGCCAAAT TGCTTGGCGG TATCGGACCT TGTGGTAGGT CGTTATGTTG	1320
	TTCTACATTT TTAGGGGATT TTGAACCAGT ATCGATTAAG ATGGCTAAGG ATCAAAATTT	1380
25	ATCATTAAAT CCAACTAAAA TTTCTGGTGC ATGTGGTCGT TTGATGTGTT GTTTAAATA	1440
	TGAAAATGAC TATTATGAGG AAGTACGTGC ACAATTACCT GATATTGGTG AAGCAATTGA	1500
30	AACGCCTGAT GGTAAACGGGA AAGTAGTTGC TTAAATATA TTAGACATTT CTATGCAGGT	1560
	GAAGCTTGAG GGACATGAAC AGCCACTTGA ATATAAATTA GAAGAAATAG AACTATGCA	1620
	TTAAGGAGGC ATTATTACAT TTGGATCGCA ATGAAATATT TGAAAAATA ATGCGTTTAG	1680
35	AAATGAATGT CAATCAACTT TCAAAGGAAA CTTCAGAATT AAAGGCACCT GCAGTTGAAT	1740
	TAGTAGAAGA AAATGTAGCG CTTCAACTTG AAAATGATAA TTTGAAAAG GTGTTGGGCA	1800
	ATGATGAACC AACTACTATT GATACTGCGA ATTCAAAACC AGCAAAAGCT GTGAAAAGC	1860
40	CATTACCAAG TAAAGATAAT TTGGCTATAT TGTATGGAGA AGGATTTTCAT ATTTGTAAAG	1920
	GCGAATTATT TGGAAAACAT CGACATGGTG AAGATTGTCT GTTCTGTTTA GAAGTTTTAA	1980
45	GTGATTAATC AAGCACACTC AAATAGTGTT ATAATTATAA ATGAATATGG TTTGGATAAG	2040
	TCTGAGACAA TGCATGTTTC AGGCTTTAAT TGTGTATAAA GTTTTGGTGA TTGCATAAGA	2100
	GATGGCGGTA CTAAATGTTA TTATTAAGTG TGCACGCagT ATCaTTAGTT ATAAAATGTA	2160
50	GCTGTTAAAA GTCAAAAATA CATCGAATGT AGTTAGGCAT ATAATATAAA AAGAGTTTTTC	2220
	AATTACTCAA TAGAAAAAGG TTGTCTTCAT AGGAGTTAAA AATGTTAAAA GAGAATGAAC	2280
55	GATTTGATCA ACTAATCAA GAAGATTTTA GTATTATTCA AAATGATGAT GTTTTTTCAT	2340

	TGGACTTATG TTCAGGCAAT GGGGTGATAC CCTTGTTATT GTTTCGAAA CATCCACGAC	2460
	ATATAGAAGG TGTGAGATT CAAAAACAC TTGTCGATAT GGCGCGACGC ACATTTCAAT	2520
5	TCAATGATGT TGATGAATAT TTAACAATGC ATCACATGGA TTTGAAAAAC GTTACTAAAG	2580
	TATTTAAACC TTCACAATAT ACTTTAGTAA CGTGTAATCC GCCTTATTTT AAAGAGAATC	2640
10	AGCAACACCA ACATCAAAAA GAAGCACATA AGATAGCGAG ACATGAGATT ATGTGTACAC	2700
	TTGAAGATTG CATGATTGCA GCCCGTCATT TATTAAGA AGGTGGCAGG CTAAACATGG	2760
	TACATCGTGC AGAGAGACTA ATGGATGTCT TGTGTAAT GAGAAAAGTG AATATTGAAC	2820
15	CTAAGAAAGT CGTTTTTATA TATAGTAAAG TAGGGAAATC AGCACAAACG ATAGTAGTAG	2880
	AAGGTCGAAA AGGTGGAAT CAAGGTTTAG AAATCATGCC CCCATTTTAT ATTTATAATG	2940
	AAGATGGTAA TTATAGCGAA GAAATGAAGG AAGTATATTA TGGATAGTCA TTTTGTATAT	3000
20	ATTGTAAAAT GTAGTGATGG AAGTTTATAT ACAGGATACG CTAAAGACGT TAATGCACGT	3060
	GTTGAAAAAC ATAACCGAGG TCAAGGAGCC AAATATACGA AAGTAAGACG TCCGGTGCAT	3120
	TTAGTTTATC AAGAAATGTA TGAGACAAAG TCTGAAGCAT TGAAGCGTGA ATATGAAAT	3180
25	AAAACCTTATA CCAGACAAAA GAAATTGCGA TTAATTAAGG AGCGATAGTA TGGCTGTATT	3240
	ATATTTAGTG GGCACACCAA TTGGTAATTT AGCAGATATT ACTTATAGAG CAGTTGATGT	3300
30	ATTGAAACGT GTTGATATGA TTGCTTGTA AGACACTAGA GTAAGTAGTA AACTGTGTAA	3360
	TCATTATGAT ATTCCAATC CATTAAAGTC ATATCACGAA CATAACAAGG ATAAGCAGAC	3420
	TGCTTTTATC ATTGAACAGT TAGAATTAGG TCTTGACGTT GCGCTCGTAT CTGATGCTGG	3480
35	ATTGCCCTTA ATTAGTGATC CTGGATACGA ATTAGTAGTG GCAGCCaGAG AAGCTAATAT	3540
	TAAAGTAGAG ACTGTGCCTG GACCTAATGC TGGGCTGACG GCTTTGATGG CTAGTGGAAT	3600
	ACCTTCATAT GTATATACAT TTTTAGGATT TTTGCCACGA AAAGAGAAAG AAAAAAGTGC	3660
40	TGTATTAGAG CAACGTATGC ATGAAAATAG CACATTAATT ATATACGAAT CACCGCATCG	3720
	TGTGACAGAT ACATTAAAA CAATTGCAAA GATAGATGCA ACACGACAAG TATCACTAGG	3780
45	GCGTGAATTA ACTAAGAAGT TCGAACAAAT TGTAAGTAT GATGTAACAC AATTACAAGC	3840
	ATTGATTCAG CAAGGCGATG TACCATTGAA AGGCGAATTC GTTATCTTAA TTGAAGGTGC	3900
	TAAAGCGAAC AATGAGATAT CGTGGTTTGA TGATTTATCT ATCAATGAGC ATGTTGATCA	3960
50	TTATATTCAA ACTTCACAGA TGAAACCAAA ACAAGCTATT AAAAAAGTTG CTGAAGAACC	4020
	ACAACCTAAA ACGAATGAAG TATATAATAT TTATCATCAA ATAAGTTAAT CACTTTATCG	4080
55	ATTaTATGAA ATTTTAAACG ATTTTATAAA CGCAAGCTGT AATTTTAAAT GGTAAGTTAT	4140

	GT TTTT TTAAT GT AAAA TAAA TACATTGAAA GTAATAAATA CCTTAACATT GAATAAGATG	4260
	AAAA TGAGAT GACGAGATAA ATGTTTCGCGT CCGTTGAAAT GCATAGAAAT CTTAGATATT	4320
5	ATTTGAAGTG AGACATTACG AGGAGGAACA GTTATGGCTA AAGAAACATT TTATATAACA	4380
	ACCCCAATAT ACTATCCTAG TGGGAATTTA CATATAGGAC ATGCATATTC TACAGTGGCT	4440
10	GGAGATGTTA TTGCAAGATA TAAGAGAATG CAAGGATATG ATGTTTCGCTA TTTGACTGGA	4500
	ACGGATGAAC ACGGTCAAAA AATTCAAGAA AAAGCTCAAA AAGCTGGTAA GACAGAAATT	4560
	GAATATTTGG ATGAGATGAT TGCTGGAATT AAACAATTGT GGGCTAAGCT TGAAATTTCA	4620
15	AATGATGATT TTATCAGAAC AACTGAAGAA CGTCATAAAC ATGTCGTTGA GCAAGTGTTT	4680
	GAACGTTTAT TAAAGCAAGG TGATATCTAT TTAGGTGAAT ATGAAGGTTG GTATTCTGTT	4740
	CCGGATGAAA CATACTATAC AGAGTCACAA TTAGTAGACC CACAATACGA AAACGGTAAA	4800
20	ATTATTGGTG GCAAAAGTCC AGATTCTGGA CACGAAGTTG AACTAGTTAA AGAAGAAAGT	4860
	TATTTCTTTA ATATTAGTAA ATATACAGAC CGTTTATTAG AGTTCTATGA CCAAAATCCA	4920
	GATTTTATAC AACCACCATC AAGAAAAAAT GAAATGATTA ACAACTTCAT TAAACCAGGA	4980
25	CTTGCTGATT TAGCTGTTTC TCGTACATCA TTTAACTGGG GTGTCCATGT TCCGTCTAAT	5040
	CCAAACATG TTGTTTATGT TTGGATTGAT GCGTTAGTTA ACTATATTTT AGCATTAGGC	5100
30	TATTTATCAG ATGATGAGTC ACTATTTAAC AAATACTGGC CAGCAGATAT TCATTTAATG	5160
	GCTAAGGAAA TTGTGCGATT CCACTCAATT ATTTGGCCTA TTTTATTGAT GGCATTAGAC	5220
	TTACCGTTAC CTAAAAAGT CTTTGACAT GGTGGATTT TGATGAAAGA TGGAAAAATG	5280
35	AGTAAATCTA AAGGTAATGT CGTAGACCTT AATATTTTAA TTGATCGCTA TGGTTTAGAT	5340
	GCTACACGTT ATTATCTAAT GCGTGAATTA CCATTGGTT CAGATGGCGT ATTTACACCT	5400
	GAAGCATTG TTGAGCGTAC AAATTTGAT CTAGCAAATG ACTTAGGTAA CTTAGTAAAC	5460
40	CGTACGATTT CTATGGTTAA TAAGTACTTT GATGGCGAAT TACCAGCGTA TCAAGGTCCA	5520
	CTTCATGAAT TAGATGAAGA AATGGAAGCT ATGGCTTTAG AAACAGTGAA AAGCTACACT	5580
45	GAAAGCATGG AAAGTTTGCA ATTTCTGTG GCATTATCTA CGGTATGGAA GTTTATTAGT	5640
	AGAACGAATA AGTATATTGA CGAAACAACG CCTTGGGTAT TAGCTAAGGA CGATAGCCAA	5700
	AAAGATATGT TAGGCAATGT AATGGCTCAC TTAGTTGAAA ATATTCGTTA TGCAGCTGTA	5760
50	TTATTACGTC CATTCTTAAC ACATGCGCCG AAAGAGATTT TTGAACAATT GAACATTAAC	5820
	AATCCTCAAT TTATGGAATT TAGTAGTTTA GAGCAATATG GTGTGCTTAA TGAGTCAATT	5880
55	ATGGTTACTG GGCAACCTAA ACCTATTTTC CCAAGATTGG ATAGCGACGG AATATTGCAT	5940

	AACCTCAAAT TGATATTAAA GACTTTGATA AAGTTGAAAT TAAGGCAGCA ACGATTATTG	6060
	ATGCTGAACA TGTTAAGAAG TCAGATAAGC TTTTAAAAAT TCAAGTAGAC TTAGATTCTG	6120
5	AACAAAGACA AATTGTATCA GGAATTGCCA AATTCTATAC ACCAGATGAT ATTATTGGTA	6180
	AAAAAGTAGC AGTTGTTACT AACCTGAAAC CAGCTAAATT AATGGGACAA AAATCTGAAG	6240
10	GTATGATATT ATCTGCTGAA AAAGATGGTG TATTAACCTT AGTAAGTTTA CCAAGTGCAA	6300
	TTCCAAATGG TGCAGTGATT AAATACTGT ATTTTAAAA ATTAGGAGAG ATAATTATGT	6360
	TAATCGATAC ACATGTCCAT TTAAATGATG AGCAATACGA TGATGATTG AGTGAAGTGA	6420
15	TTACACGTGc TAGAGAAGCA GGTGTTGATC GTATGTTGT AGTTGGTTTT AACAAATCGA	6480
	CAATGAACG CGCGATGAAA TTAATCGATG AGTATGATTT TTTATATGGC ATTATCGTT	6540
	GGCATCCAGT TGACGCAATT GATTTTACAG AAGAACACTT GGAATGGATT GAATCTTTAG	6600
20	CTCAGCATCC AAAAGTGATT GGTATTGGTG AAATGGGATT AGATTATCAC TGGGATAAAT	6660
	CTCCTGCAGA TGTTCAAAAG GAAGTTTTTA GAAAGCAAAT TGCTTTAGCT AAGCGTTTGA	6720
	AGTTACCAAT TATCATTTCAT AACCGTGAAG CAACTCAAGA CTGTATCGAT ATCTTATTGG	6780
25	AGGAGCATGC TGAAGAGGTA GGCGGGATTA TGCATAGCTT TAGTGGTTCT CCAGAAATTG	6840
	CAGATATTGT AACTAATAAG CTGAATTTTT ATATTTTCATT AGGTGGACCT GTGACATTTA	6900
30	AAAATGCTAA ACAGCCTAAA GAAGTTGCTA AGCATGTGTC AATGGAGCGT TTGCTAGTTG	6960
	AAACCGATGC ACCGTATCTT TCGCCACATC CGTATAGAGG GAAGCGAAAT GAACCGGCGA	7020
	GAGTAACTTT AGTAGCTGAA CAAATTGCTG AATTAAAAGG CTTATCTTAT GAAGAAGTGT	7080
35	GCGAACAAAC AACTAAAAAT GCAGAGAAAT TGTTTAATTT AAATTCATAA AGTTAAAAGT	7140
	GAGAAAGATC ACCGCCATAA ATGTAAACGA TGCTATATTC GTTTAATATG CTATGGTTCT	7200
	TTCTCACTTT TTTAAATTAA AATATCGTGC ATGTGGAATA CGTGCGATAG AGATGGTTAG	7260
40	AGCTTTGAAA TTAAGAATTG TAGGAAGGCG TTTTAAATGA AAATCAATGA GTTTATAGTT	7320
	GTAGAAGGAC GAGATGATAC TGAGCGTGTT AAACGAGCTG TTGAATGTGA TACGATTGAA	7380
	ACGAATGGTA GTGCCATCAA CGAACAACT TTAGAAGTAA TTAGAAATGC TCAACAAAGT	7440
45	CGAGGCGTTA TTGTATTAAC AGATCCAGAT TTCCCAGGAG ATAAAATTAG AAGTACAATT	7500
	ACTGAACATG TCAAAGGTGT TAAACATGCG TATATTGATA GAGAAAAAGC TAAAAATAAA	7560
50	AAAGGGAAAA TTGGTGTTGA ACATGCCGAC TTAATTGATA TTAAAGAAGC GTTAATGCAT	7620
	GTTAGTTCAC CCTTTGATGA AGCTTATGAA TCAATTGATA AATCTGTGCT AATAGAGTTG	7680
55	GGGTTAATTG TTGGGAAAGA TGCAAGGCGC CGTAGAGAAA TTTTAAGTAG AAAATTGCGA	7740

	GCGGATGTAA GGCAAGCTTT AGAAGATGAA TGAGGAAGTG AAAATGTTGG ATAATAAAGA	7860
	TATTGCAACA CCATCAAGAA CGCGAGCGTT GTTAGATAAA TATGGCTTTA ATTTTAAAAA	7920
5	AAGTTTAGGA CAGAACTTTT TGATAGATGT GAATATCATT AATAATATCA TTGATGCAAG	7980
	TGATATTGAT GCACAAACTG GGGTGATTGA AATTGGTCCA GGCATGGGGT CATTGACAGA	8040
10	ACAATTGGCC AGACATGCTA AAAGAGTATT GGCATTTGAA ATTGATCAAC GTTTAATACC	8100
	TGTATTAAAT GATACACTAT CACCTTATGA TAATGTGACG GTGATTAATG AAGATATTTT	8160
	AAAAGCGAAT ATTAAAGAAG CTGTTGAAAA TCATTTACAA GATTGTGAAA AAATAATGGT	8220
15	TGTTGCAAAC CTGCCGTACT ATATTACGAC GCCAATTTTA TTAAATTTGA TGCAACAAGA	8280
	TATACCAATT GATGGCTACG TGGTGATGAT GCAAAAAGAA GTGGGCGAAC GCTTAAATGC	8340
	TGAAGTAGGT TCAAAAGCAT ATGGTTCGTT ATCAATTGTC GTACAATACT ATACAGAGAC	8400
20	TAGTAAAGTA TTAACGGTAC CTAAATCTGT ATTTATGCCA CCACCTAATG TTGATTCAAT	8460
	AGTTGTAAAA CTGATGCAGA GAACTGAACC GTTAGTAACA GTAGATAACG AGGAAGCATT	8520
	CTTTAAGTTA GCAAAAGCAG CATTTGCACA AAGAAGAAAG ACAATTAACA ATAACATCA	8580
25	AAATTATTTT AAAGATGGTA AACAACACAA AGAAGTGATT TTACAATGGT TGGAACAAGC	8640
	AGGTATTGAT CCAAGACGTC GCGGTGAAAC GCTATCTATT CAAGATTTTG CTAAATTGTA	8700
30	TGAAGAAAAG AAAAAATTCC CTCAATTAGA AAATTAAATG ATTGACAAAG CAAAGCACTA	8760
	TTGTTAAAT TTAATTTTG TTTGACGAAA AC GTTGCAA TATGGTATTA TGTAACTGT	8820
	AGCGAGGTGG AGCAATATGC CAAAATCAAT TTTGGACATC AAAAAATTCTA TTGATTGTCA	8880
35	TGTAGGAAAT CGTATTGTAC TGAAAGCCAA TGGAGGCCGT AAGAAACAA TAAACGTTT	8940
	TGGAATTTTA AAAGAAACAT ATCCGTCAGT TTTCAATTGTT GAGTTAGATC AAGACAAACA	9000
	CAACTTTGAG AGAGTATCTT ATACATACAC TGATGTGTTA ACTGAAAATG TTCAAGTTTC	9060
40	ATTTGAAGAG GATAATCATC ACGAATCAAT TGCACACTAA ATAAGACATA TAGAGATGTT	9120
	AGACGTTTCT TAGTATAAGA AGTAAATATT ATGATAATTA TTTGAGTGTT GGGCATTATG	9180
45	TTCAATACTC TTTTATTTA CAAAATGTTT AACACTGATG TTTGCTTAT AGATTTTCA	9240
	GTAAATGGAT AATTGTATTT ATAAACACAA ATACAAGTAA ATACTAAGTA ATTAGATGGA	9300
	GAAAATTACT TTTTATTAA AAAACACTA AAAACAAAT TAAAATGTCA AATATTAAAT	9360
50	CTCTTTATGT TAAAATCATC ATATTAAGAT AACGAAAAGA GGGCGGAAAA TGATATATGA	9420
	AACGGCACCA GCCAAAATTA ATTTTACGCT CGATACACTT TTTAAAAGAA ATGATGGCTA	9480
	TCATGAGATT GAAATGATAA TGACAACAGT TGATTTAAAT GATCGTTTAA CTTTTCATAA	9540
55		

AAATCTCGCA TATCGTGCAG CGCAACTATT TATTGAGCAA TATCAACTAA AGCAAGGTGT 9660
 AACAAATTTCT ATCGATAAAG AAATACCTGT TTCTGCTGGC TTAGCTGGAG GTTCGGCTGA 9720
 5 TGCAGCAGCA ACGTTAAGAG GATTGAATCG ACTTTTTGAT ATAGGGGCGA GTTTGGAAGA 9780
 ATTGGCTCTA CTAGGCAGTA AAATCGGGAC AGATATTCCG TTTGTATTT ATAATAAAAC 9840
 10 TGCACTATGT ACTGGAAGAG GAGAGAAAAT CGAGTTTTTA AATAAACCAC CTTCAGCTTG 9900
 GGTGATTCTT GCTAAACCAA ACTTAGGCAT ATCATACCA GATATATTTA AGTTGATTAA 9960
 TTTAGATAAG CGTTACGACG TACATACGAA AATGTGTTAT GAGGCCTTAG AAAATCGAGA 10020
 15 TTATCAACAA TTATGTCAAA GTTTGTCTAA TCGATTAGAG CCAATTTCTG TTTCAAAACA 10080
 CCCACAAATC GATAAATTAA AAAATAATAT GTTGAAAAGT GGTGCAGATG GTGCGTTAAT 10140
 GAGTGAAGC GGACCTACTG TGTATGGGCT AGCACGAAAA GAAAGCCAAG CAAAAAATAT 10200
 20 TTATAATGCA GTTAACGGTT GTTGTAATGA AGTGTACTTA GTTAGACTAT TAGGATAGAA 10260
 GGGTTGAAAA GATGAGATAT AAACGAAGCG AGAGAATTGT TTTTATGACG CAATATTTGA 10320
 TG 10322

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5614 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

GATTGATTAA ATGTTTTTAAT CCACTTCAAT GCCTTCGATA AACTCTACAA TCGCGCTATT 60
 CATATAATTA TTCGATTTC A TTTGTTTCAGC ATATGTCTCA TTAAATCCAG ACATAACTTT 120
 40 TTTAAAWGCG AAAATTGAAA TTGGTATCGT TACTAATAAG GCACTAGCCA TACGCCAATC 180
 AATGAGCATT ATGTATAAAA AGATAGCAGC TGACAAAAGT AAGTTTCCTA TAACTTCAGG 240
 AATCATATGT GCTAAAGGTA ATTCTATTGT TTCAACCTTA TCGACAAATA TATTTTTTAA 300
 45 TTCACCTATT TTCTTAGATT CCACTACGCC TAAAGGGAGA CGCATTAAAT TTTGAGCTAA 360
 TTTTTTACGA ATTTTCAGATA AAATTCATA TGCCGTAATA TGTGATAGCA TCGTTGACGC 420
 50 TCCAAAACAA CCACTTGTG AAATATAAGC GATTAAAGCA ATAAAGATAT AAACCATAAT 480
 CGAATTAATC GTATATGTAT TGTTAATCAT CATTAAAATA ATTTTAAATA CTGCCCAATA 540
 AGGAACTAAT CCAGAAAAGA CACTGATGAT AGACAACAAA ATTGATAACA TAATTTTCCA 600

	ATATGTAACCT CCTkTCAATT AATAATCTAA ATTAAGCCGC TTATATTATT TATTTCACTG	720
	GATGATATAC ATAATATAAA TTTGTTATTT GTTAAAAATT AATACTTATT ACAAGTACAT	780
5	CATATATTAG TTGATAACGA TTATCAATGT CGCGTGGATT TGTGACACAT TTCTTTTAAA	840
	AATTCACAAG GTTATGGGGC AGAAATGATA AAGAGCCACT AATGATTTAT TATGTAGTGG	900
10	TTCTGGGAGT GGGACAGAAA TGATATTTTC ACAAATTTA TTTCGTCGTC CCACCCCAAC	960
	TTGCATTGTC TCTAGAAATT GGAATCCAA TTTCTCTTTG TTGGGTCCCT GAATATAGCC	1020
	TTGTAGAGTC TAGTACATTG ATTTGTATCC CAATGTCCCT ATAATTGATT ATTCGCTTTA	1080
15	TCTAATGATC CTATGACTCA ACTATTAAAT CATTTTTTCGA AATACTTAAT TCTAATATAA	1140
	TTAAATTCAT TTATTGTAAT ATTGCAAAAA TACATTGCAC ACCTTGTTCA TCAATGCTAT	1200
	AATTAATTAC ATAATAAATT GAACATCTAA ATACACCAAA TCCCCTCACT ACTGCCATAG	1260
20	TGAGGGGATT TATTTAGGTG TTGGTTATTT GTCACCTTTT TTATTGTTGC GCGTTCGTAA	1320
	CCAATGTGCA AAAAACGCAA CAAGACAGCC GCTTATAGCT GAAGTCATGA TGTTAATTAA	1380
	TAAATTGAAC ATCCGTCATA CACCTCCTCT CTGCGTTAAA GTAACGCCCG AGATGTTAGG	1440
25	CGACCATCAT ATTATATCAT TTATTTATTA TATTTACGC AATATTAAGG CTTAAGTAAA	1500
	GTITTTTTTTA GTGGTTTACG CTACTTTAAT TGCTATCTTT TAAAATCCAT TTAGATAATA	1560
30	TAAATGTGAT GGGTATCGTA ATAATTAAAC CAGCAAATGG TGCAATTTCT GCTGGCAAAT	1620
	TTAGCCAGGA TACAAATACA TATAATAAAA CTGTTTGTA GCTTACGTTG ACAATCTGCG	1680
	TAATTGGAAT ACTAATGAAT TTTCTCCAAG TAGGTTTTAC CCTGTAAACA AAATAACAAT	1740
35	TCAAATAATA TGAAATCACA AAAGCGACTA GAAATCCGGT AATATGACTA ATCATATATT	1800
	CAATGTGTAA TAATTTTAAAC AGCAATAAAT AGACAACATA ATAATTTAAC GTATTAATGC	1860
	CGCCACCAAT GATAAATTTT AAAATTTTCA CATGCGTTTG TGTTAGTTTC ATATGTGTAC	1920
40	TCCTCAACAT CAAAATATAT GCATAACTAC GTTCTCGAAC ATACTCGAAT ATGCGAGCCA	1980
	ATCCGCTTCA CTTCAAATAT GCTTATTTCA ATCTTTATAC CCTTTCACAG CAAATTTAGT	2040
	CTCTTTCCCC TCATCCTTAT ACGCCATTAT AATGTAACGT ATTTATCGCG TGAATCATT	2100
45	GCACTATAGA GATTACTTTA GTTCACTAGT AATTTTATAT ACAATAAGAG CGACAACAGT	2160
	AATGAGAGGA TGTCTACTAT GCAATTACAA AAAATTGTCA TCGCTCCTGA CTCATTTAAG	2220
50	GAAAGTATGA CCGCACAGCA AGTTGGCAAT ATTATAAAAC AGGCTTTTAC TAATGTTTAT	2280
	GGGAATACCC TTCATTATGA TATCATTCCG ATGGCTGATG GTGGTGAAGG TACCACAGAT	2340
55	GCTTTAATGC ATGCAACAGG TGCCACTAAG TATACAGTCA TCGTTAATGA CCCTTTAATG	2400

	GCGGCAGCGT CAGGTTTGGG TTTATTAGAA AAAGAGGAAC GTAATCCTTT ATACACATCA	2520
	TCATATGGTA CCGGTGAACT AATTAAAGAT GCATTAAATC ATGGTGCTAA GACCATTATT	2580
5	TTAGGGATTG GTGGCAGTGC AACAAATGAT GGTGGTACAG GTATGCTAAG TGCACTAGGC	2640
	GTAAAGTTTA CTGATGTAAA CGGGGACTTA TTACAAATGA ATGGTGCTAA TCTTGCTCAC	2700
	ATTGCACAAA TCGATATAAC CAATCTAGAT TCGCGATTAA AAGAGGTGAC CTTTAAAGTG	2760
10	GCCTGTGATG TTTCAAATCC TTTATTGGGT GAAAATGGTG CTACCTATAT TTATGGTCTT	2820
	CAAAAAGGCG CTGATGCAAA GATGATACCA AAGTTGGATT TCGCAATGTC GCATTATCAT	2880
15	GATAAGATAA AAATGTGCAC AGGAAAGTCC GTTAATCAAA TACCAGGTTT TGGTGCAGCT	2940
	GGCGGTATGG GCGCAGCATT ATTAGCGTTT TGTGAGACAA CTTTAACAAA AGGTATTGAT	3000
	GTGCTCTTTG ACATTACAGA TTTTCATCAA AGAATTAAAG ATGCAGACCT CGTTATTACT	3060
20	GGAGAAGGAC GCATGGATTA TCAGACCATC TTTGGTAAAA CACCCGTAGG CGTTGCGTTA	3120
	GCTGCAAAAC AATATCATAT TCCTGTCATC GCGATTTGTG GCAGTCTAGG CGAAAATTAT	3180
	CAACATGTTT ACGATTTTCG TATTGATAGT GCCTATTCTA TAATCTCTTC ACCTAGCACT	3240
25	TTAGAAGATG TCCTACAAAA TAGCGAACAA AATTTATTAA ACACTGCAAC TGACATTGCT	3300
	CGTATTCTGA AATTACAATA ATGTCAAAGT AAATCATCAG CTTTATTATT TGCAGTTAAA	3360
	ACTTGAATGA GGTGAAACCC ATGAAAAGAA CTGATAAATA CCGTGATTCA TATCAATACG	3420
30	ACAATCAAAA CCAAATCAT CGTCGTCAAT CTGAAGACGC ATCGTATAGA CAACAATATG	3480
	CTAAAGGCGA TCCTGAAGAA CACCCGGAAC GATACTATAA TGGTAGAGAT TATCGAAGAG	3540
35	AACAAATTCT TGAAGAAGAA AACGAGAAAT CCCGCCGTTT AAAAAAATGG TTATATATCA	3600
	TTATTGCCAT TCTCTTAATT ATTGTCGCTA TTTTGTGCAC ACGCGCCTTA CTTAACAATG	3660
	ATAGCGATAA AGTTAGTAAT GACCCTAAAG TCTCTCAAAA TTATAAAAA CAAGTTGAAA	3720
40	ATCAAGACGG CCAAATTAAC CAGCAAGTAG ATAATGCTAA AGAAAATATT AAAAAACAACC	3780
	AAAAAACTGA TGACATTATT AAAAATTTAC AAAATCAAAT CGACAACTTG AAGCAGCAAG	3840
	AACAAAACAA AGCTGATTCT AAGCTAACTC AATTTTATCA AGACCAAATC AACAAATTGA	3900
45	CAGAGGCAAA TAATGCACTT AAAAACAATG CAAGCCAAGG TAAAATTGAA AGCATGTTAA	3960
	ATGATATTAA TACAAAATTC GACAGTATTA AATCTAAATT AGAAAGCTTA TTAAAGATG	4020
	ACAATGGTGG CGCTAATTAA TTATTACACC TGCTTTGATG ATAAACATTA ATTCCCTATA	4080
50	CTTTATCTGT ATCACTACGT TATTCGTGAT GATGCATTAA GAGTATAGGG ATTTTTTATA	4140
	TAAACTTGTA TTCTAACTAC ATACAAATAC ACACAAAACG TATATAATTT ATATAATTAT	4200

55

TTATTGCTAA TTACGTTAGG CGTCATGACC GCTTTTGGCC CACTAACTAT AGATATGTAC 4320
 GTACCATCAT TACCTAAAGT GCAAGGTGAT TTTGGTTCTA CTACATCAGA AATTCAATTA 4380
 5 ACATTATCAT TCACAATGAT TGGTCTTGCA CTAGGCCAAT TTATCTTTGG ACCTTTATCC 4440
 GATGCTTTTG GTCGCAAACG GATTGCTGTA TCCATTTTGA TCATTTTCAT TTTGGTATCA 4500
 GGTTCGTCTA TGTTTGTTGA TCAATTGCCA TTATTCTTAA CTTTACGATT TATTCAAGGT 4560
 10 TTAAGTGGTG GTGGCGTCAT CGTGATTGCA AAAGCCTCTG CTGGTGATAA ATTTAGTGGC 4620
 AACGCACTCG CTAAATTTTT AGCATCTTTA ATGGTAGTTA ATGGCATCAT CACTATTCTT 4680
 GCACCATTAG CCGGTGGATT AGCTTTATCC GTAGCAACAT GCGTTCCTAT TTTCACAATT 4740
 15 TTAAGTATTG TGGCACTCAT CATTTTAATT GCGTCGCTT CTCAATTACC TAAAACATCT 4800
 AAAGATGAAT TAAAGCAGGT GAATTTTAGT AGCGTCATTA AAGATTTTGG AAGTCTTTTG 4860
 20 AAAAAACCAG CATTTATTAT TCCAATGCTA TTACAAGGWT TAACTTATGT AATGCTATTT 4920
 AGTTATTCAT CTGCATCGCC ATTTATTACT CAAAATTGT ATAATATGAC ACCCCAACAA 4980
 TTTAGTATCA TGTTTGCTGT TAACGGTGTA GGTTTAATCA TTGTCAGTCA AGTCGTTGCT 5040
 25 TTATTAGTAG AAAAATTACA TCGCCACATA TTATTAATCA TTTTAACTAT TATACAAGTG 5100
 GTAGGTGTTG CTTTAATTAT CCTGACACTT ACATTCCATT TACCACCTTG GGTCTTACTC 5160
 ATCGCATTCT TCTTAAATGT GTGTCCTGTG ACGTCAATTG GACCGCTTGG TTTCACAATG 5220
 30 GCTATGGAAG AACGAACAGG TGGCAGTGGT AACGCATCAA GTTTACTTGG CTTATTCCAA 5280
 TTTATCTTAG GTGGCGCTGT TGCACCATTA GTTGGCTTAA AAGGCGAATT TAATACATCA 5340
 CCATATATGA TTATTATCTT CATTACAGCC ATTCTATTAG TCAGTCTACA AATCATTTAC 5400
 35 TTTAAAATGA TTA AAAAGCA ACATGTCGCA TAACACTTCA ACATAATTAG AACCTAGCA 5460
 AAGATATCTA TCTTTGTCAG GGTTCCTCTT TATGAATTAT GAGATCGAAT CTTCAACTAA 5520
 AATTACGCCT TCATAGCAAG GACATTTCTA TTCAATCACC CTTTAACAGG CATCCAAATT 5580
 40 TcTGTAATAT ATTTTCACT TGTAGTATCA CCAT 5614

(2) INFORMATION FOR SEQ ID NO: 100:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9179 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

55

	AAAGACAATG ATATGAAGTA TATGGATATC ACAGAAAAG TGCCAATGTC GGAATCTGAA	120
	GTAAACCAAT TGCTAAAAGG TAAGGGGATT TTAGAAAATC GAGGGAAAGT TTTTCTAGAA	180
5	GCTCAAGAAA AATATGAGGT TAATGTCATT TATCTTGTTA GCCATGCATT AGTAGAAACA	240
	GGTAACGGCA AATCAGAATT AGCAAAAGGC ATTAAAGATG GGAAAAAACG CTATTACAAC	300
	TTTTTTGGTA TAGGAGCATT CGATAGTAGT GCTGTTGTA GTGGGAAAAG TTATGCTGAA	360
10	AAGGAACAAT GGACATCACC AGATAAGGCG ATTATTGGTG GTGCAAAGTT CATTGCTAAT	420
	GAATATTTTG AAAACAATCA ACTGAATTTA TATCAAATGC GATGGAATCC AGAAAATCCT	480
	GCGCAACATC AATATGCGAG TGACATTCGC TGGGCAGATA AAATTGCCAA ATTAATGGAT	540
15	AAATCCTATA AGCAGTTTGG TATAAAGAAA GATGATATTA GACAAACATA TTATAAATAA	600
	GACATCGGTG CTTAAAGGAG CTGGAACAAT TTATTGTTTC GAGCTCCTTT AGCGCATTCT	660
20	GAGTGTGTTA GTTAAATGGA TTTTAACCTA ACAAAAAACG CTATATAGCA TCAAATATGC	720
	TATATCCCAC ATCATTGTTA CAAATGTACA TGATGTAAAT GAATATTGCT GTCTAAATGT	780
	GCAATGTAATA TACAATGGTG CAGATAATAC ACTTAAAGTCC TTAAAAATGA AACGTTAGTT	840
25	CCAAGAGTCA TTTTAAACA ATAGTGCATG TGATAAAATA GAAAAGAATG AAAAATATAG	900
	AGGTGACAAT ATGAAGATAG CAATTATAGG TGCAGGCATC GGTGGATTAA CAGCTGCTGC	960
	ATTATTACAA GAACAAGGTC ATACTATTAA AGTCTTTGAA AAAAATGAGT CAGTTAAAGA	1020
30	AATTGGCGCT GGGATTGGTA TCGGAGATAA TGTGCTTAA AACTAGGTA ATCATGACTT	1080
	AGCTAAAGGT ATTAAAAATG CTGGGCAAAT CTTATCTACA ATGACAGTGT TAGATGACAA	1140
	AGATCGCCTG TTAAGTACTG TTAAATTAA AAGTAATACA TTGAATGTGA CGTTACCACG	1200
35	CCAAACATTA ATTGACATTA TTAAATCTTA TGTAAGAT GACGCAATAT TTACAAATCA	1260
	TGAAGTCACG CATATAGATA ATGAGACAGA TAAAGTTACC ATACATTTCG CGGAACAAGA	1320
40	AAGTGAAGCA TTTGATTTAT GTATTGGTGC TGATGGAATT CATTCTAAAG TGAGACAATC	1380
	TGTAAATGCT GACAGTAAAG TATTATATCA AGGGTATACA TGCTTTAGAG GTTTAATTGA	1440
	TGATATTGAT TTAAAGCATC CGGATTGTGC AAAAGAATAC TGGGGaAGAA AAGGaAGAGT	1500
45	AGGTATTGTT CCGTTATTAA ATAATCAAGC ATATTGGTTC ATTACAATTA ACTCGAAGGA	1560
	AAACAATCAT AAATATAGTT CGTTTGTTAA ACCTCATTG CAAGCATACT TTAATCACTA	1620
	TCCAAATGAA GTTAGAGAGA TCTTAGACAA ACAAAGTGAA ACAGGTATCT TATTGCATAA	1680
50	TATTTATGAT TTGAAACCAC TCAAATCTTT TGTTTATGGT CGTACTATTT TACTAGGAGA	1740
	TGCAGCACAT GCGACAACGC CTAATATGGG GCAAGGTGCT GGACAAGCAA TGGAAGATGC	1800

55

	TAAAATACGT	GTCAAACATA	CTGCAAAAGT	AATTAAGCGT	TCTAGAAAAA	TCGGTAAAAT	1920
	TGCCCCAATAT	CGTAGTCGTT	TATTTGTTGC	AGTTAGAAAT	CGTATTATGA	AAATGATGCC	1980
5	AAATGCATTA	GCAGCTGGAC	AAACTAAATT	CTTATATAAA	TCGAAAGAAA	AATAATACAA	2040
	CAATATGAAA	ACCCCCGTAT	GTTGAAACGA	GAGCTCAACA	TATGGGGGTT	CTTGTTTTTA	2100
	TAATGTTATT	ATAATAAATT	CAATTATTAG	TTAACGACAA	ATTGTGGTTT	CTCACCTTGA	2160
10	ACGGCACTAA	TTGCAGCATT	AGCAACAATT	TTAGACATCA	TGTCACGTGC	TTCAAATGTA	2220
	GCATTACCAA	TATGCGGTGT	TAATACTACA	TTATTAAGTG	ATTTTAAGTC	ATCGGTAATA	2280
15	TCTGGTTCAA	ATTCATATAC	ATCAAGTGCA	GCACCTTCAA	TTTCATTATC	TTTCAATGCT	2340
	TGCACTAGTG	CTTGTTTCGTG	CACGATTGGA	CCACGAGAGG	CATTGATTAA	ATACGCCGTA	2400
	GATTTTCATCA	TTTTAAATTG	TTCTGTATCA	ATTAAATGAT	GCATTTTAGG	ATTATAAGCA	2460
20	GCGTTGATAG	TGATAAAATC	TGCATTCTTT	AATAGTGTAT	CTAAATCTAC	ATATTTTGCA	2520
	CCGATTTCTC	GTTCTTTTTTC	TTCTTTGCGA	TTAGGTCCAG	TGTATAGCAC	ATCCATGTCA	2580
	AATGCTCTTG	CACGACGAGC	TACTGCACTA	CCAATTTTCA	CTAAACCGAT	AATGCCGATT	2640
25	GTTTTCCCAG	ATACTTCTCT	ACCTCTGAAA	AATAAAGGTG	CCCATCCATC	AAATCCAGTT	2700
	GTACGTGATA	ATTGGTCCCC	TTCAACAATA	CGACGCGCTA	CTGCAAGTAC	TAATCCAATT	2760
	GTTAAATCAG	CAGTCGCGTT	TGTTGATGCT	TTAGGTGTGT	TTGTAACATC	TATACTTTTT	2820
30	TCTCGGGCAT	ACTCGATATC	AATATTATTA	AAACCAGCGC	CATAGTTGGC	AATGATTTTT	2880
	AAGTCTTTAC	CAGCATCGAT	AACATCTTTA	TCAACGTTTG	TAGATAATAA	ACTAATTAAG	2940
35	GCAGTCGCGT	TTTTAACACC	TTTAATTAAA	GTGTCTTTAT	CGACTAATCC	TTTACCTTCA	3000
	TACATTTCAA	CTTCAAAATG	TTCTTGTAAG	AGTTTTTAAAC	CTACTTCTGG	TATtGCACCA	3060
	GCAACATAAM	CTTTTTCCAT	AAAAGATCAC	TCCTTTTATC	TTAGTATAGT	AGAAGATTAG	3120
40	ACAGTATACA	ACTATGTCAT	GATGTCTTGT	GTATCAATGA	TGTAAGCGCG	TACTTTTGAT	3180
	GGAGGCGATA	TAACTTAGGC	ACTGTAGAAC	TATGAATATT	GTAATGTGGA	AAAAGTGGAT	3240
	CAATTAAATT	AGATAACGTA	GTTTTAAAGT	TAATAGTATT	AGAAAAAATT	AATATTTTGA	3300
45	ATATGGGAGG	AAATATAAAT	AAGTAGGTGG	CAACGAAAAA	TAGCAAAAAA	AGAGCTTCTC	3360
	CTATAAAGGA	AAGCTCAAAG	TTTTTTGATG	ACATATGTAC	TAGAATTAAG	TTTCAAGACA	3420
	ATATGTATCA	TCGTGTTTAT	ATTAAATATG	GATGTAGTTG	TAGTTACCTG	CTTCACTTGC	3480
50	AGAAATAGTT	CTAGAACTTA	CTGAGAAAGG	TCCGCCACTA	TAATTCATTT	CTGAAATTGT	3540
	AACTGAACCA	TCACTGTTTA	CACTTTCTAC	ATATGCAACG	TGACCAAATG	GTCCTTCAGA	3600
55							

	AGCAGCAGCC CAATTATTAG CATTTCCTCCCA AGTAGAACCG ATTTCTCCGC CAACTTTATC	3720
	ATATACATAC CAAGTACATT GTCCTGCAGT GTATAAGTTA CCAGAATGTG AAATTGATGA	3780
5	TGTAGTTGTC GTAGTTGTCG TAGTCGTTGT AGTTTGAGTC GTGTTGTAGT TATAGTTGTT	3840
	GTAATTTGTA TAATTTTCAG CAGCATCTGC ATGATGTGCT TGACCTACTA ATGCTGTGCC	3900
	GATTCCTGCT GTTAACGTAG TTGCTGTTAC TAATTTTTTC ATGAATAAAG TCCTCCAAAG	3960
10	TTCTATATCT TTTTTTATAA ATAAAACGTA GCGACTGTTT TATTCTCACA TCTCGAATTG	4020
	ATGACAATAG TTACTTTAAC AAAATLAATG cTTCTGTGG GGAATGTTAT TGATTTGTAA	4080
15	AAGAATAAAA AAACCTTGAC TAATTTTGTA ATAAAAATTA GTCAAAGTTA CAATGAGATT	4140
	AACAGATAAT TAATAGGAAA TATTTATTTG TAATATGTTT AAATAAATCG AATTGTTAAA	4200
	GGTATTATAT ATTCTTGGCC ATTATAATAT TTGACACACG CAATAATTGT GAATACAAAA	4260
20	GATAATATTG AGAAAGCGAA TATGGATAAA ATACCGATAA ACGTAATGAT GAAACCTATA	4320
	ATAATAATGA AATCAATATC TGTAGCAATT AGGAAAACGC CTATTAAAGT GATAACGACT	4380
	AAAACGATAG ACCAAATAAT ATAAGAAATC GTATAGTTAA GATAATTTTT TCCAGCACGA	4440
25	TCAACTAGTT TCGATTATC TTTTTTCAAT AACCATATTA TCAGTGGACC AATAATAGAT	4500
	GTGAATAAAC TTAATAAATA GATAAGCATC GCCATAATGT TCTCATCATT GGATTTGCGA	4560
	TTCGGTTGAT GATTTGTTAC GTCGTTTATT TCAGTTGTCA TATTAGACAC TCCTTTGAAA	4620
30	ATTGTAATAT TATCTTTAAC TATAACAAAA TATAATCAAA AATAAACATG TTTATTAAAC	4680
	AATTATTAAA AATAAAAATA ATTGGTGGAC GTCGGCGTTT AAATAGGTTA ATTTAAGGTT	4740
35	ATATATACTT AACATTTATA ATGATGCGTA ATGAATTCGC ATCATTTTTA TATTGTCTTA	4800
	CGTATAATTT GTTTTTAATT TTAACCAAAG ATAGAAAGAG GGTGTTTTAT GAAAATAGCA	4860
	ATTGTAGGAT CAGGAAATGG CGCAGTTACG GCAGCAGTAG ATATGGTGAG CAAAGGCCAC	4920
40	GATGTTAAAT TATATTGTCG TAATCAATCT ATAAGTAAGT TTCAAAACGC AATCGAAAAG	4980
	GGCGGATTTG ATTTTAATAA TGAAGGTGAT GAACGTTTCG TAAAATTCAC TGATATTAGT	5040
	GATGATATGG AATATGTTTT AAAAGATGCT GAAATTGTTT AAGTGATTAT TCCATCTTCA	5100
45	TACATAGAGT ATTATGCTGA TGTAATGGCA GAGCATGTAA CTGATAATCA GTTGATATTC	5160
	TTCAACATGG CTGCAGCAAT GGGGTCAATT CGTTTTATGA ATGTTTTAGA AGATAGACAT	5220
	ATTGAAACAA AACCACAAC AGCGGAAGcT AATACGTTGA CGTATGGTAC GCGTGTGAT	5280
50	TTTGAAAATG CAGCAGTTGA TTTATCTCTA AATGTACGTC GTATCTTCTT TTCAACATAT	5340
	GATAGAAGCT GTCTAAATGA TTGTTATGAC AAAGTTTCAA GTATTTATGA TCATTTAGTA	5400

55

	CCAACATTAT TGAATGTCGG TCGCATTGAT TATGCTGGCG AGTTCGCTTT ATATAAAGAA	5520
	GGAATTACTA AACATACAGT TAGATTACTT CATGCAATCG AATTAGAACG TTTGAATTTA	5580
5	GGCCGTAGAT TAGGTTTTGA ATTATCAACA GCTAAAGAAT CACGTATTGA ACGTGGTTAT	5640
	TTAGAACGTG ATAAAGAAGA TGAACCATTA AATCGTTTGT TTAATACAAG CCCAGTATTT	5700
	TCACAAATTC CAGGACCAAA TCATGTAGAA AGCAGATATT TAACTGAAGA TATTGCATAT	5760
10	GGTTTAGTAC TATGGTCAAG CTTAGGTCGT GTTATTGATG TACCGACACC AAATATAGAT	5820
	GCAGTAATTG TAATTGCATC AACCATTTTA GAGAGAGACT TCTTTGAGGA AGGCTTAACA	5880
15	GTTGAAGAAA TTGGTTTGA TAAGCTTGAT TTAGAAAAAT ATTTAAATA AATGATGGCT	5940
	TGAAGATAGA AAAGGATATA GCATTATGCA AAAGCAATAA ATTGAAGAAA AGAGGTTTCT	6000
	CATCAATAAG CGnAGGGGAC GATAGATGAT GAAAAGAAAA CCCACCTTTT TAGAATCAAT	6060
20	TTCGACAATG ATTGTAATGG TTATTGTTGT TGTAACAGGC TTTGTGTTTT TTGATATTCC	6120
	AATTCAAGTA TTATTAATTA TTGCCTCAGC ATATGCCACA TGGATTGCAA AACGTGTAGG	6180
	CTTAACATGG CAAGATTTAG AAAAAGGCAT TGCAGAACGT TTAAATACTG CAATGCCTGC	6240
25	AATTTTAATT ATACTAGCGG TAGGAATTAT AGTAGGCAGT TGGATGTTTT CTGGCACAGT	6300
	GCCAGCCTTG ATTTATTATG GCTTAGATTT ATTGAATCCA AGCTATTTTT TAATATCAGC	6360
	CTTTTTTATA AGTGCTGTTA CATCTGTAGC AACTGGTACA GCATGGGGCT CTGCATCAAC	6420
30	TGCAGGGATT GCACTTATTT CTATTGGTAA TCAATTGGGG ATTCCTCCAG GGATGGCAGC	6480
	GGGTGCTATT ATAGCAGGGG CTGTGTTTGG CGATAAAATG TCACCATTAT CAGATACAAC	6540
35	TAATTTAGCG GCGCTTGTTA CTAAAGTTAA TATATTTAAA CATATACATT CGATGATGTG	6600
	GACGACGATA CCTGCATCAA TCATAGGTTT ATTAGTATGG TTTATTGCTG GATTTCAATT	6660
	TAAAGGGCAT TCAAATGATA AACAGATTCA AACTTTGTTA TCAGAGCTTG CACAGATTTA	6720
40	TCAAATTAAC ATATGGGTCT GGGTTCCCTT AATTGTGATC ATTGTTTGTT TGCTATTTAA	6780
	AATGGCTACA GTGCCAGCTA TGCTAATATC AAGCTTTTCT GCCATTATAG TGGGGACTTT	6840
	TAATCATCAT TTCAAATGA CAGATGGTTT CAAAGCAACA TTTAGTGGTT TTAACGAATC	6900
45	AATGATACAT CAGTCTCATA TTTCATCCAG TGTGAAAAGC TTGTTAGAAC AGGGTGGTAT	6960
	GATGAGTATG ACCCAAATAT TAGTAACGAT ATTTTGCGGA TATGCATTG CAGGTATTGT	7020
	AGAAAAAGCA GGATGTTTAG AAGTCTTATT AACTACTATT TCTAAAGGCA TCCATTCTGT	7080
50	AGGAAGTTTA ATATGTATTA CTGTTATTG TTGTATTGCG CTTGTATTG CTGCAGGTGT	7140
	TGCTTCGATT GTAATTATTA TGGTGCGTGT GTTAATGAAA GATTTGTTCG AAAAATACCA	7200

55

	AATACCATGG GGAACATCAG GTATTTACTA TACGAATCAA CTTTCATGTCT CTGTTGAAGA	7320
	ATTTTTCATA TGGACAGTAC CATGTTATTT ATGCGCAATT ATAGCAATTA TCTATGGTTT	7380
5	TACAGGATA GGTATTAAAA AGTCATCGAA TTCACGTTTA ACTTAATGTG AGCGTGGAAT	7440
	ATATATAATA TGTGAAACA CTTTAATCAT TTATAATTGT AGCGTTATA ATTTGAAAAG	7500
	GTTTTAACTT AGAATAAATA TCCTCTATGC ATATACTGAA TATGTTTTGT AGCGGAACAT	7560
10	GTTGATATAT GTAATGTAAG TTTTATGTCA TGATTTGTAA TGAATAAATT AATTGAGAAT	7620
	TTGAAGGCAA GTATATTTGT AAGTACTTTA ACTAAAAATT TATCAATGTA TAGCCGATTT	7680
	GACATGCCTA AATTTGGGTG TGTCAATGGC TGTATGTTGT TTATTCTTTA TTACAGAGTG	7740
15	AATCGGATTG GTGAAAATCG AAATTTTGAG ATTTTACCA ATTCGATTTT TTTCATAGAA	7800
	ATTAaaaaag CCAACAAGGC TCTTGAAACC TTGTTGGCGT AAACATAGCC ATCACTAATT	7860
20	AGTGAATGAA GTTATAACCA GCAGCTTGGC TAGCTGAGAT TGTACGTGAA GTTACAACAC	7920
	CTGGGCCATA ACCATAGTTC ATTTCTGAAA CTCTTACTGA ACCATTGCTG TTAACACTTT	7980
	CAACGTATGC AACGTGACCG TATGCACCTT GAGTTGTTTG CATAATTGCA CCAGCTTTTG	8040
25	GTGTATTGTT CACTGTGTAA CCAGCTCTTG CAGCTGCGTT AGCCCAGTTA CTTGCATTGC	8100
	CCCAAGTTGA ACCGATTTTA CCACCTACAC GATCAAATAC GTAGTATGTA CATTGACCAG	8160
	AAGTGTATAA GTTACGTCCT GAAGTATAAC CACTTGAGAT TGAACGGCCA TTTGATGATG	8220
30	GAGCCATAGT TGTAGTTACT TGAACATTGT TGCTTGAAGT GCTGTAGCTT GCACCTAAAC	8280
	CACCAGTACG GTAGCTGTTT GTGTTGTAAC TATTATAGTT ATTGTAGTTA TATGATTGAT	8340
	TATTATTTGA GTAGTTGTTG TAACGGCTGT AGTTATTGTA GCTATAACCG TTGTTGTAAT	8400
35	TGTTATAGTT ATTGTAACCA TTGTAGTAGT AATAGCTGTA GTAGCCATTA TCTTGGTTTA	8460
	ATTGACTTGG ATGCCAGTTA CCTTCCATG TGTAATGGTA GTTACCTGT GCATCAATAG	8520
40	TGTAAGTATA GCTATATGAT GTTGGGTCGT TTGGATTATA ACCGTAGTTA TCTTGCTCAG	8580
	AAGCATGAGC TTGATTTCTT GATGCAATTG CGATTGTAGC GAATCCTGCA GTTGCGATAG	8640
	TAGCTGTAGC GATTTTCTTC ATTTTAAAAA TATCCTCCTA AAAATTTTAA ATCTAAAATA	8700
45	TTTTCGTAAT GTCCGTGTGA CAAAATTAAT GTTATAAGTT ATCTCTCGTA ATTAAACGAC	8760
	AAGAAAGACT ATAACAGAAA TTAGCGTCCT TGTGTGCTTT GTTAACGTTT TGTAATTTTT	8820
	TGCTAATATC TTGACACAAT AGAATTTTAA AAGTATAGAA ATTTGCATTT TGCAAAACTT	8880
50	ATAACTACGG CATTCTTTGT GAAACTGAA TGTTCGAAA ATAAGTCTGT TACAAATTTG	8940
	TAATATTACT GAAAATTCTA AATGTATATT TTGTGCATAA TATAGGACTT TTAATCAGAA	9000

55

GGATGAAAAT GTATATTTAA TGGATAAAAT ATCCTAATTT AGCATAAAAA AATGTTTTTAA 9120
TAAAAGTATT ATTTGATATA ATCGATTAT GTTTTGTTAC TGCTAAAAAA CATGTGGCG 9179

5 (2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1868 base pairs
(B) TYPE: nucleic acid
10 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

15 CCTTCAGCCA TTTGACTTCG ACATGAGTTG CCTGTACATA TAAAATAAAT TGTTTTTTTA 60
GTCATAACAA TCTCCTAATT AATTAAAATA TGATAAGTGT TAGATACAAC CCTATGAGGG 120
20 TTATAAATAG TACTGGAATT GTAATGATGA TACCAGTTTT AAAGTATGTG CCCCAAGAAA 180
TCTTAACATC TTTTGTGTT AAGACGTGTA ACCACAGTAA TGCTAGCTAA GAGCCTATCG 240
GTGTAATTTT TGGACCTAAA TCAGAACCGA TAACATTCGC ATAAATTAGG CCTTCTTTTA 300
25 ACATGCCATG GACATTTGAT TGACCAATAG CAATCGCATC TATTAAAACT GTAGGCATAT 360
TATTCAATTAT TGATGATAAA AACGCTGAAA TGAAGCCCAT TCCCAAATA GTGCTAAATA 420
GACCGTAATT GGAAATATAT TCTAATATTT TAGCCAATAT TAAAGTAATG CCAGCATTTTC 480
30 TTAAGCCGAA TACGACGATA TACATACCAA TTGAAAATAA TACTATATTC CAAGGTGCGC 540
CCTTAATGAC TTGCTTAATA TTTACAGCAT TTGATTTACG AGCCAACATT AGAAAAATAA 600
AAGCAATGAT TCCAGTGAAA ATTGATACCG GAATTTTAGT AAATTTACTG ATTAGATAGC 660
35 CGAAAAGTAA TATAACTAGA ACAATCCaTG AAATTTTAAA TAGCTTTAAA TCATTAATGG 720
CATCTTTAGG ATGCTTTATA TTATTATCAT CAAACGTTTT AGGTATCGCT TTTCTAAAAT 780
ATAACCACAA TACTATAATA CTTGCTAAAA GCGAGAATAA ATTAGGTATA ATCATTCTAC 840
40 TAAAATATCG AACGAATCCT ACATGAAAAT AATCAGCAGA TATAATATTC ACTAGATTGC 900
TCACGATTAA AGGTAAAGAA GTTGTGTCAG CTATAAAACC ACTCGCAATA ATAAAAGGGA 960
45 ATATGGCCCG CTTACTAAAA CCTATATTTT TAACCATCGC TAATACAATA GGCGTTAAGA 1020
TTAAcGTGCG CCATCATTTG CGAAAAATGC AGCAACAATG GCACCCAATA ATATGATATA 1080
AACGAACATT TTAAACCAT TGCCTTTTGA AGCATGAAGC ATGTGAATAG CTGACCATTTC 1140
50 GAATAATCCA ACTTTATCTA ATATTAATGA AATAAGAATG ACTGAGACAA AAGTCAAAGT 1200
AGCATTCCAA ACAATACCTG TTACTTCGAA AACATCGGAA AAACCTTACAA CACCAGTAAT 1260

55

TAATACAAAT AATAAGTTA CTAGAAAAAT GAGTGTGCT AAAGTTGTCA TCATTAGCAT 1380
 TCACCACTCT TAAGTTATG ACAAATACAT CGTTGTTAG AGGTATGAAC CTTAGACAAG 1440
 5 TTATTAATTA CGGACTCAA AATATTATGA TTgAGCTGGT ATAAATGTTT ATTTCCGATT 1500
 TTTTCGTGTCG TAACTAAGTT GGTTTTTACT AATGCTTTCA TATGTAGCT AAGTGTAGGT 1560
 TGAGAGAATT GAAAATGTGC TAACAAATCA CAAGCGCATA ACTCTCCACA AGAAAGTAAA 1620
 10 TCTAGTATTT CTAATCTGCT TGAATCTGAT AAACTTTTA AAAATGTTGC TAGTTCCTTA 1680
 TACGTCATAA CATACTCCT AGACGTAAA TAGATTATCA TCTATATAGA TGAATGTCTA 1740
 15 TGTTCCTTTG GTATATTACA CGATATGACT ATGTAATTA AATTGTTTT TAGTATTAAA 1800
 AGGGTATTAA AGATAAATTA TAGATATTGA TTTTGCAAAA TATACTCTTT GTTCTGCATT 1860
 GAAAAAGG 1868

20 (2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15249 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

30 ATTTATGAAA TCCATAGCA TAAACATTAT TCTTGCATCG GCTATACAAA CAGTTACCGC 60
 AAGCAAATTT GTATATCAAC CTGGAATTGT GTTCACGTCA ATGGCAATG CCGATGATGT 120
 GTTATCAGGC GATAGTTATT TTATGGCTGA ATTAATCT ATTAAGCGTA TTGTTGAAAT 180
 35 TCCAGATAAT CAAAAATAT ACTGCTTTAT AGATGAAATT TTAAAGGTA CCAACACAAC 240
 TGAACGAATT GCCGCTTCAG AATCAGTACT ATCATTTTTA CATGAAAAAT CTAACCTTAG 300
 40 AGTTATTGCA GCAACACATG ATATTGAGTT AGCTGAACTC TTAAACAAC GTTATGAAAA 360
 TTACCATTTC AATGAGGTAA TAGAAAATAA TAACATACAT TTTGATTACA AAATTAAGCC 420
 TGGCAAAGCA AATACAGTA ATGCCATCGA ATTATTAAAA ATCACTTCAT TTCCAGCAAA 480
 45 AATATATGAA CGAGCAAAAG ATAATGTCCC GAAAATTTAG CATTTAACCT TAAACATAAA 540
 AACGTCAGCT ATCACATGAC AGAAGACTAT GAACAGTTTC AATAATGTTC ATAGTAATCA 600
 TGTTAATAAC TGACGTTTAT TTTATTCTGC AGAATACTCT TCTAAATCTA TATTGCTGTG 660
 50 CCCATTTAAT GCTAAATCAG CAAATCGACC TTGCTGATAC AAATAGTGGC CGGCAACGCC 720
 TATCATTGCA GCATTATCTG TGCATAATTT AGGACTTGGG ATAGTTAATT GAATGTCATT 780

55

	AACAATTAAT CGCTGAACAC CATATTCTTT ACAAGCTTGA ATAGCTTTAA ACGTGAGCAC	900
	CTCTACAACA CTGTTTTGAA AGCTCGTTGC TACGTTAGCT TCAATGATTG GaATATTTTT	960
5	TTGTCGTTGA TTGTGAAGTT GATTGATTAC GGCACCTTTC AACCCACTAA AACTAAAATC	1020
	ATAACTATCT TTATCCAACC AAACACGAGG GAATGAATAA GTATCTTCAC CTTCAGCAGC	1080
	CAACCGATCA ACTTGTGGAC CACCTGGATA ATTTAAACCA ATTGTTCTGT CCACCTTTATC	1140
10	ATAAGCCTCA CCTACTGCGT CATCTCGTGT TTCACCAATG ACTTCAAATG ATAAATGATC	1200
	CTTCATATAA ACTAATTCAG TATGTCCACC TGAAACAATA AGTGCAATTA GCGGGAATGT	1260
15	TAATGGCTCT TCTATGTGAT TAGCATATAT ATGTCCTGCA ATATGATGAA CAGGAATAAG	1320
	TGGCTnATCG TAAGCAAATG CCAATGCTTT GGCTGCATTA ACACCTATTA GTAACGCACC	1380
	AATTAGTCCA GGGCCTTCTG TAACCGCTAT GGCATCAATA TCTTCTATTG ATACATCGGC	1440
20	ATCCCCTAGA GCCTCGTTTA TTGTTGCTGT TATACCTTCA ACGTGATGTC TACTTGCCAC	1500
	TTCGGGAACG ACACCGCCAA ATCGTTTATG ACTTTCAATC TGACTTAAAA CTGTATTTGA	1560
	TAAAATATCT CTGCCATTTT TTATAACACT AACGCTTGTT TCATCACAAC TTGTTTCAAC	1620
25	AGCTAGTATT AATATATCTT TAGTCATTTA AATTCACCCA CATAACCATT GCGTCCTCAC	1680
	CTTCACCATA ATAATTTTTA CGTTTACCAC CATATTGAAA TCCTAAATTT TCATATACAT	1740
	GTTGTGCCAC TTTATTATTA ACTCTTACTT CTAAACTCAT CACATCACAA GTGTGACTTG	1800
30	CATAGTTTAT TCCGTATTTT AAAAGCATTT GACCTAAACC ATAGCCTCTA TAATTATCAT	1860
	CGATTGCAAC TGTTGTAATT TGAGCTTGAT CGATAACAAT CCATAAACCT AAATAACCAA	1920
35	TAATTTGTTG TTCAAATTCt AAGACAAAAT ATTTGCAAA GTTATTTTGC TCTATTTTAT	1980
	GATAAAATGC GTCAATTGTC CAAGAACTGT CATTGAAACT CCGACGCTCA AGATCAAAGA	2040
	CTTGTTGGCAC ATCTTCTTTA GTCATCTCTC TAATGTTTAA TTGTTCTTTT GACTGTTGAT	2100
40	CCAATTTCTG TCCGCCTCAG CTAATTTATG GTATTTAGGA GTAAATGTAT GTACGTCTGA	2160
	AGGTTTATCT AGCAATTGAT ACATGACTGA TGCATTTGGT AGctGCGCAA TCACTTCACC	2220
	TTGTAATTCA TCTTGTAATT TTACAGTATC TTTCCCAATA TAAATAAATG GTTGGTTTAA	2280
45	ATCTTCTAAA AAAGCTCGCA ATGCCTCTAT CGACATATAT TGATCTTCTA AAATAGTCAC	2340
	TAATTGACCA TTTTGCCACT GGAATATGCC TGTATAAACT GCTTGTCTGC TTGCATCAAA	2400
	CACAGGAACC AATAATTTAT CAGTATGATC GATTGTTGCT GCCAATGCCT TTAATGATGA	2460
50	AACACCATAT AATTTAACAT CTAACGCATA CGCTAATGTT TTAGCAACAG TAACACCGAT	2520
	ACGTAAGCCA GTATATGAAC CAGGACCTTC AGCAACAATA ATCGCATCTA ATTGCTGTTT	2580

55

	TTGTTTAGAA TCCGTAGTTA TTTCAGCTAA AACTTCATCG TTTTGCATCA ATGCTACTGA	2700
	TAATGGTTGA TTCGATGTAT CAATGAGCAG CGAATTCATG GATAATTGCC TCCTTAATTT	2760
5	GTTCATAATG TTCTCCTTGC GCGAACAACCT CAATTTGTCT TGTATTTTCA GATATTGTTG	2820
	AAATGTTAAT AGATAAATGC GTCGCTGGAA GTAAATCTTT TATAAATTGA CTCCATTCAA	2880
	TAACAGTAAT TGCCTGATCT TCGAAAAATT CATCAAATCC TAAATCTTCA TCAGAATCTT	2940
10	CTAAGCGATA ACAATCCATA TGATGCAATT TTAAATTTTT ACCCCTATAT GATTTAATGA	3000
	TGTTAAATGT CGGGGAATTA ATCGTACGTC TTACACCAAG AGCTTTTCCT ATAAATTGCG	3060
15	TTAACGTTGT TTTACCTGCT CCTAAATCTC CGTTAAGTAA AATCAAATCA CCACTTTTCA	3120
	ATTGCTCAAC TAAAAATATA GCAAATTGAT TCATTTTCATC TAAATTATTT ATCTTTATCA	3180
	ATGTTGATTC TCCTATATTA TGCTTTTCAT TCATAAAAAT GATTATCCAT TGTTCATCG	3240
20	TATCTAACTT TATATTTAAC CTTTATATTG TAACAAATTT CAACTTAAAT TTCTTATCTT	3300
	TGAAACAGAT TATCTATTCA AAGTTAATTG TAAGAAAATT TAAAATATTT GTTGACATAC	3360
	TAAAGCAGAT ATAGTAAATT AAATTTATCA AATTTT TAGA CAATTCTAAC TATTAAAGTG	3420
25	ATATATACCA TTCACGGAAG GAGTATAATA AAATGCTTAA TCAATATACT GAACATCAAC	3480
	CGACAACCTC AAATATTATT ATTTTATTAT ACTCTTTAGG ACTCGAACGT TAGTAAATAT	3540
	TTACTAAACG CTTTAAGTCC TATTTCTGTT TGAATGGGAC TTGTAAACGT CCCAATAATA	3600
30	TTGGGACGTT TTTTATGTT TTATCTTTCA ATTACTTATT TTTATTACTA TAAACATGA	3660
	TTAATCATTAA AAATTTACGG GGAATTTTAC TATGCGAaCG AgcATGATCA AAAAAGGAGA	3720
35	TCACCAAGCA CCAGCAAGAA GTCTTTTACA TGCCACGGGC GCGCTAAAAA GTCCAACCTGA	3780
	TATGAACAAA CCATTTGTAG CTATTTGTAA CTCTTATATT GATATTGTTT CTGGACATGT	3840
	TCACCTTGAGA GAGCTTGCAG ATATAGCTAA AGAAGCAATT AGAGAAGCCG GTGCCATTCC	3900
40	ATTTGAATTC AATACAATTG GTGTTGATGA TGGAATAGCT ATGGGACATA TCGGAATGCG	3960
	ATATTCTCTA CCATCACGTG AAATTATTGC AGATGCAGCT GAAACTGTAA TTAACGCTCA	4020
	TTGGTTTGAC GCGTATTTT ACATTCCTAA TTGTGACAAG ATTACACCCG GTATGATTTT	4080
45	AGCAGCCATG AGGACAAACG TACCAGCTAT CTTTGTCTCT GGTGGACCAA TGAAAGCTGG	4140
	CTTATCTGCA CATGGAAAAG CATTAAACTT TTCATCAATG TTTGAAGCAG TCGGCGCATT	4200
	TAAAGAAGGA TCGATTTCTA AAGAAGAATT TTTAGATATG GAACAAAATG CCTGCCCTAC	4260
50	TTGTGGTTCA TGTGCTGGGA TGTTTACTGC AAATTCAATG AACTGTTTGA TGGAAGTTTT	4320
	AGGTCTAGCA TTACCATACA ACGGTACTGC ACTTGCAGTC AGTGATCAGC GACGAGAAAT	4380

55

	TATCGTTACT CGCGAAGCAA TTGATGATGC ATTTGCACTT GATATGGCTA TGGGTGGTTC	4500
	AACAAACACG GTACTGCATA CGTTAGCCAT TGCCAATGAA GCTGGTATTG ATTATGACTT	4560
5	AGAGCGCATT AATGCTATTG CCAAACGCAC GCCATATTTA TCAAAAATAG CACCTAGTTC	4620
	ATCGTATTCA ATGCATGATG TGCATGAAGC TGGTGGCGTC CCAGCAATTA TTAATGAATT	4680
	GATGAAGAAA GATGGCACGT TACACCCAGA TAGAATCACA GTTACTGGCA AAACGTTACG	4740
10	TGAAAATAAC GAAGGCAAAG AAATTAAGAA CTTTGATGTC ATTCACCCTC TTGATGCACC	4800
	ATATGATGCA CAAGGCGGTT TATCTATCTT ATTTGTAAT ATCGCCCCTA AAGGCGCAGT	4860
15	TATTAAAGTT GCGGCGTTG ATCCATCTAT CAAAACATTT ACTGGGAAAG CAATTTGTTT	4920
	CAATTCGCAT GATGAAGCTG TTGAAGCAAT AGACAATCGT ACCGTTCTGT CAGGCCACGT	4980
	CGTTGTCATT AGATATGAAG GACCTAAAGG TGGACCAGGT ATGCCTGAAA TGTTAGCACC	5040
20	TACTTCCTCT ATTGTTGGTC GCGGCTTAGG TAAAGATGTT GCATTAATTA CTGATGGGCG	5100
	TTTTTCCGGT GCCACAAGAG GTATTGCAGT TGGTCATATT TCCCCTGAAG CTGCATCTGG	5160
	TGGACCAATT GCCTTAATTG AAGATGGTGA TGAGATTACT ATTGATTTAA CAAATCGTAC	5220
25	ATTAAACGTA AACCAGCCTG AAGATGTTCT AGCGCGTCGC CGAGAATCTT TAACACCATT	5280
	TAAAGCGAAA GTAAAAACAG GTTATCTAGC TCGTTATACT GCCCTAGTAA CTAGCGCAAA	5340
	TACAGGTGGC GTCATGCAAG TCCCTGAGAA TTTAATTTAA TTTATTTTAA TATTGGAGAT	5400
30	GGTTAAATG TCTAAACTC AACATGAAGT AAACCAAAAT ATTGACCCTT TAAAAATGGC	5460
	TGAATCACTT GAACCTGAAC AACTAAATGA AAAAAGTTTA AATGATATGC GTTCAGGATC	5520
35	AGAAGTGCTA GTAGAAGCTC TACTTAAAGA AAATGTGGAT TATTATTTCG GTTATCCTGG	5580
	TGGTGCCGTA CTACCTTTAT ATGACACGTT TTATGATGGT AAAATCAAAC ATATTTTAGC	5640
	AAGACACGAA CAAGGTGCTG TTCATGCTGC AGAAGGTTAT GCACGTGTAT CTGGTAAamT	5700
40	GGCGTCGTTG TAGTTACAAG CGGTCCaGGT GCAACTAATG TAATGACAGG TATTACGGAT	5760
	GCACATTGCG ACTCTTTACC TCTAGTTGTA TTCACTGGAC AAGTTGCTAC ACCAGGCATT	5820
	GGTAAAGATG CATTCGAAGA AGCGGATATT CTATCTATGA CTTACCAAT TACAAAACAA	5880
45	AATTATCAAG TGAAACGTGT TGAAGATATC CCTAAAATCG TACACGAAGC TTTCCATGTA	5940
	GCTAATTCTG GACGCAAAGG TCCTGTAGTG ATTGATTTTC CAAAAGATAT GGGTGTTTTA	6000
	GCTACAAATG TGGATTTATG CGACGAAATC AATATTCCAG GTTATGAAGT TGTACAGAA	6060
50	CCAGAAAATA AAGACATTGA CACTTTCATC TCACTTTTAA AAGAAGCGAA AAAGCCTGTC	6120
	GTATTAGCCG GCGCAGGTAT TAATCAATCA AAATCAAATC AATTATTAAC ACAGTTTGTT	6180

55

	GATACACTAT TTTTAGGTAT GGGAGGAATG CATGGTTCTT ATGCTAGTAA CATGGCATT	6300
	ACTGAGTGTG ATTTACTCAT TAATTTAGGT AGCCGCTTCG ATGATAGATT AGCAAGCAAA	6360
5	CCTGATGCCT TTGCACCTAA CGCCAAAATT GTACATGTAG ATATTGATCC TTCAGAAATC	6420
	AATAAAGTTA TTCATGTAGA TTTAGGTATT ATTGCAGACT GTAAAAGATT TTTAGAATGT	6480
10	TTAAATGATA AAAATGTTGA GACTATAGAA CACAGTGA CT GGGTTAAACA TTGTCAAAAT	6540
	AATAAGCAGA AACACCCATT TAACTTGGT GAAGAAGATC AAGTATTTTG TAAGCCACAA	6600
	CAACAATCG AATATATCGG CAAAATTACA AATGGTGAAG CAATTGTTAC TACAGACGTG	6660
15	GGACAACATC AAATGTGGGC AGCTCAATTT TATCCATTTA AAAATCACGG ACAATGGGTT	6720
	ACAAGCGGTG GTTTAGGAAC AATGGGATTC GGTATTCCTT CGTCAATTGG TGCCAAATTA	6780
	GCTAATCCTG ATAAAACAGT CGTATGTTTC GTCGGTGACG GTGGTTTCCA AATGACAAAC	6840
20	CAAGAAATGG CACTTTTACC CGAATATGGT TTAGATGTCA AAATCGTACT AATCAATAAT	6900
	GGAACATTAG GTATGGTTAA ACAATGGCAA GATAAGTTCT TTAATCAACG CTTCTCACAC	6960
	TCAGTATTTA ATGGTCAACC TGATTTTATG AAAATGGCAG AAGCATATGG CGTCAAAGGT	7020
25	TTCTTAATCG ATAAGCCAGA ACAACTGGAA GAACAATTAG ATGCAGCGTT TGCTTATCAA	7080
	GGACCAGCTT TAATTGAGGT TCGTATTTCC CTTACTGAAG CTGTAACCCC AATGGTTCCG	7140
	AGTGGCAAAT CAAATCATGA AATGGAGGGC TTATAATGAC AAGAATTCTT AAATTACAAG	7200
30	TTGCGGATCA AGTCAGCAG CTAAATCGAA TTACAAGTGC TTTTGTTGCG CTACAATATA	7260
	ATATCGATAC ATTACATGTT ACACATTCTG AACAACTGG GATTCTAAC ATGGAAATTC	7320
35	AAGTCGATAT TCAAGATGAT ACATCACTTC ATATATTAAT TAAAAAATTA AAACAACAA	7380
	TTAATGTTTT AACGGTTGAA TGCTACGACC TTGTTGATAA CGAAGCTTAA TTTTAAGACA	7440
	AAGGCAATGA TGCCTAATT AGTTATAGAT ATATCATAGG CTGCTAGTTA ACATCTGCCA	7500
40	CTATTACAAA GTTATATTTT AGAATTTTCG AAACACAAAA TATTTAATTA TTTGGAGGAA	7560
	TTTATTATGA CAACAGTTTA TTATGATCAA GATGTAAAA CGGACGCTTT ACAAGGCAAA	7620
	AAAATTGCAG TAGTAGGTTA TGGATCACAA GGTACGCGC ATGCACAAAA CTTAAAAGAC	7680
45	AATGGATATG ATGTAGTCAT CGGCATTGCG CCAGGTCGTT CTTTGTACAA AGCTAAAGAA	7740
	GATGGATTTG ATGTGTTCCC TGTGTCAGAA GCAGTTAAGC AAGCTGATGT AATTATGGTG	7800
	CTATTACCTG ATGAAATTCA AGGTGATGTA TACAAAAACG AAATTGAACC AAATTTAGAA	7860
50	AAACATAATG CGCTTGCAAT TGCTCATGGC TTAAACATTC ATTTTGGTGT TATTCAACCA	7920
	CCAGCTGATG TTGATGTATT TTTAGTAGCT CCTAAAGGAC CGGGTCATTT AGTTAGACGT	7980

55

	CAAGCACGTA ATATTGCTTT AAGTTATGCA AAAGGTATTG GTGCAaCTCG TGCAGGTGTT	8100
	ATTGAAACAA CATTTAAAGA AGAACTGAG ACAGATTAT TTGGTGAACA AGCAGTACTT	8160
5	TGCGGTGGTG TATCGAAATT AATTCAAAGT GGCTTTGAAA CATTAGTAGA AGCGGGTTAT	8220
	CAACCAGAAT TAGCTTATTT TGAAGTATTA CATGAAATGA AATTAATCGT TGATTTGATG	8280
	TATGAAGGCG GTATGGAAAA TGTACGTTAC TCAATTTCAA ATACTGCTGA ATTTGGTGAC	8340
10	TATGTTTCAG GACCACGTGT TATCACACCA GATGTTAAAG AAAATATGAA AGCTGTATTA	8400
	ACTGATATCC AAAATGGTAA CTTCACTAAT CGCTTTATCG AAGACAATAA AAATGGATTG	8460
15	AAAGAATTTT ATAAATTACG CGAAGAACAA CATGGTCATC AAATTGAAAA AGTTGGTCGT	8520
	GAATTACGCG AAATGATGCC TTTTATTAAA TCTAAAAGCA TTGAAAAATA AGATAGACCT	8580
	ACAATGAGGA GTTGTTAAAT ATGAGTAGTC ATATTCAAAT TTTTGATACG AACTAAGAG	8640
20	ACGGTGaACA AACACCAGGA GTGAATTTTA CTTTGTATGA ACGCTTGCGT ATTGCATTGC	8700
	AATTAGAAAA ATGGGGTGTA GATGTTATTG AAGCTGGATT TCCTGCTTCA AGTACAGGTA	8760
	GCTTTAAATC TGTTCAGCA ATTGCACAAA CATTAACAAC AACGGCTGTA TGTGGTTTAG	8820
25	CTAGATGTAA AAAATCTGAC ATCGATGCTG TATATGAAGC AACAAAAGAT GCAGCGAagC	8880
	CGGTcGTGCA TGTTTTTATA GCAACATCAC CTATTCATCT TGAACATAAA CTTAAATGT	8940
	CTCAAGAAGA CGTTTTAGCA TCTATTAAAG AACATGTCAC ATACGCGAAA CAATTATTTG	9000
30	ACGTTGTTCA ATTTTCACCT GAAGATGCAA CGCGTACTGA ATTACCATTG TTAGTGAAAT	9060
	GTGTACAAAC TGCCGTTGAC GCTGGAGCTA CAGTTATTAA TATTCCTGAT ACAGTCGGCT	9120
35	ACAGTTACCA TGATGAATAT GCACATATTT TCAAAACCTT AACAGAATCT GTAACATCTT	9180
	CAAATGAAAT TATTTATAGT GCTCATTGCC ATGACGATTT AGGAATGGCT GTTTCAAATA	9240
	GTTTAGCTGC AATTGAAGGC GGTGCGAGAC GAATTGAAGG CACTGTAAAT GGTATTGGTG	9300
40	AACGAGCAGG TAATGCAGCA CTTGAAGAAG TCGCGCTTGC ACTATACGTT CGAAATGATC	9360
	ATTATGGTGC TCAAACTGCT CTTAATCTCG AAGAACTAA AAAACATCG GATTTAATTT	9420
	CAAGATATGC AGGTATTCGA GTGCCTAGAA ATAAAGCAAT TGTTGGCCAA AATGCATTTA	9480
45	GTCATGAATC AGGTATTCAC CAAGATGGCG TATTAAAACA TCGTGAAACA TATGAAATTA	9540
	TGACACCTCA ACTTGTTGGT GTAAGCACGA CTGAACTCC ATTAGGAAAA TTATCTGGTA	9600
	AACACGCCTT CTCAGAGAAG TTAAAAGCAT TAGGTTATGA CATTGATAAA GAAGCGCAAA	9660
50	TAGATTTATT TAAACAATTC AAGGCCATTG CGGACAAAAA GAAATCTGTT TCAGATAGAG	9720
	ATATTCATGC GATTATTCAA GGTTCTGAGC ATGAGCATCA AGCACTTTAT AAATTGAAAA	9780

55

	AAGAGGGTCA TATTTACCAG GATTCAAGTA TTGGTACTGG TTCAATCGTA GCAATTTACA	9900
	ATGCAGTTGA TCGTATTTTC CAGAAAGAAA CAGAATTAAT TGATTATCGT ATTAATTCTG	9960
5	TCACTGAAGG TACTGATGCC CAAGCAGAAG TACATGTAAA TTTATTGATT GAAGGTAAGA	10020
	CTGTCAATGG CTTTGGTATT GATCATGATA TTTTACAAGC CTCTTGTAAG GCATACGTAG	10080
10	AAGCACATGC TAAATTTGCA GCTGAAAATG TTGAGAAGGT AGGTAATTAA TTATGACTTA	10140
	TAACATTGTT GCCCTACCTG GTGATGGAAT CGGTCCAGAA ATTTTGAACG GATCTCTATC	10200
	ATTGCTTGAA ATTATAAGTA ATAAATATAA CTTTAATTAT CAAATAGAGC ACCACGAATT	10260
15	TGGTGGTGCC TCTATTGATA CATTCCGCCA GCCTTTAACT GAGAAAACCT TAAATGCGTG	10320
	TAAAAGAGCA GATGCTATTT TACTGGGTGC AATCGGTGGA CCTAAATGGA CAGATCCTAA	10380
	CAATCGACCA GAACAAGGAT TATTAAAATT GCGTAAATCC TTAAATTTAT TTGTAAATAT	10440
20	ACGCCCCACT ACCGTTGTCA AAGGCGCTAG TTCTTTATCA CTTTAAAGG AAGAACGCGT	10500
	TGAAGGCACA GATTTAGTTA TAGTCCGTGA ATTGACAAGT GGTATTTATT TTGGAGAACC	10560
	TAGACATTTT AATAATCAG AGGCCTTAGA TTCTCTTACT TATACAAGAG AAGAAATAGA	10620
25	ACGCATTGTT CACGTAGCAT TTAAATTGGC CGCTCAAGA CGAGGAAAAC TAACATCAGT	10680
	TGATAAAGAA AATGTATTAG CTTCTAGTAA ATTGTGGCGC AAAGTCGTAA ATGAAGTAAG	10740
	TCAATTATAT CCAGAAGTAA CAGTAAATCA CTTATTTGTT GATGCTTGTA GTATGCATTT	10800
30	AATCACAAAT CCAAAACAAT TTGACGTCAT CGTATGTGAA AACTTATTTG GCGATATTTT	10860
	AAGTGATGAA GCTTCAGTGA TTCCTGGTTC ACTTGGTTTA TCACCTTCTG CTAGTTTTAG	10920
35	TAACGATGGT CCAAGATTGT ATGAGCCTAT TCATGGATCA GCACCAGATA TTGCAGGTAA	10980
	AAACGTTGCC AATCCATTG GAATGATTCT ATCTTTAGCG ATGTGTTTAC GTGAAAGCTT	11040
	AAATCAACCA GATGCTGCAG ATGAATTAGA ACAACATATT TATAGCATGA TTGAACATGG	11100
40	GCAAACGACA GCAGATTTAG GCGGCAAATT GAATACTACT GATATTTTCG AAATTCTATC	11160
	TCAAAAATTG AATCACTAAG GGGGAGATGT AAATGGGTCA AACATTATTT GACAAGGTGT	11220
	GGAACAGACA TGTGTTATAC GGGAAATTGG GCGAACCGCA ACTATTATAC ATTGATTTAC	11280
45	ACCTTATACA TGAAGTTACT TCTCCTCAAG CATTTGAAGG ACTTAGGCTT CAAAACAGAA	11340
	AATTAAGACG CCCAGATTTA ACATTTGCAA CACTCGATCA CAATGTTTCT ACTATTGATA	11400
	TATTCAATAT TAAAGATGAA ATTGCAAACA AACAAATCAC AACATTACAA AAAAACGCCA	11460
50	TAGATTTTGG GGTGCATATT TTTGATATGG GTTCTGATGA ACAAGGTATT GTTCACATGG	11520
	TAGGACCTGA GACAGGACTT ACACAGCCTG GCAAGACAAT CGTTTGTGGT GACTCTCACA	11580

55

	ATGTTTTTCGC AACTCAAACG CTATGGCAAA CAAAACCCAA AAACCTTAAAA ATCGATATTA	11700
	ATGGTACCTT ACCAACAGGC GTCTATGCTA AGGACATTAT TCTGCATTTA ATTAAAACGT	11760
5	ATGGTGTGGA CTTTGGTACA GGCTATGCTT TGGAAATTTAC TGGCGAAACA ATTAAAAACC	11820
	TTTCAATGGA TGGTCGAATG ACTATTTGTA ACATGGCTAT CGAAGGTGGT GCCAAATACG	11880
	GCATAATCCA ACCTGATGAT ATAACATTG AATATGTTAA AGGGAGACCA TTTGCCGATA	11940
10	ACTTCGCTAA ATCAGTTGAT AAGTGGCGTG AGCTATATTC TGATGACGAC GCGATATTTG	12000
	ATCGTGTAAT TGAACCTGAT GTTTC AACAT TAGAACCACA AGTGACATGG GGAACATACT	12060
	CTGAAATGGG TGTTAATTTT AGTGAACCAT TCCCTGAAAT CAATGATATC AACGATCAAC	12120
15	GTGCGTATGA TTATATGGGG TTAGAACCAG GTCAAAAAGC TGAAGACATC GACTTAGGGT	12180
	ATGTTTTTCT CGGTTTCATGT ACAAATGCTA GACTATCAGA TTTGATTGAA GCTAGTCATA	12240
20	TTGTTAAAGG AAATAAAGTT CATCCAAATA TTACAGCTAT TGTCGTACCA GGTTCCTCGTA	12300
	CAGTAAAAAA AGAAGCAGAA AAATTAGGTC TAGATACTAT CTTTAAAAAT GCAGGATTTG	12360
	AATGGCGTGA ACCAGGATGT TCAATGTGTT TAGGCATGAA TCCTGACCAA GTACCTGAGG	12420
25	GCGTACATTG TGCATCTACA AGTAATCGAA ACTTTGAAGG ACGACAAGGC AAAGGTGCAA	12480
	GAACACATTT AGTATCCCCT GCTATGGCAG CAGCAGCAGC TATTCATGGT AAATTTGTGG	12540
	ACGTAAGAAA GGTGGTTGTT TAAATGGCAG CAATCAAACC TATTACAACA TATAAAGGTA	12600
30	AAATAGTCCC TCTCTTCAAC GACAATATCG ATACAGACCA AATCATTCTT AAGGTACACT	12660
	TAAAGCGTAT TTCAAAAAGT GGCTTTGGTC CATTTCCTTT TGATGAATGG CGGTACTTAC	12720
	CTGATGGTTC AGATAATCCT GATTTCAATC CTAACAAACC ACAATATAAA GGGGCTTCTA	12780
35	TTTTAATTAC TGGAGATAAT TTTGGATGTG GTTCAAGTCG TGAACATGCT GCTTGGGCTC	12840
	TTAAGGACTA TGGTTTTTCAT ATTATTATTG CAGGAAGTTT CAGTGACATA TTTTATATGA	12900
40	ATTGCACTAA AAATGCGATG TTGCCTATCG TTTTAGAAAA AAGTGCCCGT GAACATCTTG	12960
	CACAATATGT TGAAATTGAG GTCGATTAC CAAATCAAAC TGTGTCATCA CCAGACAAGC	13020
	GTTTCCATTT TGAAATTGAT GAAACTTGGA AGAATAAACT TGTAAATGGC TTAGATGACA	13080
45	TTGCAATCAC CCTACAATAT GAATCATTAA TAGAAAAATA TGAAAAATCa CTTTAAGGGA	13140
	GTTGAATATT ATGACAGTCA AAACAACAGT TTCTACGAAA GATATCGATG AGGCATTTTT	13200
	AAGACTTAAA GATATTGTCA AAGAAACACC TTTACAATTA GACCATTACT TATCTCAAAA	13260
50	GTATGATTGT AAAGTCTATT TAAAACGAGA AGATTACAA TGGGTACGTT CTTTTAAATT	13320
	AAGAGGTGCT TACAACGCTA TTTCTGTTTT ATCAGATGAA GCTAAAAGTA AAGGTATTAC	13380

55

	AAACGCTGTT ATCTTTATGC CAGTCACTAC ACCTTTACAA AAGGTAAATC AAGTAAAGTT	13500
	CTTTGGAAAT AGTAACGTTG AAGTTGTACT CACTGGTGAT ACATTTGATC ACTGTTTAGC	13560
5	TGAAGCTTTA ACTTATACAA GTGAACATCA AATGAACTTT ATAGATCCAT TCAATAATGT	13620
	TCATACAATT TCTGGACAAG GTACGCTTGC TAAAGAAATG CTAGAACAAG CAAAGTCTGA	13680
	CAATGTTAAC TTTGATTATC TATTTGCCGC AATTGGTGGT GGCGGTTTAA TTTCAAGTAT	13740
10	TAGTACTTAC TTTAAAACCT ATTCACCTAC CACGAAAATT ATAGGTGTTG AACCTTCAGG	13800
	TGCAAGTAGT ATGTATGAAT CTGTTGTGGT AAATAATCAG GTAGTCACAT TGCCTAATAT	13860
	CGATAAATTT GTGGACGGTG CATCTGTAGC TAGAGTTGGC GATATTACAT TTGAAATTGC	13920
15	AAAAGAAAAT GTAGATGATT ACGTTCAAGT AGATGAAGGT GCAGTTTGTT CTACGATTTT	13980
	AGATATGTAT TCAAAACAAG CAATTGTAGC AGAACCTGCT GGCGCATTAA GTGTAAGTGC	14040
20	GCTTGAAAAC TATAAAGATC ATATTAAAGG TAAAACAGTG GTTTGTGTCA TTAGTGGTGG	14100
	TAATAATGAT ATTAATCGAA TGAAAGAAAT TGAAGAACGT TCATTACTAT ACGAAGAAAT	14160
	GAAGCATTAC TTTATCTTAA ATTTCCCTCA ACGTCCAGGT GCATTGAGAG AATTTGTAAA	14220
25	TGACGTATTA GGACCTCAAG ACGATATTAC TAAATTTGAA TACTTAAAAA AATCTTCTCA	14280
	AAATACAGGT ACTGTCATTA TTGGTATTCA ACTTAAAGAT CATGATGATT TAATACAAC	14340
	CAAACAACGT GTAAATCATT TCGATCCTTC CAATATTTAT ATTAATGAAA ATAAGATGTT	14400
30	ATATTCATTG TTAATTTAAC ACATAGTAAG AAAAACAGTC ATAAATTGAT TTCTAATTGA	14460
	AATCATCTTA TGACTGCTTT TTATTATACT TTACATTTCT CGTTTCGTCA GATTCAAACG	14520
	TTTTCACTTC GCCAAGCCAT CTTTCTTTGT GTTTGCTTTT aTTTTGACGT TTTAGACATA	14580
35	AAAAAaGAGA CCTTGCGGTC TCAATGCGGC TCATCGCATC CACTTTTTGC CTGGCAACGT	14640
	TCTACTCTAG CGGAACGTAA GTTCGaCTAC CATCGACGCT AAGGAGCTTA ACTTCTGTGT	14700
40	TCGGCATGGG AACAGGTGTG ACCTCCTTGC TATAGTCACC AGACATATGA ATGTAATTTA	14760
	TACATTCAAA ACTAGATAGT AAGTAAAAGT GATTTTGCTT CGCAAAACAT TTATTTTGAT	14820
	TAAGTCTTCG ATCGATTAGT ATTCGTCAGC TCCACATGTC ACCATGCTTC CACCTCGAAC	14880
45	CTATTAACCT CATCATCTTT GAGGGATCTT ATAACCGAAG TTGGGAAATC TCATCTTGAG	14940
	GGGGGCTTCA TGCTTAGATG CTTTCAGCAC TTATCCCGTC CACACATAGC TACCCAGCTA	15000
	TGCCGTTGGC ACGACAACCTG GTACACCAGA GGTATGTCCA TCCCGGTCCT CTCGTACTAA	15060
50	GGACAGCTCC TCTCAAATTT CCTACGCCCA CGACGGATAG GGACCGAACT GTCTCACGAC	15120
	GTTCTGAACC CAGCTCGCGT ACCGCTTTaA TGGGCGAACA GCCCAACCCT TGGGACCGAC	15180

55

GTGGAACTT

15249

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14051 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

5	GTGGCAATAT TTCTAGTTCT CGTTTTGATA AGATTTTAAA AGGATCTGTT GTGTTTGCAG	60
15	TGTCCTGATT TGAATTAGAT ACAAATTCAT TCACTAAAGA TGTTGTAAGT TTCATATCTA	120
	CATATGTTTC ACCTTTATAT ACAGTTCGAA TAGCTAACAA TAATTGTTCA TCAGGTGCAT	180
20	TTTTCAATAT GTAACCTTTC GCACCATTAC GCAACACATG GAACAAATAC TCCTCATCAT	240
	CAACATTGT TAATATTAGT ATTTTAGTTT CAGGAAAAC TGCAGCAATT TTA CTG TAG	300
	CGATAAGACC TGACTCACCT GGTGGC ATAC TTAATCCAT TAGTAACACA TCAGGTTTAT	360
25	ATTCATTAC TTTTGGTAA GCTTCGACGC CATCTGCAGC CGTTGCAACA ACTTCCATAT	420
	CATTTTGATA ATTTAAATC ATAGAGAACC CCGTACGGAC AACAGCGTGA TCATCGGCAA	480
	TGACTATTTT CAATTTTATT CCCCCAATGT ATGTTTCAAA TTGGAATGTT CAATGTAACA	540
30	TTGGTACCCT CACCAATTTT CGTTTCAATA TTGACGCTAC CGCTGACTAA CTCAGCTCGC	600
	TCATTCAATC CATATAAACC GAGTCCAGAA CCTTTAGGCT TAGAACTTGG ATCAAAACCA	660
35	TTTCCCGCAT CTATCACTTC TGCTACCAAA TGGCGCCCAG TTTGACGGAT ACCTACATTT	720
	ATTTCAATTA CATCAGCGTA TTTCAACGCA TTTAAATAG CTTCTTGCAC TACTCGATAA	780
	ACAACCGTTT CAATATCACT ATCAAAGCGA GTATTTTAA TATTTGATGT ATATATGATT	840
40	TTTATTCAT AATTTTCTTC AAAGTGTTA AAATATGATT TAAAAGCTGC TTCAAGGCCT	900
	AGATCATCCA AAGAAGCGGG TCTTAATTCA ACCGACATAT TACGTATATC ATCAATTAAT	960
	TTAGCGACAA TATATTCAAT ATTTTCTGCG TCTTCCAAA GCTTAGTTGT ATCTTCTTGA	1020
45	TATTTTAATA ATCTCAATTG AACATCTACA TTGAGCATTT CTTGAATCAC ACTATCATGT	1080
	AACTCTCTAG AAATTCGCTT TCTTTCATTT TCTTGGGCTG AGATTGTTTT ACGCATCATA	1140
	CGTTGTTGAT GCAATTTCTC TTGCTGTTCA ATTTGTGATG AAACATTTTG AAGCGTAAAT	1200
50	GCATGAATTC CCCTGTCTTG ATCAATCAAC TGATATGTTG CTGTAAATGG CATCACTTTT	1260
	TGATCTTTCG TCTTCATAAA TACTTGGAAG TTCGTAGCTT GTACTTGCAT CGATTCTAAG	1320

	ATCGCATTCTG CCACAGCACT GTAATTATCT TCTTCAGATA ATATATCTTT AGCAGCATCA	1440
	TTTATTGCAA TAATTTTACC GTTATCATCA GCAAAACTA TCTTTTCGAT TGAATGCTCA	1500
5	TAATATTTTT TCAATAAGT ATCTAACTGT ATACTGTCCT CATTATCAT GACTTACACC	1560
	CTAATTCATC TCATTATTTA TCATCATTGA AAATACCAA CTTACGTTGA ATATCATCAT	1620
	TATCAAATAT TTTTGGTAAA GGACGACCAT CTCTTTGACC AAATAATAGT ACGCCATACA	1680
10	CTTGATTCTT ATACCAAAGC GGCAGTCTA AAAGTCTGT TAATGATTCTG CTCAATAAAA	1740
	TTGGATAGTC AATCTTTTCT TCAGGCCCTA AAGCTAAACC AACATTGGCT ATTACCATAC	1800
	GCTTTCCTGT TTTTATAACA GTTCCAGCTA ATCCACGACC TTTTCTTAAA ATAATCAATT	1860
15	TAAATCGATT ATTTTTATTA CCTGAAACAT AGTGCCATT TATTGGAGAT GATGGTTTGT	1920
	TAGATTCATA GAAAGCGATT GCCGCAAAAT CATAACCCTC TTCTTTGCGT ATTTTATCTA	1980
20	ATGTCTCTTG AAATCTACGA TCTTCAATTA TTGCTTCTGG TGTCAAATCC TTTACCTCT	2040
	TATGCTTACA CTTTATTCTT ACGGTAAATA ATATATCTGC GATTTATATA TGTCAAAGGT	2100
	ACACTCCAAA CATGCACCAA ACGTGTAAT GGCCAACAAG CCATAATAGT GAAACCTAAC	2160
25	AATATATGCA TTTTAAATGC AATCGGCACA CCACTCATCA ATGACGCATC TGGTTTTAAC	2220
	ATAAATAATT GTCTAAACCA AATTGATAAT GAAGTTCTGT AGTTAAAGTC TGGATGTTGT	2280
	ATATTTGTTA CTAATGTTGC GTAACATCCC ATAAATACGA TAAGTAATAA TAAGAAATTT	2340
30	ACAAATATAT CCGACGCTGA ACTTAATCTT CGAATACTTT TCGTAGTAAC ACGTCTCGCT	2400
	GTTAATAAAA ACATCCCTAT CAAAGTTATT ATACCAAAGA TGCTACCAAT ATAAACAGCG	2460
	CCTATATGAT ATAAATGCTC AGACACACCC ACTGCATCCA TCCATGGTTT CGGTATTAAC	2520
35	AATCCAACTA CGTGTCCAAA AAACACTGGA ATAATACCTA AGTGAAATAA TAACTTCCC	2580
	CACATCAACC TTTTCTTTC TATTAATTCA CTAGATTAG CTGTCCAAGA AAATTTATCA	2640
40	TAACGATAAC GTGCAATATG ACCTGCGACA AAGACAACTA AACATAAATA CGGAAATATA	2700
	ACCCATAAAA ACTGATTAAG CATGATGTTT CACTCCTTTT GGTGATGTCA AACATAATTT	2760
	CAATGTTTTT CTAAGTGCTT GAATCACATA GGCATATGGA TTGTTATCTT CACCAAGTGC	2820
45	ATTGCCCATC ACATATGTTT CATCCTCAAT AATCATAATG ATTAATTGAA TATTCTCTTC	2880
	AGCTCTTGA TCATTTGCCC ATTCTGCCAC TTGCAAAAAT TGAAGCATCA ACGGTAGATA	2940
	ATCAGAAAGT TCATTATCTA CCATTCTAG TCCAAACATT TCATATAATA CCTTTAATTT	3000
50	AGCTAACATT TGCCACGTT CTTTTGCGT ATCAAATTTG TTATACGTCA TATATAATGG	3060
	TGCTTTTTTC GTAAATCAA ATGTATCTGT ATAAATCGCT TTGATTTCTG ATAATGAAAA	3120

55

	TGTTTCTTCA AAAGTTTTTG GATGAAAAGT TAATTTTTCT GGAAACATA ACTGTTGTGC	3240
	CATATATCCA AAACTTTCTT GATATTTTTT AAAATTATCG AAATTAATCA CGGAAAATCC	3300
5	CTCCATAGAA ATTCTCATT AAAATTTCTT GACCAGTTTT CCCTGAACCT ACTGCAACGC	3360
	CACAGCCTTC ACAGTTATCT CCAAATGCT CGCCGCCGTA ATTGTATCCT GTACTACCTT	3420
	GTGCGTGATA CGTATCTAAA TAGGTTTCTT TGTGTGATGT TGGAATAACA AATCGATCTT	3480
10	CATATTTGGC TAGTCCTAAT AAACGATACA TGTCTTTAGT TTGGCGCTCG GTTATACCTA	3540
	ATCGCTCTAA TCGAGACGTG TCAAATGGCT GTTGAGTAAC TTGAGATCTC ATATAACTTC	3600
15	TCATCATTGC CATACGTTGT AGGGCTCCTT TTAGTGGCTC TGTATCTCCT GCAGTGAAAA	3660
	TATTAGCTAA GTATTCAATA GGTAAACGCA TTTCTTCAAT GGCTGGGAAA ATCGCATCTG	3720
	GATTTTGAGT TGTATTTTTA CCTTCAAAAT AGCTCATAAT TGGGCTAAGT GGTGGGCAAT	3780
20	ACCAAACCAT CGGCATCGTT CTAAATTCAG GATGTAACGG AAATGCAAGT TTATATTCAA	3840
	TTGCTAACTT ATAAATTGGA GAGTTTTGTG CAGCTTCAAT CCAATCGTAA CCAATACCAT	3900
	CTTTTTCAGC TTGAGCAATG ACTTCTTCGT CAAATGGGTT TAAGAATATA TCTAATTGTT	3960
25	TTTCATATAA ATCTTTCTCG TCTACTGCTG AAGCTGCTTC ATGAACTCGA TCTGCATCAT	4020
	ATAATAAAAC ACCTAAGTAA CGCATACGTC CTGTACAAGT TTCAGAGCAT ACCGTAGGCA	4080
	TACCCGCCCTC GATTCTCGGG AAACAGAAAAG TACACTTTTC AGCTTTGTTC GTTTTCCAAT	4140
30	TGAAGTAAAC TTTCTTATAT GGACAACCTG TCATACAGTA ACGCCATCCA CGACATGCGT	4200
	CTTGGTCAAC TAATACAATG CCATCTTCAT CACGTTTATA CATAGCACCT GAAGGACACG	4260
35	ATGCAACGCA ACTTGGATTC AAGCAATGTT CACATAAAGC TGGTAAATAC ATCATAAAAG	4320
	TTTCGTCAAA TTGGAATTTA ATATCTTCTT CTATTTTTTG GATGTTAGGA TCTTTTGGAC	4380
	CTGTAAACATG ACCACCTGCT AAGTCATCTT CCCAGTTAGG TCCCCATTCA ATTTCAATGT	4440
40	TATCCCCCGT AATTTCTGAA TACGCTCTAG CAACTGGCGA ATGCTTCCCT GATTTGCGAG	4500
	TTGTTAAATG TTCATAATTA TAGTTCCATG GCTCATAATA ATCTTTAATT AATGGCATAT	4560
	CTGGGTTATA AAAAATTTTA CCTAAAGCAA TTTTGAAT TCTACTCCA GATTTTAATT	4620
45	CAAGTTTCCC TTTACGATTT AGTACCCAAC CACCTTTGTA GTGTTCTTGG TCTTCCCAAC	4680
	GTTTCGGATA CCCTACACCT GGCTTCGTTT CTACGTTGTT GAACCACATG TACTCAGCAC	4740
	CTGGACGATT TGTCCaAGTG TTTTACATG TCACACTACA CGTATGGCAT CCTATGCATT	4800
50	TATCTAAATT TAATACCATC GCAAcTTGCG CTTTAATCTT CAAGCCAATT AACCTCCTTC	4860
	ATCTTTCTAA CTGCTACATA TAAATCCCTT TGGTTCCCAA TTGGTCCATA ATAATTAAAG	4920
55		

	GGCGCGTTGT GTGAACCACC ACGTGTATCT GTAATTTCTG ACCCAGGCGT TTGAATATGT	5040
	TTATCTTGTG CATGATACAT AAACATTGTA CCTTTAGGCA TACGATGCGA AATAACTGCT	5100
5	CTTGCCGTTA CAACACCATT ACGGTTATAC ACTTCTAGCC AATCATTATC TTGGATATCG	5160
	TGTTTTTCAG CATCTTCATT TGATATCCAA ACCGTTGGAC CACCTCTAAA TAGTGTCAAC	5220
	ATATGCTTAT TATCTTGATA CATTGAGTGT ATATTCCATT TTCCATGAGG CGTTAAATAA	5280
10	CGCagTACCA AAGCATCTGT ACCACCTTTA ATTTTCTTAT CTCTATTCCC AAATACCATT	5340
	GGCGGCAATG TCGGTTTATA TACTGGTAAG CTCTCCCCAA ATTGTTGGAA AACTTCGTGA	5400
	TCCACATAAT AACTTTGACG TCCTGTTAAT GTTCTAAAAG GTACTAGACG TTCTATATTC	5460
15	GTTGTAAATG GTGAATATCG TCGACCTTGT TTATTTGAAC CTGGGAATAC TGCTGTCCGT	5520
	ATTACTTCTC GTGGTTGTGA AGTTATATTT AAAAACGAAA TTTTCTCAGC AGCGCGTTCTG	5580
20	CTAGAAATAT CTTTTAACGG CATTCCAGTT TGTCTTCGA GATCTTCATA TGATTTTTGT	5640
	GATAATTTAC CATTCGTAGC AGATGAAATA CTTAGTATTG CATCAGCTAC ATTACGTGCT	5700
	GTATCAATAC GTGGACGATT CGCTCTCACA GAATCATCAT TTGTATCACT CCACGTACCT	5760
25	AACATACTTT TTAATTCTTC ATATTGTTCA CTGACACCGA AACTTACACC ATGTGCTCCA	5820
	ACTTTCCCTT TTTCAAGTAC AGGACCAAGC GTGACATATT TGTCGTAAAT TTTAGTGTAG	5880
	TCGCGTTCTA CAATTGCAAA GTTAGGCATT GTACGTCCAG GTACCGCTTC AATTTACCCC	5940
30	TTCGACCAAT CTTTCACTAC GCCGTATGGT GTTGAAATTT CTTGCTTTGT ATCATGACTA	6000
	AGTGAGTTG TCACAACATC TTTAAACGTT CCAGGTAAAT AGTCTTTTGC CATTTCTGAA	6060
	AATGCTTTTG CCAACGTTTT ATAAATATCC CAGTCTGAAC GCGATTCCCA TAACGGATCA	6120
35	ATGGCAGGAT TGAAAGGATG TACATATGGA TGCATATCCG TTGATGATAA ATCATGTTTT	6180
	TCATACCAAG TCGCTGCCGG CAAAACAATG TCAGAATATA ACGGTGTTGC CGTCATTCTG	6240
40	AAGTCTAAAG AGACCACTAA ATCTAACTTA CCTGTTGTTT CTTACGCCA CGTAATTTCT	6300
	TCTGGCTTTT CATCTTCATT TGGTGTAGCT AATAACCCTG ATTTTGTGCC AAGTAAATGC	6360
	TTCATAAAGT ATTCTTGACC TTTTGCAGAA CTTGAAATTA AGTTTGAACG CCATATAAAT	6420
45	AATGATTTTG GATGATTCTT TTTCAAATCA GGATCTTCTA TTGCAAATG TGTGTTTTT	6480
	GATTTCACTT CATCAATTGC ACGTTGCAAA ATCGCTTCAT TTGAATCTAT ACCTTCATCT	6540
	TTAGCTTCTT CTGCAAACAA CAACTATTT TTATTAAATT GTGGATATGA TGGTAACCAA	6600
50	CCAAGTCTAG CTGCTAAAC ATTATAATCA GCTGGATGTT GATGCTTTAA CTCCTCTGTT	6660
	TTAGCTAATG GAGATTTTAA ACGATCTACA TTTGACTCTT CATATTTCCA TTGGTCTGTT	6720

55

	AATGCGACAG TACTCCATCC TTCAATCGGA CGACATTTTT CTTGTCCCAC ATAGTGAGCC	6840
	CAACCGCCAC CATTACACACC TTGACAGCCA CATAACATAA CTAAGTTTAA GATTGAACGA	6900
5	TAAATCGTAT CTGAGTTAAA CCAATGGTTA ATACCCGCAC CCATGATAAT CATTGAACGC	6960
	CCTTCAGTAT CGATAGCGTT TTGCGCAAAT TCTTTCGCTA CTTGAATGAC AACACTTTGT	7020
	TTTACGCCTG AAATGGCTTC TTGCCAAGCA GGTGTATATT TTGATTCTGC ATCGTCGTAT	7080
10	CCTTTTGATT CTAATTTATG ATCAAAACGA CGCAGCCAT ATTGACTTGC CATTAAGTCA	7140
	AAAATTGTAG CAATACGGAC TTTGTCACCA TTTGCTAAAG TGACTTGTCG AGTTGGAATT	7200
15	GGACGATTGA ATATCCCATC TCCATCACTA TCAAAGTATG GGAATTGAAT TGTTCCTAAT	7260
	TCGTATCCAC CTTCTGTCTG TGATAATGTA GGGTTAATTT TAGAACCATC TTCTGTTTCT	7320
	AGTTTTAAGT TCCACTTCTT ACCTTCTTCC CAACGTTGAC CCATTGTGCC ATTAGGTACT	7380
20	ACTAACTAT CGCTGATTGC ATCATGAATA ACTGGCTTCC ATTCGCCTTG CTCGTGTTGT	7440
	TGACCTAAGT CACTCGCTCT TAAAAATCGA CCCGCTTTAT ATCCATTTTC ATCTTCATCC	7500
	AGCATGATAA GAAACGGCAT ATCTGTATAT TGTTTAGCGT AATTTATAAA GCGTTCATTA	7560
25	GGTTGATTAA CATAATGTTT TTGTAAAATA ACATGCGTCA TTGCTTGTGC AATTGCAGCA	7620
	TCTGAACCAG GATTCGGTGC TAGCCAGTTA TCTGCAAATT TCACATTTTC TCGTAACTCT	7680
	GGTGCTACTG AAATGACTTT TGTACCTTTA TAGCGGACTT CAGTCATAAA ATGTGCATCC	7740
30	GGAGTACGTG TTAAAGGTAC ATTAGAGCCC CACATAATAA TGTATGATGC GTTATACCAG	7800
	TCATTGATT CAGGCACATC TGTGCTCTT CCCCAAATTT GTGGAGAGGC AGGTGGTAAA	7860
35	TCTGCATACC AGTCATAAAA ACTAAGCATT TCACCACCAA GCAAATTGAT GAATCGAGCA	7920
	CCTGCTGCAT AACTAATCAT TGACATCGCT GGAATAGGTG TAAATCCTGC GATTCGATCT	7980
	GGACCATATT TTTTATTGT ATACAGTAAT TGTGCTGCGA TTATCTCTGT AACGTCTTTC	8040
40	CAATTTGAAC GCACGTGCCC TCCCATACCT CGGGCTTGCT TATATTGTTT GGCTTTGTCT	8100
	TCATTTTCAA CAATAGACGC CCATGCAGCA ACGCGATTAC CATTGTTTTTCT TCTAATGCT	8160
	TCAGTCCATA AATCCCAGAG TTTCCACGA ATATATGGAT ATTTGATTCTG AAGCGGACTG	8220
45	TATTCATACC AAGAGAATGA CGCACCTCGT GGACATCCTC TCGGTTTATA TTCAGGCATA	8280
	TCCGGACCAC AACTTGGATA GTCAGTTTGT TGATTTTCCC AGGTAATCAC ACCATTTTTTCT	8340
	ACAAATACTT TCCAAGAACA TGAGCCTGTA CAGTTAACAC CATGTGTTGT TCTTACTTCT	8400
50	TTATCGTGGC TCCAACGTTT TCTGTACATT TTTTCCATT CTCTACTTTT ACTTTCTAGG	8460
	ATCGACCAAT TCCCATTAAT TTTTCTGTT GGCTTAAAGA AATTCATCC AAATTTTCCC	8520

55

	TAAAATGCCC AAGACTATTG CTTTAATTAG ATTGTACATT TTTTCACAAA CATAAAATAT	8640
	TAGGGAATCA CCTAATTACT TAAGGAATTT CCCTATCAAT AACGGGATTT CATTGAAATA	8700
5	ATACACAATC ATGTATGGTC ATGCTTATTG CCAATCTAAA TCGTTCAAAT TTGGCACAAC	8760
	GACAAATAAG GCTTCAACAC GAATATATTC TCTCGGTTGA AACCTTACTT ATTCATTTAT	8820
	TTTTTATAAA TTAGTGACAT AACACTGTAT TAGCATCTGC ACGATCGGTT GAAATATATG	8880
10	TTACATTTTC TTGCTGCTTA ATAAATGCAT CATAGTAATC ATATTGCGAC GAATGATATG	8940
	TGCCATTGCA TGTATCAITTT GGGTTTAGCA AACAGCCATA ACCTTCGTCA TATAAATGTT	9000
	CACAGAGCAT AAGGGCGTCA TGTTTAGAAC CACTTACTAC ATAAAATTGC TTCATAGGAT	9060
15	CATATGATTT AGGAGTGTTT TCAGTATAAT CAACRACTTC CCCTATAATA CATATACCTG	9120
	GTTTCGCCTC AATTGAATAG TGTTCGAATT TTGAAATAAT ATTACTTAAA CGCCCCTTAA	9180
20	CAACAACTC GTTAAAACAC GATGCTTGAA AGACAATCGC TATCGGGTAA TCAATATCTG	9240
	TGTATTGTTG TATCTGTGTG ATAATTTTCC CTAAACGTTT TACCCCCATA TAAATTGCTA	9300
	ACGTGCCACC ATTCACTAAG GAATTGACAT CCACTTCATT TTCTTCTGAA TCTTTAAAGT	9360
25	GACCTGTAGA AAATGTCACA CTTTTAGCAA CTGTACGCAT TGTCAAACCT GTCTGCATAG	9420
	TAGCAACTGc tGCGCTCGCT GATGTCACCC CTGGTACAAT TTCAAACGCA ATATGATGTT	9480
	CATTTAGTAT GTCGACTTCT TCTTGACACAC GACCAAATAT CGCTGGATCG CCACCTTTAA	9540
30	GTCTAACAAAC CTTGTTATAT CGACGCGCTG CTTCCACGAT ACAGTCATTT ATTTTTTCTT	9600
	GCTGAATATG TTTTGCATAC GGCTTTTTAC CAACATCGAT AATTCAGTA GTCAAATTCG	9660
	CATATTGTAA AATTAAACGA TTCACTAATC GATCATATAG AATGACATCC gCTTCACGTA	9720
35	TTAAACGCTC AGCCTTTTTT GTCAAATAAT TCGGATTACC TGGACCCGCA CCTATCAAGT	9780
	AAACCTTGCC ATATTCCTCT ACAGACATAT ATATACGTTT CCGTCTGTAA CTTCTACCTC	9840
40	ATAAACATCT ACACAACCTT CATCAGGTTT TTGAACAATA CCTGTATTTA AATCAATTTT	9900
	TTGATCGTGG AGCGGGCAAA ATACATATTC CCCACTCACT GTCCCTTCAG ACAATGGTCC	9960
	TTGTTTGTGT GGACAGATAT TGTGAATCGC ATGAATTTTG CCACTTTCTG TTAaaaaaCA	10020
45	CCCTACCTCT TTGCCTTTGA CAATAACCTT TTTTCCAATT AGGGGTGTTA ATTCATCTAT	10080
	AGTTGTCACT TTAATTTTTT CTTTTGTTTC CATGTATTAC ACCTTCTCCA CTTCAAAAAT	10140
	TCTACGTGCT TGAGCATTGC TAGTTATTGC TTCCAAGGT TCAGCTTCGA CTGCTTTTTT	10200
50	AGCATCCATA ATGCGTTCAA ATAGTTCATT TTGTCTTTCT GGGTCAAGTA AGACTTCTTT	10260
	TACATTTTCA AATCCAAGTC TTCTTAACCA TGCGCGTGTT CTTTCAGCAT ATATACCTGT	10320

55

	AGTTGTTAAA AATTCAGCTT TTTCAACTTC TGTACCACCA TTACCACCGA TATAGATTTG	10440
	GAATCCATTT TCAACTGAGA TAATACCAAA ATCTTTAACA CCTGATTCAA CACAACCTCT	10500
5	TGGGCAGCCT GATACACCCA TTTTGAATTT ATGAGGTGTA TCGATGTATT CAAATGTTTT	10560
	TTCTAAACGA ATGCCAAGTC GTGTCGTGTA TTGCGTACCA AATCGACAAA ACTCTTTACC	10620
	AACACAGCTT TTAAGTGAAG GTGTTTTCTT ACCATAAGCT GATGCTGAAC GCATACCTAG	10680
10	GTCTTCCCAT ATATTTGGTA ATTCTTCTTT TTTAACTCCA TACAAACCAA CACGTTGTGA	10740
	ACCTGTCACT TTAAGTAGTG GCACATGATA TTTCTTAGCC ACTTCTCCTA GACGAATCAG	10800
	TTGGTCTGCA TCTGTAACAC CCCACGCGAT TTGAGGTATA ACAGAAAATG TACCATCATT	10860
15	TTGAATATTC GCATGGTAAC GTTCGTTAGC AAATCTTGAT TCTCTTTCAT CTTCATGATC	10920
	ATGTGGATAA ACCATGTTTA AATAATAGTT GATTGCTGGT CGACATTTTG GACATCCACC	10980
20	TTTATTTTTA AAGTTTAAAA CATGTCGAAC TTCTTTAGAT GTTTTTAAAC CTTTCGCTCT	11040
	TATTTGCGTT ACTATTTGAT CGCGTGTCAA ATCAGTACAA CCACATATAC CAGCAGGTTT	11100
	TGCGGCAACA AAGTCATCTC CTAAGGTGTG CTGCAATATT TGAGCAATTT GCGGTTTACA	11160
25	TTTACCACAT GAATTCCCCG CTTTTGTTTT AGCCGTTACT TCTTCAACTG TTGTAAAGCC	11220
	ATTTTCCGTA ATCGCATTTA CTATAGTACC TTTATCAACA CCATTACAAC CACAAATTGT	11280
	TTCATCATCA GCCATATCAG CAATTGATAG CGATGCCTCT TCTCCACCTT TAGTAAGCAA	11340
30	TGATACAAGT GTGTAATCTT CAGTGGATTC ACCTTTTTTC ATCATGTTAT AAAAGCGTGA	11400
	ACCATCATCG ATATCACCAT ATAGTACTGC ACCAACTACA TTACCGTCTT TTAAGAAAGAT	11460
	TTTTTTATAG TTATTATCAA CACTATTAAA TATTTCAATA CCTTTAATTT CTGCATTTTC	11520
35	TACAATTTGA CCAGCACTAT ACAAGTCACA CCCAGAACT TTTAATGACG TAAATGTTGT	11580
	TGATCCCTTG TATCCGTTCT TTTCTTTATT TGTTAAATGA TCAGCTAATA CTTTACCTTG	11640
40	TTCATATAGT GGTGCAACGA GTCCATAAAC TTTGCCGTTA TGTTCTGCAC ATTCACCAAC	11700
	TGCATATACA TTGCTATCAC TTGTTTGCAT CACATCATTG ACAACAATAC CACGATTAAC	11760
	ATCTAGACCT GATTCTTTGG CTACTTCTGT GTATGGTCGT ATACCTACTG CCATAACAAC	11820
45	TAAGTCTGCC GGAATCTCGC GTCCATCAGC CAATTTAACA CCTCAACAT CATCTTCTCC	11880
	TAAGATTTCA GTTGTGTTGG CTTGCATTTT AAACCTCATA CCTTGCTTTT CTAGATCTGC	11940
	TTTAAGCATA TTTCCAGCTT TACGGTCTAG TTGCATTTCC ATCAACCATT CAGCTAAATG	12000
50	TAAACACGTT ACTTCCATAC CTTGATCTAA TAAACCACGT GCACACTCTA AACCTAGTAA	12060
	TCCTCCACCA ATTACAATTG CTTTCTTTTT AGTCTTAGCA ATGTTTCATCA TTTGTTTCAGT	12120
55		

EP 0 786 519 A2

	GAATGCTTTA GAACCTGTCC CAAAAATCAA TTTATCGTAT GATACTTCAA TACCATTTCG	12240
	AGTAGTAACT GATTGATTG CTCTATCTAC TTCAATTACA GGATCATTTG TAATTAACTC	12300
5	GATACCATGT TCCTCATACC ACTCATATGG ATTCATAATT GTTTCTTCAA CTGTCATTTT	12360
	ATTTTGTAAA ATATTTGAAA GCATGATGCG GTTATAGTTT GGATAAGGTT CTTTACCTAT	12420
	TACCGTAATA TCATATAAAT CGTTGGCGCG CTCTAATATT TCTTCGATTG TTCGAATGCC	12480
10	CGCCATACCG TTACCAATCA TTACTAGTTT TTGCTTTGCC ATAAAATATG CCCCTTTACT	12540
	CCATAATATT TATTTCAAAA AAAGGTATTA ATTTTTCGTT AGTGCTTTTA TATTTTCATT	12600
	GGAATCATTa AGCTTTCTAA TCTATCGTTA ATGATTGCT TTAATAATTGG GTCGAAGTTA	12660
15	ATTGAAGGTG TGAAGTGTAT ATCTGTATTA ATAACCATGT CATTCAATTG CTGCTTCACT	12720
	TTGTTAACAA GTCTTCCGTC ATATAAAAAT AATGGTACGA CAATCAATTT TTGATACCGT	12780
20	TTCGAGATGC TTTCTAAATC ATGTGTAAAA CTAATCTCTC CATATAGCGT TCTCGCATAT	12840
	GTCGGCTTGC TAATTTGCAA ATTTTGAGCG CATATTTGTA ACTCTTCGTG TGCCTTAGTA	12900
	AACTTTCCAT TAATATTGCC GTGTGCAACA ACCATAACTC CAACTTGTG TTCGTCACCT	12960
25	GCTAATGCGT CACAAATACG TTGTTCAATT AATCGTCTCA TTAAAGGATG TGTGCCAAGT	13020
	GGCTCGCTTA CTTCTACCTT TATGTCTGGA TACCGTCGTT TCATTTCATG AACGATATTC	13080
	GGTATATCCT TGAGATAATG CATTGCACTA AAGATTAGCA ATGGTACAAT TTTAAAATGG	13140
30	TCAACCCAC TTTGAATCaA CGTCGTCaTT ACCGTCTCTA AATCCTGATG CTCACTTTCT	13200
	AAAAACGCAA TATCATAGTG ATGTATATCA TCTTTTACTA ATTCAGAAAT AAATGCTTCT	13260
	AACGCTTGaT TCTGTCTGCC GTGCCTCATG CCATGTGCAA CAATGATATT CCCATTCA	13320
35	TTTACCAACC CTTTCACACG TATTGTATAC CAAATCATTT TGT'TTTTGTG AAAAGAATCA	13380
	CATTATAATG TAAAATCAGG GAATTCCTG ATGCTGTAG TCATGCATAT TCCTTATACA	13440
40	TTTTCCCTTT TTGTTAAATC AAAAAAGCG ACCGATATAT GAATCCCTAC TCAACATTTA	13500
	TTTGAGCAAG CATTAAATATA TCGGTGCTT GTAGTGATA TTATTATCTT AAAATGGTGG	13560
	TTGGCCTAAT ATTGTTTCGT CAAAGCGCTC GGGTATCAAT ACTTTGCGCA TGATCACACC	13620
45	TAAATCGCCA TCATCATTTT CATGTTCGCT GTATATTTCA TAACCTCTTT TTTCAATAAT	13680
	TTTAAGTAAC CACGGATGCA ATCTTGCGA TGTACCTAAA GTAACGCGG CTGACTTTAA	13740
	CGTATCTCGC AAAAATGCTT CTTCAACATA AGTAAGTAAT TGGCTACCAT AGCCTTTCCC	13800
50	TTCACTACTCA GGATTTGTG CAAACCACCA GACAAAAGGA TAACCCGAAA TACTTTTCAC	13860
	ACTTCCCCAA GGATATCTAA CCGTAATCGT AGATATAATT TCATCATCAA TTGTCATGAC	13920
55		

CCAATCAATA CCTAGTTCTC TTAGAgGCGT AAATGCTTCA TGCATGAGTT CTTGCAATTT 14040
TTCTGCATCT T 14051

5 (2) INFORMATION FOR SEQ ID NO: 104:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1885 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

TAATCCTCAA CTTnGATTAT ATGGCTTGGG CGCATATGAA CTGCTTAGTT TAGTGTATGA 60
 CATTCATACA GTTCGCATGA CTATCATACA ACCTCGAATA GATAACTTTT CTACTGAAGA 120
 GTTACCAATC TCAAGATTAC TTCAATGGGG AACCGATTTT GTTAAACCCT TAGCCAGACT 180
 TGCTTATAAC GGTGAAGGTG AGTTTAAAGC AGGTAGTCAT TGTAGATTCT GTAAGATAAA 240
 GCATTGATGT AGAACACGTG CAGAATACAT GCAAAATGTG CCTCAAAAGC CACCACATT 300
 GTTGAGTGAT GAAGAGATTG CAGAACTTTT ATATAAACTG CCTGATATCA AAAAATGGGC 360
 TGATGAAGTA GAGAAATATG CGTTAGAACA AGCGAAAGAG AATGATAAAA CGTATCCAGG 420
 TTGGAAGCTA GTCACGGGAC GTTCAAGGAG AGTGATAACT GATACAAAAG CAGTCCGAGA 480
 CAGGTTAGTT GAAGCGGGTT ATAAACCTGA AGATATTACA GAAACCAAGT TACTTAGCAT 540
 TACGAATTTA GAAAAATTAA TCGGCAAAAA AGCATTTTCT AAAATTGCAG AAGGCTTTAT 600
 AGAAAAGCCG CAAGGTAAAT TAACACTTGC TACCGAGTCT GATAAACGAC CAGCTATAAA 660
 GCAATCTGCT GAAGATGATT TTGACAAACT ATAAAAATTA AAAAGGACGG TATATAAACA 720
 TGAAGCAAA AGTATTAAAT AAACTAAAG TGATTACAGG AAAAGTAAGA GCATCATATG 780
 CACaTATTTT TGaACCTCAC AGTATGCAAG AAGGGCAAGA AGCAAAGTAT TCAATCAGTT 840
 TAATCATTCC TaaATCAGAT ACAAGTACGA TAAAAGCCAT TGAACAAGCT ATAGAAGCTG 900
 CTAAGAAGA AGGAAAAGTT AGTAAGTTTG GAGGCAAAGT TCCTGCAAAT CTGAACTTC 960
 CATTACGTGA TGGAGATACT GAAAGAGAAG ATGATGTGAA TTATCAAGAC GCTTATTTTA 1020
 TTAACGCATC AAGCAAACAA GCACCTGGTA TTATTGACCA AAACAAAATT AGATTACGG 1080
 ATTCTGGAAC TATTGTAAGT GGTGACTATA TTAGAGCTTC AATCAATTTA TTTCCATTCA 1140
 ACACAAATGG TAATAAGGGT ATCGCAGTTG GATTGAACAA CATTCAACTT GTAGAAAAAG 1200
 GCGAACCTCT TGGCGGTGCA AGTGCAGCAG AAGATGATTT TGATGAATTA GACACTGATG 1260

55

TTGAGGTGTC AAGAATTTGA AATTTATGAA TATAGATATT GAAACATACA GCAGTAACGA 1380
 TATTTTCGAAA TGTGGTGCCT ATAAATACAC AGAAGCTGAA GATTTTCGAAA TTTTAATTAT 1440
 5 AGCTTATTCG ATAGATGGTG GAGCGATTAG TGCGATTGAC ATGACTAAAG TAGATAATGA 1500
 GCCTTTCCAC GCTGATTATG AGACGTTTAA AATTGCTCTA TTTGACCCTG CTGTAAAAAA 1560
 GTATGCATTC AATGCTAATT TCGAAAGAAC TTGTCTTGCT AAACATTTTA ATAAACAGAT 1620
 10 GCCACCTGAA GAATGGATTT GCACAATGGT TAATTCAATG CGTATTGGCT TACCTGCTTC 1680
 GCTTGATAAA GTTGGAGAAG TTTTAAGACT ACAAAGCCAA AAAGATAAAG CAGGTAAAAA 1740
 TTTAATTCGT TATTTCTCTA TACCTTGTAACACCAACAAA GTTAATGGAG GAAGAACrAG 1800
 15 AAACCTACCT GAACATGATC TTGAAAAATG GCAACAATTT ATAGATTaCT GTATTTCGAGA 1860
 TGTAGAAGTA GAAATGGCGA TTGCT 1885

20 (2) INFORMATION FOR SEQ ID NO: 105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2656 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

30 TAATCCTTAG TTCACTGnCA AATTTCAAAA CACCAGTTCC CTCTATCTGC ATCCATAGAA 60
 ACTGnATGTT TGTGTCAATA ACCGGATTAT ATTGTGATGn TGTTTGTAAC TCGATTAAAGT 120
 TATCATCTTT CGAAAAATTA TCTACTACCA TTATTCAACC ACCTTTCCTT CGAATAAACT 180
 35 CCATTTACCA ACKCCACCAG TACCAAAGTT TCTAACTAAA AATTGATGTG CAGACGGGAA 240
 GTTATTACGT CTTAATACTT GTGTTGTATT ACCTGGTGTA TTCGATTTTA CTTCTAATAT 300
 40 CCAACCTGCA ATACCTTTAA AGTCTTTAGG AAAATCAGTA AATCGGTTTG ATTCTTCAGT 360
 AGTGATATAG AAATCTAAAC CAACGATTTT TAAATCTGAT AATTTTGTA TACTCTTAGG 420
 GATATGTTCC CAATAACCGG CGTTTTGCGG GCAGAAATTC CATGCTCCGT TGTTTTTCTT 480
 45 ATTGAAAATG TCAATGACAC GTTCGAATTT AAGCATATTT CTACCTGTGC TGTTTCTGGt 540
 AAGTACTTGT CTTAGAGCAC CATTATAGTG TCCAGSCAGT ACATCCAAGA ACCACCCTGC 600
 ATCTCTAAAC GCTTTCGGTA ACGGGAAATC TAATGCATTT TGTGTGTCTT GaCGTATAGA 660
 50 TATAGTAATG ACCAACTTCC GTAATATCAC TTAGATATGC TGGGTCTGT ATTGGTAACG 720
 GTTTAACACG TCCGCCTGAA TCAGTCATTG ATACTTGAGG TGCGATGTTT TTCAAGAATT 780

55

	TAGTTACCCC GATTAGAAGT GCTTTACGTC CTGTTTCTAG ATCGTAATAC ATATCTAGAC	900
	CCTCAGCCTC TTGGAAATCT CCTTTAAAGT TGTATTTCAC ACCGCCTATA TCGATGCGAC	960
5	GTTTAAATAA CAATTCTTTC GTTTTGATAT CGAAGCCTTG TAAGTAGTTA GGGTTGGCTG	1020
	TATTCGAATC ACCTGTATAC CAATATAAGA TACCTGCATC ATAAGTGATA CCTTGCATAG	1080
	GTTGTGTATC TGAAGTGTAT TCCATAGGTA TATCCATTTG ATACAATACT TTGTCTATAC	1140
10	CTTTATCAAT ATCGTCAGCA CTTCTAACCT CAACAAAGTT CAACGAATTC TTAAGTTGTC	1200
	TTTCAGTGGG TTTATATTCA CGTCTAAAAA TCATTAAATT TTCTACCGGA TTATAAATCG	1260
	CTGACGTATA TCTGTCGTTA AATATATTCTG GCATGACATC TTGCATTTCA TTACCATAAG	1320
15	TTATTTCTCC AGTTCCTATAT TGGAAACGTA CAACTTGTT GTTTTTGTGA CTGTCCAATA	1380
	CAGCTGAATA AATCCATAAT TCTCCATCAA TGTATCTATA CGCATTGTGT GTACCGTGAC	1440
	CGCCGTTTTT AACAAAGCAAT CTATCAATAA ATTGTCCGTT GGGCTTCAAT CTAGATAACA	1500
20	TGTAATGATT ACCTGGACGA GCTTGCCTCA TATAAATAAT TTTCGTTCTA GGGTCTACCC	1560
	AAAATGATTG CATTACTGCA TTTGTATATG GCGATAAATC AGTGATAAAT TCCGGTTCTT	1620
25	GCTCTTTTGG TTCGAATCGG TATTCTGTCTG CTCGATATTC TTTATAGTGT TCATCTACAG	1680
	CTTTCTCAAC CTTTTTAGTG AAAACATCTA GTGTTGAATA ATCATGATAC AAACGATCTT	1740
	GCAATGTCTT ATGACCATAA CCTGTATTAT CAACGCGCGC GTCTTTTAcT TCGTTGATAC	1800
30	CGTCGCCGTT ATGACCTAGT ACCATGTTGC TAAATCGACC GTTTAAATAT GTTAAAAAGT	1860
	CAGAGACGTT ACTTGTAACA TTTAAATGTT CATACTTTAT TTGTTCTCCA TCATGTGCGA	1920
	ATACCTCTTT ATTCTGTGG TATTCAAGAG AGAAATTAAA ATCCGTCAGC ATGTCTGAAA	1980
35	TAAGTTTAAA GTTATACTCA TTTTCATCTA CATATCTGTA GTCAAAGACT CTACTTAAAT	2040
	CTGTAATTAG TTTATTACTC ATGTTTTCTT CCTTTACTAT CCATAAAACT GATmATAATT	2100
	TTTAATAAGC TCATACATAA TAACTTCATG ACCTCTTTCA TTAGGATGTA ATCCATCAGG	2160
40	CATGCTAGAT TTTCTAAATG CTGGATTATA TGGTTTGAAA TAATCTGTGT GATAAGCATC	2220
	ATATACTGGT ACATCCAATT CACTACAAGC CAATATCTGA GCATTGACAT AATCCTCTAA	2280
45	AGTTAACCTT AGTTTGTTTT TGTCCGTATC TTTACGGCGT ATCGTTGTAC CACTCATAGG	2340
	GCATTGCCTA GTAGCTGTCA TTACAAGTAT TTTTGAAGCT GGATTATTTT TCCTGATAAC	2400
	TTCAATTGCA GAACAAAAGG CGCCGTAAAA CGTTTTAGTG TCGGTTTTAT CAGTGCCTAT	2460
50	CGGTACGCCT GCCCAATAAC CATGTAACCA GTCATCATCT GTACCTTGTA ATATGATTAG	2520
	GTCTCCTCTT ATTTGCTCTG CTTGTCTaTA AATGCTGTTT TCTaCCGCTT CTTTACCTAT	2580

55

CTTGCCCTAAC ATTTCT

2656

(2) INFORMATION FOR SEQ ID NO: 106:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4854 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

15	AAAATGAGGG TTCTAGCGGA AATTACCAAA AGCGTGGTTC ATACTATGGG CAGCGTAATC	60
	GTATTTCAAA AGAAAAAACA CCTAAATGGT TAGaAAATAG AGATAAACCT AGTGAAGAAG	120
	ATTCGGCTAA AGATAATAGC GTAGATGATC AACCAATTAGA GCAAGATCGA CAAGCATTTC	180
20	TAGATAAAATT ATCTAAAAAA TGGGAGGAGG ACAGTCAATA ATGAAGCAAT TTAAAAGTAT	240
	AATTAACACG TCGCAGGACT TTGAAAAAAG AATAGAAAAG ATAAAnCAGA AGTAATCAAT	300
	GACCCAGATG TTAAGCAATT TTTGGAAGCG CATCGAGCTG AATTmACGAA TGCTATGATT	360
25	GATGAAGACT TAAATGTGTT ACAAGAGTAT AAAGATCAAC AAAAACATTA TGACGGTCAT	420
	AAATTTGCTG ATTGTCCAAA TTTCGTAAAG GGGCATGTGC CTGAGTTATA TGTTGATAAT	480
	AACCGAATTA AAATACGCTA TTTACAATGC CCATGTAAAA TCAAGTACGA CGAAGAACGC	540
30	TTTGAAGCTG AGCTAATTAC ATCTCATCAT ATGCAACGAG ATACTTTAAA TGCCAAATTG	600
	AAAGATATTT ATATGAATCA TCGAGACCGT CTTGATGTAG CTATGGCAGC AGATGATATT	660
	TGTACAGCAA TAACTAATGG GGAACAAGTG AAAGGCCTTT ACCTTTATGG TCCATTTGGG	720
35	ACAGGTAAAT CTTTTATTCT AGGTGCAATT GCGAATCAGC TCAAATCTAA GAAGGTACGT	780
	TCGACAAATTA TTTATTTACC GGAATTTATT AGAACATTAA AAGGTGGCTT TAAAGATGGT	840
40	TCTTTTGAAA AGAAATTACA TCGCGTAAGA GAAGCAAACA TTTTAATGCT TGATGATATT	900
	GGGGCTGAAG AAGTGACTCC ATGGGTGAGA GATGAGGTAA TTGGACCTTT GCTACATTAT	960
	CGAATGGTTC ATGAATTACC AACATTCTTT AGTTCTAATT TTGACTATAG TGAATTGGAA	1020
45	CATCATTTAG CGATGACTCG TGATGGTGAA GAGAAGACTA AAGCAGCACG TATTATTGAA	1080
	CGTGTCAAAT CTTTGTCAAC ACCATACTTT TTATCAGGAG AAAATTTTCA AAACAATTGA	1140
	ATTTTAAAT GATTGGTGTA TAATGAATAC AAATCTAAAT CGTTTAAATG ATTGAAGACA	1200
50	AGATGATCTA ATCAATATTA CACAGAAAGC CATTGTTTGA TGAGAATATG GTTAATAAAT	1260
	TAGATGATTA CTACTTCATT TATGGTATTT GTAATGAATA CCCGGATCAA GACCGTTATC	1320

	CTCGTCCCTT GTATAGGGGC GGGATTTTTT GTTTTTTTCA GACATAAATG TTTGTTGGTG	1440
	TCATAAATTC CCTGTTTATT GTTAATAGGT TTAATGTTAA AACGATGATT GTTGTTCAAT	1500
5	TTTTTAACGA GGTGAGATAA AAGTATTTAT AAAGCAAATA GGAGGGTTTA ACATGGAACA	1560
	AATTAATATT CAATTTCCAG ATGGTAATAA AAAGGCGTTT GATAAAGGTA CTACTACTGA	1620
10	AGATATAGCA CAATCAATTA GTCCTGGATT ACGTAAAAAA GCTGTTGCCG GCAAATTTAA	1680
	CGGGCAACTT GTAGATTTAA CTAACCGCT TGAACTGAT GGATCAATTG AAATTGTGAC	1740
	ACCAGGTAGT GAAGAagcGT TAGAGGTATT ACGTCATTCT ACTGCACATT TAATGGCACA	1800
15	CGCGATTAAA AGGTTATATG GTAATGTTAA ATTTGGTGTA GGTCTGTAA TAGAAGGTGG	1860
	ATTCTACTAT GACTTCGACA TTGACCAAAA CATCTCATCT GATGACTTTG AACAAATTGA	1920
	AAAAACAATG AAACAAATCG TTAACGAAAA TATGAAAATC GAACGAAAAG TGGTTTCACG	1980
20	AGATGAAGTG AAAGAGTTAT TCAGCAATGA TGAATACAAA TTAGAATTAA TCGACGCGAT	2040
	TCCTGAAGAT GAAAATGTAA CATTATATAG TCAAGGTGAT TTTACTGATT TATGTCGTGG	2100
	AGTTCACGTT CCATCAACAG CTAAATTTAA AGAGTTTAAA CTATTATCTA CAGCAGGTGC	2160
25	ATACTGGCGT GGAGATAGTA ACAACAAAAT GTTACAACGT ATATACGGTA CTGCTTTCTT	2220
	TGATAAAAAA GAATTGAAAG CACATTTACA AATGTTAGAA GAGCGTAAAG AACGTGATCA	2280
	TCGTAAAATT GGTAAAGAGT TAGAACTATT CACAAATAGC CAATTAGTTG GTGCTGGTTT	2340
30	GCCATTATGG TTACCTAACG GTGCAACAAT TAGACGTGAA ATTGAACGTT ACATTGTTGA	2400
	TAAAGAAGTT AGCATGGGAT ATGACCACGT TTATACACCA GTACTTGCTA ATGTTGATTT	2460
35	ATACAAAACA TCTGGTCACT GGGATCACTA TCAAGAAGAT ATGTTCCAC CAATGCAGTT	2520
	AGATGAAACT GAATCTATGG TATTACGTCC AATGAACTGT CCACATCATA TGATGATTTA	2580
	TGCGAATAAA CCACATTCAT ATCGTGAATT ACCTATCCGT ATCGCTGAGC TAGGAACGAT	2640
40	GCATAGATAT GAAGCAAGTG GTGCTGTATC AGGATTACAA CGTGTTCTGT GTATGACTTT	2700
	AAATGATTCA CATATCTTTG TTCGACCTGA TCAAATTTAA GAAGAATTCA AACGCGTTGT	2760
	AAACATGATT ATTGATGTGT ATAAAGACTT TGGTTTCGAG GATTATAGCT TTAGATTAA	2820
45	TTATAGAGAC CCTGAAGATA AAGAAAAGTA CTTTGATGAT GATGATATGT GGAATAAAGC	2880
	TGAAAATATG CTTAAAGAGG CAGCGGATGA GCTTGCTTA TCGTACGAG AAgCGATTGG	2940
	TGAAGCGGCA TTCTATGGTC CGAAACTAGA TGTTCAAGTT AAAACAGCGA TGGGTAAAGA	3000
50	AGAGACATTA TCAACAGCAC AACTTGATTT CTTATTACCA GAACGTTTTG ATTAACTTA	3060
	TATTGGTCAA GATGGTGAAC ATCATCGTCC AGTTGTTATT CATCGTGGTG TTGTATCAAC	3120
55		

AGCGCCAAAA CAAGTTCAAA TCATTCCAGT TAACGTTGAT TTACATTATG ATTATGCGCG 3240
 CCAATTACAA GATGAATTGA AATCTCAAGG CGTTCGTGTA AGTATTGATG ACCGTAATGA 3300
 5 AAAAATGGGT TATAAAATCA GAGAAGCTCA AATGCAAAAA ATACCTTATC AAATCGTAGT 3360
 TGGGGATAAG GAAGTTGAAA ATAATCAAGT GAATGTGCGT CAATATGGAT CGCAAGACCA 3420
 AGAAACAGTT GAAAAAGATG AATTTATCTG GAATCTAGTT GATGAAATTC GTTTGAAAAA 3480
 10 ACATAGATAG ACAGTTGTCTG CAATAAAATG CTTTAAACT TTTATTGCGT ATCAAGTTTT 3540
 ACAGGGTTGA TTATGCGTGA TGAATCCTGT ATATTACAAG TTAGTTAAAA TATTAAATTG 3600
 AGTTAGAGGT TGCATGTTTA ATTAGTAACT TGTCAGAAGT ATTTATGGTA CATAAGTTGA 3660
 15 ACAAGTGAAG GGTAAAGATG CCGAAATAGA TATAAACCAT AAATTATATC TATTGGGACA 3720
 GTTTTCGAAT AGGAACTGTA CTGTCACAGA ATGTGATGTG CTACCTTATA TAGATAATTG 3780
 20 CCAAAGTGGT TGCATATCTT AAAGGTATGT AGCCACTTTT TTACTTTTAA TATCACTATG 3840
 TTCTGTAAAA AAGGGTATGA AAGTGAATAA AGGTTATTTA TTTCTTGGCC TCTAAACAT 3900
 GGAAAGGGAG CTTATATGTC AAAAGTTCAA AATGAAAGTA ACAATGTTGT CAAAAGGGGA 3960
 25 CTTAAAGATC GTCATATTTT TATGATTGCG ATTGGGGGTT GTATTGGTAC AGGTTTATTT 4020
 GTAACCTCTG GTGGAGCAAT TCATGATGCA GGTGCTTTGG GTGCATTAAT AGGATACGCA 4080
 ATTATCGGAA TAATGGTATT TTTCTTAATG ACGTCACTTG GCGAAATGGC TACGTATTTG 4140
 30 CCAGTATCAG GTTCATTTAG TACATATGCT ACAAGATTG TTGATCCATC TTTAGGGTTT 4200
 GCGCTTGGTT GGAACATTG GTTTAACTGG GTAGTGACTG TAGCAGCAGA TATTACGATT 4260
 GCAGCACAAG TCATTCAATA TTGGACACCA TTGCAAGGCA TACCCGCTTG GGCATGGAGT 4320
 35 GCGTTGTTCT TAGTTATAAT TTTTAGTCTG AATTCGTTAT CAGTTCGCGT CTATGGTGAA 4380
 AGTGAATACT GGTGGGCATT GATAAAAGTG GTTACAGTTA TTGTTTTCAT TGCAATTGGT 4440
 TTATTAACGA TTGTCGGAAT CATGGGTGGT CATGTTGTAG GATTGGAAT ATTTAATAAA 4500
 40 GGTGAAGGTC CAATCTTGG TGGCAACTTA GGAGGAAGTT TGTTATCAAT TCTAGGTGTA 4560
 TTCTTAATCG CTGGTTTCTC ATTCCAAGGT ACTGAGTTAA TTGGTATTAC GGCTGGTGAA 4620
 45 TCAGAAAATC CTGAACGTGC TGTGCCGAAA GCAATTAAAC AAGTATTCTG GAGAATTTTA 4680
 TTATTTTACA TTTTAGCCAT TTTGTTATC GGTATGTTAA TTCCTTATGA TAGTAGTGCA 4740
 TTAATGGGGG GTAGTGATAA TGTAGCAACG TCTCCATTCA CATTAGTGTT TAAAAATGCT 4800
 50 GGATTTCGCT TTGCAGCATC ATTTATGAAT GCAGTCATTT TAACGTCTGT GTTA 4854

(2) INFORMATION FOR SEQ ID NO: 107:

(A) LENGTH: 2488 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

10	ATCAAAAATT GATTGTTTT C NTTTTTTGT TTCAGCGCGG GATCTTTTAC GTCTTTTGTG	60
	AAAACGaTTT TATTATTAAC TACTTTTACT GGATAACTTT TGTATGTCGA GTCAGTAGCA	120
	TTTTTCTAT CGTTTGTAGT TGTGTCATAT TCACCagTTA TTTTATGTGT GTTCTTATCT	180
15	ACCTTTAACA ACATACGGTC TTCTTTTAAA AGCTCATCTG ATCCAACAAC TGAATAAGAG	240
	GATTCTATAT ACCATGTGTC TTGATCATT TTTTCATAAT GGGGATTATC GTGACCATCA	300
	ATTCATAAA GCGTTTCTAA GTTTTTAATA GGATACGTAC TTAGTACTTT TTTAAGACCA	360
20	TCTTTCAAAT GAATTTGTTC CCACTTCATT GCCAAAACA TATCGCCACT GACTACAATT	420
	GAAATAATAA TAATTGCTGC TAAGTTAAC CAGAAATTT TATGTGCTTT CATACATTCC	480
	CACCGTTTCT CAAAATACTT CATTAACTAT ATAATAATAT ATTTTGAAAA ATATTTACAT	540
25	CAGTATTAAA GTGAATATCA AATTTTAAAT TTATGAAAT AATAGATATT TATAAAAAGC	600
	GGAAAAGAGA TACAATAAAA AACTGCATGA CGTTTGAGAC GTCACACAGT GTAACATAAA	660
	ATTTAAAAAG TTGTTGCTAA TTTTTCAGCA TTATTAATAC TAGTTGCTTT AATTTCTTCA	720
30	GTCTTATGAG GTTCAGCATT GTGTCCTTCA ATAATGATTG TTTCATATGA TGGCACACCT	780
	AAGAATGTCA TAATTGTTCT TAAATAACGG TCACCCATTT CAAAATCAGC AGCAGGTCCT	840
	TCAGTATAAT ATCCACCACG TGATTGAATG TGTAATACTT TTTTGTGAGT TAGTAAACCT	900
35	TGTGGTCCTT CAGCAGAATA TTTAAAAGTT TTACCTGCAA TTGAAATAGC ATCAATATAT	960
	GCTTTAACTA CAGGTGGGAA AGAAAGGTT CACATAGGCG TTACAAATAC ATATTTATCT	1020
40	GCACTTAAAA ATTCTTCTAA AATGTCACCTC AATCTTGAAA CTTTCATTG TTTCATCATCA	1080
	GTTAACGTTT CGCCATTACT CATTTTTCCC CAACCAGTTA ATACATCTTT GTCAATAACT	1140
	GGAATATAAG TTTCAATATA ATCAATATGT TTTCACTTCAT CATCAGGATG TTGTTGTTGA	1200
45	TATGTTTCGA TAAATGCTTT ACCAGCCGCC ATAGAATTTG ATACCAGTTC ATTAAAAGGG	1260
	TGTGCTGTAA TATATAATAC TTTTGCCATT TGAAAATTCT CCTCTGkTTC TGTTATTTTC	1320
	TTAAGTATAA TTATTATACT CGATATAAAA TTTAATATCA ATCAAAATAT TCAAATTACC	1380
50	ATCATTTTCT TCATCTATAT nTGGCAGTAC TACTAAAGTA TGAGTGCATT TAATTATGaa	1440
	ATAGTTGATT TaGAATATAT ACTTAATACC CAAAATATAT GAAGGATGGA TGCCACTATG	1500

55

ATTATTTATA TAGATGACAT TCAAAAATGG TTTAACCAAT ATACCGATAA ATTGACACAA 1620
 AATCATAAAG GACAAGGACA CTCAAAATGG GAAGACTTTT TTAGAGGGAG TCGGATTACT 1680
 5 GAGACTTTTG GTAAATATCA ACATTCACCA TTTGATGGTA AGCATTATGG CATTGATTTT 1740
 GCATTGCCAA AAGGTACACC AATTAAAGCG CCGACGAATG GTAAAGTAAC ACGTATCTTT 1800
 AATAATGAAT TGGGCGGCAA GGTATTACAG ATTGCCGAAG ACAATGGAGA ATATCACCAG 1860
 10 TGGTATCTAC ACTTAGACAA ATATAATGTC AAAGTAGGTG ATCGAGTCAA AGCAGGTGAT 1920
 ATTATTGCAT ATTCAGGCAA TACAGGTATA CAAACGACAG GCGCACATTT ACATTTTCAA 1980
 15 AGAATGAAGG GTGGCGTAGG TAATGCATAT GCAGAAGATC CAAAACCGTT TATCGATCAG 2040
 TTACCTGATG GGAACGTAG CCTATATGAT TTGTAGTTAT AGAAGGGTGC CCGCAGTCTA 2100
 AAAAATTAAG CAATCATTGT GTGAGTATGA TACTTACATA ATGGTTGCTT TTTTCAATGA 2160
 20 AAATCGTAAT GCTAAGTCAT ACTTGTTTGA TTTAGATATT ACTTAAAATG TAAGACAAGG 2220
 TTGTTAGCAT TGGCAGTGAA ATATCGCACA TAAAAACAT TATTGTCACA CTAGAAAATA 2280
 GTTGTGCACT ATATCAATTT TCTGTATAAA AGTTTAATTC TGACAGTAAT GTAAACGTTT 2340
 25 ACAATTTATG ATTGACATTA ATAATGACTG AATATATGAT TTATGTAAGT ATTTGTGCAA 2400
 CGTTTTTACA AAGTGTATTG CACaAycAAA CTGTAAACaA aGTATGGGGg GCCATAACAT 2460
 30 GGCAGAACTA AGTTAGAGCn TATTAATAA 2488

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 4093 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

40 TTTTCTTTAT TTCAAmCTGT ATATTaATGA TGTCACCTCA TTTGATACGA TTCTTGATAA 60
 CCTATTCAAA ATTCCGCCAA ATAACATAAA TATTATATAA ATGCCGATAC TTTTAATCAT 120
 45 TTTCTACTTT TTCTTCGATA CGGAACCTTG TTTTCGAATT GAACACTTCA CCAGCTTTTA 180
 AAATTGACGG TGCTTTTTTCA CCATATAAAT TAATATCATT TGGTAAAAAT TGTGTTTCTA 240
 50 ATGTAAAGCC AGAATGTGGT TTATAAATAT TAAATGGACT ATCCCACTCA TCAGGCTGGT 300
 TAAAAGTAAA GAACACAACA TGAGGCATAT CTGTATCGAC CTCTAACATA AATTCATGAT 360
 55 TTTCAACATA CATTTTATGT TCACCAACTG TAAATGGGTG ATCGAGACCA CCAAAACGTG 420

	TATCTTCAAA CACTTCATGT AAATCTAGAA TATCACCTGT AACAAATATTT CGCTCATCTA	540
	ATACATACAT ATCTAATTGA TTA CT TGAAA TGCGATGATT ATCAACGACA TTATTATCTC	600
5	GATTCAAATT GAAGTACACA TGATTTCGTAG GACTAAACAA TGTGTCTTCT GATGCAACTG	660
	CTTCGTATTC AATCGACCAT TGGTGATCCG CATCATAAAT ATGTGTAATC GTCACATCGA	720
10	TATCACCCGG GAAATGATCA TCAGCTGATT TCAACACCGT CTTAAATATA ACTTTAATTT	780
	GAGCAATTTT ATTTCTAATT TCATAATCAA ATA ACTTATT GTCCAAACCA TGACATCCAC	840
	CATGTAAATG ATGTTACCCG TTGTTTTTTTT CTA ACTGATA TTCTTTACCT TTCAACTTAA	900
15	ATTTAGCATT ATCAATTCTA CCGCTATATC TTCCTATAGA AGCACCAAAT TTAAAAGGAT	960
	TACTATGATa AAATTCATCC GCTTCAACAA CATTTCCAAG AACAAATATTA TTATCATGAT	1020
	ATTTCCAAGA CACTACTCTT GCTCCATAAT TCGTAAAAAT AATTTTAGTT TCATCATTAT	1080
20	CAATTTTGAT TAAATCTACA CCTTGTCTTT GGTGCTCAAC TTCAACTATC ATTTTACTT	1140
	CTCCCTTCTA ACCACAAGTG TTCAAGCTCT GCTGGGTAGC AACATTACTA AAACACCTAC	1200
25	AATACAAATG ATTGCACCGA TAACATCATA TTTATCTGGC ATTTGTTTAT CTACGACCAT	1260
	CGCAAAAATC AAACATCATGA TGATAAATAC GCCACCATAT GCTGCATATA CTCTCCGAA	1320
	TGATGGAAAT GATTGAAATG TCGCAATGAC ACCATATAAC ATGAGTATCG CACCGCCTAT	1380
30	TAGCCCAACA AGTGAAGACT GTCCTTCCCT AAGCCACAGC CAAATCAGGT ATCCCCCACC	1440
	TATTTACAT AAGCCAGCTA ATATAAATAT AAAAATCGGA TATAACATGA AATCACTCCA	1500
	TCACACATTT GCTATCAATA ATCTATCGGC TACATATCAT TTGTTTACAT TTCTTCTTAC	1560
35	TTACATTCC CATTTTAAAA AGTTCGTTTT CACATTCATA TTGTACACTT TTTTAGACAT	1620
	TATTCTATAG CTAAATATAA AAAAATAAGA GTAACACGCT TTCATCATCA TTTTATATGA	1680
	TAAATGTGTG TCACTCTCAT CAATTTTATT TTTTAAATAC ACGTTTCATT GAATTAAATA	1740
40	AGCCACGTTT AAATGTAAGT ACTGAATCTT TATATGTTTT AATTGCAATC CATATCAAGA	1800
	CAGCTACCAT TACAATTGAG ATTAAAGAAC TTAAGATGAC CTCATATATT TGAAGCCCTG	1860
45	AAGTTTGAGC GCGTACAACT AATTGAAATG GCGCTAAAA CGGAATATAA CTGTGATTA	1920
	AAGCAAGTTG TCCATCAGGA TTATTTATCG TGAATATCGC GATATAAAAT GCAATCATAC	1980
	CAAGTAATGT CAGTGGCATC AAAGATTGAT TTAAATCTTC TATTCTAGAT GTTAATGATC	2040
50	CGAGGATGGC TGCAAGTAAT ACATACGCCG TAATTCCAAC AATACTACTT ATAATTCCGA	2100
	CAATAATAAT TTGCCAAGAC AATTGATTCA TTTCCACGTT AAAACCTTGT AGCAAGTCTT	2160
55	TTAAGTCAAA GGCAAAAATG CATATAACTG CCATCAATAC AATTAAAATA ATCTGAGTCA	2220

	TAATAATCAT TTCAATGACA CGCGATGTTT TCTCACTAGC AATTTCCATA GCTATTTGAG	2340
	ATGCATAATT TAAAACAATG AAGAACATTA GAAAGATAAT GCCATmaGcT AAAGCATAGT	2400
5	TGAAAATCTT TTGTCCTTCT GATACTTTAT CGACTTCATC ATTAGAAATC ACCTTATTAT	2460
	CAACTTTACT TTGTGCTTGT AATTTTTGTA AGTCTTCTTT GTTGATATTT AATCCCCGG	2520
10	CTACCATATT TGTTTGAATA GCTGTAAGCA GTGCTTGTAC TTTTGTGAA TCTTCATGAC	2580
	TTACTCGCTT CTCACTAATG ATTGTCCTT GTAACGTGCG ATTTTGATT CACCTTGATAA	2640
	TATAAGCTTT ATCAAGTTTA TGTTTTTTTA CTTCTTTTTC AGCATCTTCT ATAGAAACTT	2700
15	TAGTAAACTT AGCATCACTA TGAAATGTAT TCGCCTGTTG CTTGAAAACC TTATAGATTT	2760
	GTTCAATTCGG TGCTGCTACA CCAATTTTAT CTGGACCATC ATCAAACATG TTAATAATCT	2820
	TATCAATGTT AGATAGGCCA ATCATTAAAG CAGCAATAAT AATCATAAAA ATTACAAATG	2880
20	ATTTAGCTTT AATTTTTTTG ATATATGTCA AAGTAAATGT CGCCCAAAC TTATGCATCC	2940
	TTGCCACCAA CCTTCTCAAT GAATATATCT TGTAATGATG GTTCTACAAC TTGGAATCGT	3000
	TTAACATAAC CTTGATGTGC CACAACCTGA TAAATATCTT TGGCTACGTC TTCATTCTCA	3060
25	ATCGTCAACT GAAGACCTTG CTTCATGTTT TCACTATGAA TGATGCCTCT AATGTTTGT	3120
	AAATCTGGTA GTGTTGTTTC TGATTCAATG ACAACTTTCT TGTTACCATT AGATGCACGT	3180
30	ACATGATTGA TATCACCAGA AACACAAGT TGACCTTTAT CTAAAATACA AACATCATCA	3240
	CATAATTCTT CAACATGCTC CATACGGTGA GAACTATAAA CGATTGTACT GCCCCAATCA	3300
	TTTAAGTCTT TAACTGCTTc TTTTAATAAC TCAACATTAA CTGGGTCTAG ACCACTGAAA	3360
35	GGCTCATCTA ATATTAGTAA TTCTGGTTTA TGTAACATAC TTGCTAACAG CTGAATTTTT	3420
	TGTTGATTCC CTTTTGATAG ACTATCAATT CGTTTTTGC GGTTCAGT AATATCAAAA	3480
	CGCTCAAGCC AATACGATAT TTGCTGTTGT ATTTCTGTTT TTGACATTCC CTTTAAAGTT	3540
40	GCCAAATATT TCAATTCTTC TTCAACTGTC AATTTCCCAT GTAAACCGCG TTCTCCGGT	3600
	AAATAACCAA TACGATTGTA CATTGTTTTA TCTAGTTTTT TACCGTTATA CGTrrTGTGT	3660
45	CCTTCAGTTG GTTCACTTAA GCCTAAAATC ATACGAAATG TCGTTGTTTT ACmTGCACCA	3720
	TTTCTTCCTA GAAAACCTAA CATTTTACCT GATTCTAACT TTAATGAAAT ATCATTTACT	3780
	GCCGTCATCT TGCCAAAACG TTTCGTAACA TGTTCAATTA CAAGTCCCAT ACTTTGCCTC	3840
50	CTAAAAaAT ATGTATTTAT CTTAATATAA CATTTCATT CTCTATAAAT GCAATATTTT	3900
	TAAAATGAAT TTATTTTTTAA AATTTCTGAA ATTGAAAAAT TTAAATAGTG CCATTTTTGC	3960
55	ATGTTAAGTA TCATTAGCAC TAGATATGTT TTTTCCATGC CTTTATTGCC TTATTTGTAA	4020

CTTnCCGGTG TTT

4093

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17846 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

15	TGCCAAACTA CCTTTTGACA GTCGTTGCTG TACTTCAGGA TGATCAATCA CATATnTTAC	60
	TTTATCAAAT AGGGCATCTT CATCATTTTT AGTAATTAAA TAACCATTGA AATCTGAAGT	120
	AATCAGTTTCG TTAGGTCCAT ATTAAATATC ATAACATAATA ACTGGAACAC CATGTGCTAA	180
20	AGATTCAAGT AGCGCTAAAG AGAAACCTTC CATGTTACTT GTTATTAAAC TCAAATAGGC	240
	ATCGCTATAT TCTTGGTCTA GATTGCTTAA AAAGCCGCGT AAGTAAACAT GATTTTCCAA	300
	TCCATATTTTT TGTATCAATT CATTTAATTT TTTACTTTCA GAaCCAAAAC CATACATATG	360
25	AaGCTCTATT TTTGGGACAT ACGATACTAA GCGTTTAATT AATTCAATTT GTTGATGTAA	420
	TTGTTTTTCA GGTGAATAAC GAGCAACGGA AATTAATTTA AACTGCGCT GATCTAATGT	480
	TTGGACTGGT GTATCAATTG TTCACTATA GCCGACAGGA ATATTAACAA CTGGAATAGT	540
30	ATGGTTAATA CGTTTTTCAA CATCTAATTT TTGCTGCTCA GTAGAAACGA TAATTGCACG	600
	ATATCGAGAT AAATTTTCAA ACATCGCTTT ATATACATTT TTAAATGGCG ATGAATCTAA	660
35	TGCATCAATA TTTTAAATGT GTGTACTGTG AAGCACAGCT ACTACTGGGA TTGACTCAGG	720
	CGTTAAGTTG AAAATAGGTG CTGTGTACAC ATTACGATCA CTGAAAAATA AATCCCCATG	780
	TTGATATAGT TGTTTAATGA AAAATGCGCC TAATTCCGTT TCATTATTAA AGAAATATTG	840
40	TTTGTTAGCA TAGTAAACAA TAATTTTTTG TACTTCTGGT TTGCCATCCT TGTAAGAAAA	900
	ATACTTTTCT AATTTTGTGT CACCTTCTGG ATTATAGAAA AATTCACATA ATGTTTGTG	960
	TTTATCAACA AGAATCCTAC TACAACCTAA AAAGCCACGC ACATCATAAA AATCACGTTT	1020
45	TACTTtTCGT CTTTGACTAT CAAAATGATT TACATAATCT AATATACGAT ATTTAGGATC	1080
	TTGAAAATGG GCATACATTA AGAAACGCTC TTGATCATAT ATTCTAAAGT CATGACTATT	1140
50	TTCAACATGT TTTAAAGTAT AATGACATTC ATCAGTCCAA TACGACAACC AGTCAAATGG	1200
	TTCAATGCGT TCTAAATATG TTGCTTCTTG GAAGAAATCA TACATATTAA TATAGTCAGA	1260
55	ACTAGTAATA TAATTTTGGG CATTTCTATA TAAATATCTA TTCCATGACA GAAATACACA	1320

CCCAGTTAAA TTAACACCTA AACTATTACC TACAAAATAA TTCATTTACA ACACCACTTA 1440
 TATCTATTTT TTATAATTAT ATCACATAAT ATTTAATTAC TTCTTTTAAC TGGAAGATGT 1500
 5 GTTTATTTAT AAAACAACAA ATTTTGATAT TTATAATGAT AGTAGTTATT CAATCACTAC 1560
 GACCCaATAT ATCATkGTAG AGCTTAGGAT ATTGATTTAT GACTCAGGCA CATCAaATGa 1620
 10 GAgGATTTAT AAaRGAGATA TACAACTCTA GAAGGTATAA TAAAAACGCG CAACTAATGT 1680
 TACGCGTTTG AATTAATCAT ATGATATTAT TTGCGATACT TTAATTTAGC GAAAgCATCA 1740
 TGTGATGGA TAGACTCTTC ATTACGACAT TCGATATCGA AACCGTCTAA CCAATCAAAT 1800
 15 TCAACTAAGT CCGCGGCAAT TAAACGAATT AAGTCTTCGA CAAAACGTGG ATTTTCATAT 1860
 GCACGCTCTG TCACACGTTT TTCATCAGGA CGTTTTAAAA TAGGGTATAG AATTGAACCT 1920
 GCATTAGCTT CCATTGCATC TAAAATTTTA TTTTATAGT CATCAACTAT GTCTTGATCT 1980
 20 TTATTAATAT ATGTTTTAAC AGTGACAACA CCACGTTGGT TGTGCGCTGA ATACTCACTT 2040
 ATTTCTTTTG AACAAGGGCA TAGCGTTGTG ACAGTTGCTT CAATAGTAAG TTCTTTACGT 2100
 GTAnCTTTAT CACCGTCAAT TGCTAATCCA TAAGTGACAT CGGCATTACC AACTGCTTTA 2160
 25 ATATTTGTGG TTGGACTATA GCGATCAAAG AACCATTTC CAGAAACATC AACGCCTGCC 2220
 GCATTTTGTT TCATATTCGT TTGTAAAGTG CGTAACACCT GATAAAGTGT ATTAAATTCA 2280
 AGTTCAATAC CATTATCATA GTGCTTTTCA ACACTTTCGA TTATACGGCT CATATTAATA 2340
 30 CCTTTTTCGT CTTTTGTAA ACTTGTTGAA AAACTAAATG TGCCAGCTGT TTGATACTGG 2400
 TCAACAAGTA CAGGGTACAC TAAGTTTTTA ATACCAACTT CTCTATTTT AAATAAAAAA 2460
 35 TCTTTATGTG TACTTTGTAA ATCTGTCATT TCGTTCTTAG TAGTAGGTTT CGTGCCTTCA 2520
 ATAGGATCTA CGGAACCAAA GTGTTTCCAA CGACCTTCTC GTGTCGATAA ATCAAATTCA 2580
 GTCATTTTTT TCCTCCGTTA AGATTTAAAG TGATATGTCC AATATGGTTC GACTGTTAAA 2640
 40 AAGCTGTGTT GTTTACCATC GATTTCAAGG CTTGCTAATT GTTTTAAAAA TGGACCTGTT 2700
 TGAGAAGCAT GTGCTTCAAA TGCCTTAATT TTAAGTTCTT TAAAATCTGT AATATCATT 2760
 TGAATATCAG GTTCTCCAAG AGCTTCGGTT GCATCATTAC TGAACGCAAC TAAAGTTAAA 2820
 45 CGAGGGCGTT CTTCTTTAGG CATGCGTTCA ACCGTTGAA TTACAGCGTC TGCTGTTGCT 2880
 TCGTGATCAG GATGTACTGC ATATCCAGGA TAAAATGAAA TAATCAATGA TGGATTTGTA 2940
 50 TCATCGATTA AAGATTTAAT CATACCATCT ATATGTTTCAT AGGGTTCAAA TTCGACAGTT 3000
 TTGTCACGTA AACCCATTTT TCTTAAATCA GTAATACCGA TAACTTTACA AGCTTCTTCT 3060
 AGTTCACGCT CACGAATACT TGGTAATGAT TCGCGTGTG CAAATGGGGG ATTACCTAAA 3120

55

	TAATTTGCTA ATGTGCCTGC AGATGAGAAG GTTTCATCAT CAGGATGTGG AAATATTACT	3240
	AATACATGTC TTTCGTCACT CATGTTGATG CCTCCTCTAT AAATTAAATG GTCGCTCACT	3300
5	AATTTGAAGT GCTGCAGCGA GTTGACCTTC GTAATTAAAA CCTGCAATTA AAAATTCATC	3360
	ATGCTCATTG ACCTCAAAAT GCGTTAGACC TTGTACATAA ACCCAACCAC CATTTGATAG	3420
10	TTTAAGACCA ATGCCATAAG GTTCTTTATT ACCACCTTTT AGTTGTGCAT GCGTATATGT	3480
	TATTTGTATG TTTCTTAAAA AAGTACCAGC ATTAATAACA CGTTGATCGA AATGGTTTCGC	3540
	ATAGGCCCCA TTTGTCGTTT CAACATGCAG ATACACAGGT TTATGTTCAA AAGAAGCAAG	3600
15	TAAATCTATA ACTTCTTGTT CTTTAATTGG TTCCAACACG TTCACTCCTT ACACTATCAA	3660
	TGTGTTTATC TTTCTATTTT ACTAAAACT ATTCGATAAT TGTATACGAT TGCTCAATTA	3720
	TTTATAAATT AATTTTCATG AAGGGTAATT ACTCAGGATT ACGTAATCAT ACAGCATTAG	3780
20	TTTTTTACTT TTAATAATCA AAAATTTGTT GGAATTTGAA AAGTGTTAAA CATTAAAAAT	3840
	GATGCTATAT TAATGGTGTA TGAATGAATT CATAAGTTTT TAAAATGTAT TAAATTTGTG	3900
	GAGGCATGTA AACAATGAAA GTATTAACT TAGGATCGAA AAAACAAGCA TCATTCTATG	3960
25	TTGCATGTGA GTTATATAAA GAGATGGCAT TTAATCAGCA CTGTAACTA GGTTTAGCAA	4020
	CTGGTGGTAC AATGACAGAT TTGTATGAAC AACTTGTTAA GTTGTTAAAT AAAAATCAGT	4080
30	TAAACGTAGA CAATGTATCC ACGTTTAATT TAGACGAATA TGTAGGTTTA ACCGCATCAC	4140
	ATCCGCAAAG TTATCACTAT TATATGGATG ACATGCTTTT CAAACAATAT CCTTATTTTA	4200
	ATAGAAAGAA CATTCATATT CCAAATGGAG ATGCCGATGA TATGAATGCG GAAGCGTgCA	4260
35	AAATATAATG ACGTTTTAGA ACAACAAGGT CAACGTGATA TTCAAATTTT AGGTATTGGT	4320
	GAAAATGGTC ATATTGGATT TAATGAACCT GGTACGCCGT TTGATAGCGT TACTCATATC	4380
	GTTGATTTGA CTGAAaGTAC TATTAAGGCT AATAGTCGAT ATTTTAAAAA CGAaGATGAT	4440
40	GTTCCAAAGC AAGCCATTTT GATGGGACTT GCTAATATTC TTCAAGCCAA ACGTATCATT	4500
	TTACTCGCAT TTGGTGAAAA GAAACGTGCT GCTATTACAC ATTTATTAAA TCAGGAAATT	4560
	TCTGTTGATG TTCCAGCCAC ATTACTTCAC AAACACCCGA ATGTTGAGAT ATATTTAGAC	4620
45	GACGAAGCTT GCCCGAAAAA TGTTGCGAAA ATTCATGTCG ATGAAATGGA TTGATTGCAA	4680
	TGTTTAATTA AGAAATGCCT CGGGAAAGGT TCCAATAGAA AGATAAAAAG CATTGGAAGG	4740
50	ATGATTTTTA GTGGAATTAC AATTAGCAAT TGATTTATTA AACAAAGAAG ACGCGGCTGA	4800
	GTTAGCAAAT AAAGTAAAAG ATTATGTAGA TATCGTAGAA ATCGGTACGC CAATCATTTA	4860
	CAACGaAGGT TTACCAGCAG TTAAACATAT GGCAGACAAC ATTAGTAATG TAAAAGTATT	4920

55

CGCGGATGTA ATTACAATAC TAGGTGTTGC AGAAGATGCA TCAATTAAAG CAGCTATTGA 5040
 AGAAGCTCAT AAAAATAATA AACAATTACT AGTTGATATG ATTGCTGTTT AAGATTTAGA 5100
 5 AAAACGTGCA AAAGAACTAG ATGAAATGGG TGCTGATTAT ATTGCAGTAC ACACTGGTTA 5160
 TGATTTACAA GCAGAAGGGC AATCACCATT AGAAAGTTTA AGAACCGTTA AATCTGTTAT 5220
 TAAAAATTCT AAAGTTGCAG TAGCAGGTGG AATTAAACCA GATACAATTA AAGATATTGT 5280
 10 CGCTGAAAGT CCTGATCTTG TTATTGTTGG TGGCGGAATC GCAAATGCAG ATGATCCAGT 5340
 AGAAGCTGCG AAACAATGTC GCGCTGCAAT CGAAGGTAAG TAATATGGCT AAATTTAGTG 5400
 15 ACTATCAATT AATTCTAGAT GAATTAAAGA TGACTTTGTC ACATGTTGAA GCGGATGAGT 5460
 TTTCAACTTT TGCATCCAAA ATACTACATG CTGAACATAT ATTTGTAGCT GGCAAAGGAC 5520
 GTTCAGGATT CGTGGCGAAT AGTTTTGCAA TGCGCTTAAA TCAGCTCGGC AAACAGGCAC 5580
 20 ATGTTGTTGG AGAATCAACG ACACCTGCGA TTAAGTCGAA TGATGTATTT GTAATTATCT 5640
 CTGGTTCAGG TTCCACGGAA CATTTAAGAT TATTAGCAGA CAAAGCAAAA TCAGTAGGTG 5700
 CTGACATCGT ATTAATTACT ACAAATAAAG ATTCTGCAAT AGGCAATCTA GCTGGGACGA 5760
 25 ACATCGTTTT GCCTGCAGGT ACAAATATG ATGAACAAGG CTCGGCACAA CCATTAGGAA 5820
 GTTTGTTTGA ACAAGCATCT CAATTATTTT TAGATAGTGT TGTAATGGGA TTGATGACTG 5880
 AAATGAATGT TACGGAACAA ACGATGCAAC AAAATCATGC TAATTTAGAA TAAATAAAG 5940
 30 ATAGTCGATA ATATGATGCC TAGGCAGAAA TATTATCGAT TATTTTTTTA TTAAATAAT 6000
 AAATTATAGT ATAATATCAA TAATAAACGA ATAGGGGTGT TAATATTGAA GTTTGACAAT 6060
 35 TATATTTTGG ATTTTGATGG TACGTTGGCA GACACGAAAA AATGTGGTGA AGTAGCAACA 6120
 CAAAGTGCAT TTAAAGCATG TGGCTTAACG GAACCATCAT CTAAAGAAAT AACGCATTAT 6180
 ATGGSAAATAC CTATTGAAGA ATCATTTTTA AAATTAGCAG ACCGACCATT AGATGAAGCA 6240
 40 GCATTAGCAA AGTTAATCGA TACATTTAGA CACATATC AATCTATTGA AAAGGACTAT 6300
 ATTTATGAAT TTGCGGGTAT AACTGAAGCC ATTACAAGTT TGTATAACCA AGGGAAAAAA 6360
 CTTTTCGTGG TGTCTAGTAA GAAGAGTGAT GTATTAGAAA GAAATTTATC GGCTATTGGA 6420
 45 TTAAATCACT TGATTACCGA AGCTGTTGGA TCCGATCAAG TAAGTGCATA TAAACCAAAT 6480
 CCTGAAGGCA TACACACAAT TGTGCAACGC TACAATTTAA ATAGCCAACA AACGGTGTAT 6540
 50 ATTGGTGATT CAACGTTTGA TGTTGAGATG GCACAACGTG CTGGTATGCA ATCTGCAGCT 6600
 GTCACTTGGG GTGCACATGA TGCAAGGTCA TTACTTCATT CAAATCCGGA TTTTATTATT 6660
 AATGATCCAT CAGAAATTAA TACCGTATTA TAAACTTGT TAAACAGAG AATACCATGG 6720

55

	ATTTAAAATA AATATTTATT AAACATTATG AATTTTAAAGAGTAATGTC TGACTCGTTG	6840
	ATAATTTATT TTTGTAAAAA TAAATTAAAG TAATGACAAA GTTATTGAAG TAAATTGAGT	6900
5	ATAAACATTT AAATACGATG TCGAAAATGG CGATAGCATA TCACTTACAT GAAGTTGTGT	6960
	GctATCGCTA TTTTtagTTA TAATTCCAAA AAGTTAATCG TTCGATGATT TAAGAATTAT	7020
10	TATTGTTTAA TTCAAATGTA TGAGGGTATA AAATCATTGA ATTTAATTCTG ATAAAGCGAA	7080
	ATTTTTGAAC AAACATACTT TTGTATTTAT ATAAAAGTTT AAATCTTAT AAATTTGACA	7140
	AAACTAATTA ACTCCGTATA ATTATGAAAC ATACAAGAGG GAGTGTATGA ATTCATGGAT	7200
15	TTTAATAAAG AGAATATTAA CATGGTGGAT GCAAAGAAAG CTAAAAAAC CGTTGTTGCA	7260
	ACCGGTATCG GTAATGCAAT GGAATGGTTC GATTTTGGTG TCTATGCATA TAcAACTGCG	7320
	TACATTGGAG CGAACTTCTT CTCTCCAGTA GAGAATGCAG ACATTGACA AATGTTGACT	7380
20	TTCGCAGCAT TAGCCATTGC GTTTTTATTA AGACCAATTG GTGGTGTCTG ATTTGGTATT	7440
	ATTGGTGACA AATATGGACG TAAAGTTGTA TTAACATCTA CAATTATTTT AATGGCATT	7500
	TCAACATTAA CCATTGGATT ATTGCCAAGC TATGATCAAA TTGGACTTTG GGCACCAATA	7560
25	CTATTATTGC TTGCAAGAGT ACTACAAGGG TTTTCAACAG GTGGAGAGTA TGCGGGGGCA	7620
	ATGACATATG TTGCCGAATC ATCTCCAGAT AAGCGTCGTA ACTCATTAGG TAGTGGACTA	7680
30	GAAATTGGGA CATTATCAGG TTACATAGCT GCTTCAATTA TGATTGCTGT ATTAACATTC	7740
	TTTTTAACAG ATGAACAAAT GGCATCATT GGTGGAGAA TCCCATTCTT ACTCGGTTTA	7800
	TTCCTAGGAT TATTCGGCTT ATATTTACGT CGTAAGCTGG AAGAATCACC AGTTTTCGAA	7860
35	AATGATGTTG CAACACAACC AGAAAGAGAT AACATTAACT TTTTACAAAT CATCAGATTT	7920
	TATTACAAAG ATATATTTGT ATGTTTTGTA GCTGTTGTAT TCTTCaATGT TACAACTAT	7980
	ATGGTAACTG CATATTTACC AACCTATTTA GAACAAGTTA TTAAATTAGA TGCAACGACA	8040
40	ACAAGTGAT TAATTACTTG TGTATGGCA ATAATGATTC CATTAGCATT AATGTTTGGT	8100
	AAGTTAGCGG ATAAAATAGG TGAAAAGAAA GTATTTCTAA TTGGTACTGG TGGGCTAACA	8160
45	TTATTCAGTA TCATCGCATT TATGTTATTA CATTACAAT CATTGTTGT AATAGTAATC	8220
	GGTATATTTA TATTAGGATT TTTCTTATCA ACTTACGAAG CGACAATGCC AGGGTCGTTA	8280
	CCAACGATGT TTTACAGTCA TATAAGATAT CGAACTTTAT CAGTAACATT TAATATCTCT	8340
50	GTTTCGATAT TTGGTGGTaC GaCGCCATTA GTkGCAMCaT GGTtaGTTAC GAAAAC TGGA	8400
	GATCCATTAG CmCCTGCGTA TTATTTAACA GCAATCAGTG TTATTGGCTT TTTAGTTATT	8460
55	ACATTCTTAC ATTTAAGTAC AGCAGGAAAA TCTCTAAAAG GTTCGTATCC AAATGTAGAT	8520

	GAACGTAAGA ATTAGAGATT TTAATAAAAA GTATAAATCA ATCGTATATA AGCACTTTAA	8640
	AGCTAGTAGG TTCIGCTAAC TTAAAGTGC TTTTAAATT GAGAACTGTA ATTAGCCGTA	8700
5	ATAAAGTTTT TGTATATACA TAAACCCCCA CTGCAATGAT TATCGCAATG GGGGAAAGAG	8760
	GGGACTTAAA GCATATGTTT AGCTTTGAAT ACTTAAATT CTCTTGCTAT TGAAATGTTA	8820
	GGATGTAAAT ATGTCTTAGA GTATTTTGTC CAACGCAATT AATATTGAGA CTCTAACCTT	8880
10	CAATATTATT ATAGAGAACA CAAACTTAAA TAGATTGGGT GACTTATTTG TGTCAGTTAT	8940
	TGCGATTGCG ATAACCTCTT TTCTCTATAT ACATATAGTA ACGTCTTATC TAATAAAAAA	9000
	CATGGTACTA CAGTATCAAA TTTATCTAGG GCTTAAGTTT GATTTTTATA ATAGGCAGGT	9060
15	TTACCTGATA AAAATACTTA TTCATTATAT AATGTTAACA ATATGTATTT TAAAGTTTAC	9120
	ATTGAGTGAG GGATATTGAT GAACGTAATT TTAGAACAGT TGAAAACACA TACTCAAAAT	9180
20	AAACCTAATG ACATAGCATT ACATATCGAT GATGAAACAA TTACATATAG TCAACTAAAT	9240
	GCCCCATCA CTAGCGCAGT TGAATCTTTG CAGAAATATT CACTTAACCC TGTCGTTGCT	9300
	ATTAATATGA AATCACCAGT GCAAAGTATT ATTTGTTATT TAGCTTTGCA TCGTTTACAT	9360
25	AAAGTGCCTA TGATGATGGA AGGTAAATGG CAAAGTACTA TACATCGTCA ATTGATTGAA	9420
	AAATATGGTA TTAAAGATGT AATTGGAGAT ACAGGTCTCA TGCAGAATAT AGACTCACCG	9480
	ATGTTTATTG ATTCAACGCA ATTACAGCAC TACCCCAATT TATTACATAT TGGTTTTACT	9540
30	TCAGGGACAA CTGGACTGCC AAAAGCATAT TATCGTGATG AAGATTCATG GTTGGCTTCT	9600
	TTTGAAGTTA ATGAAATGTT GATGTTAAAA AATGAAAATG CAATAGCAGC CCCTGGACCA	9660
	CTATCGCACT CGTTAACATT ATATGCGTTA TTGTTTGCTT TAAGTTCCGG TCGTACTTTT	9720
35	ATAGGACAGA CCACTTTTCA TCCTGAAAAG TTAATTAATC AATGTCATAA AATATCATCA	9780
	TACAAAGTTG CTATGTTTCT TGTTCCAACG ATGATTAAAT CATTATTGTT AGTTTACAAC	9840
40	AATGAACATA CAATCCAATC ATTTTITAGC AGTGGAGATA AGCTGCATTC TTCTATTTTT	9900
	AAAAAGATAA AAAATCAAGC AAATGACATA AATTGATTG AATTTTTTGG TACATCGGAA	9960
	ACCAGTTTTA TCAGCTATAA CTTGAATCAG CAAGCACCAG TTGAATCAGT AGGTGTGCTA	10020
45	TTTCCAAATG TGGAAATTGAA AACCAACGAAT CACGATCACA ATGGTATAGG AACTATTTGT	10080
	ATAAAAAGTA ATATGATGTT TAGTGGCTAT GTAAGTGAAC AATGTATAAA TAATGATGAA	10140
	TGTTTGTGTT CTAATGATAA TGGCTATGTA AAAGAGCAGT ATTTATATTT AACGGGACGT	10200
50	CAACAGGATA TGTTAATTAT TGGTGGTCAA AATATATATC CAGCACATGT TGAACGCCTT	10260
	TTAACGCAAT CTTGAGCAT TGATGAAGCA ATTATCATCG GTATTCCAAA TGAGCGTTTT	10320

55

CAATTTTAA AAAAGAAAGT GAAaCgnTaT GAAATTCCAT CGATGATTCA TCATGTAGAA 10440
 AAGATGTATT ACACTGCAAG tGGTaaaATT GCTAGAGAAA AAATGATGTC GATGTATTTG 10500
 5 AGAGGTGAAT TATAATATGA ATCAAGCAGT CATAGTTGCA GCTAAACGAA CTGCATTGG 10560
 GAAATATGGT GGCACTTTAA AACATTTAGA GCCaGAACAA TTGCTTAAAC CTTTATTCCA 10620
 ACATTTTAAA GAGAAGTATC CAGAGGTAAT ATCTAAAATA GATGATGTAG TTTTAGGTAA 10680
 10 TGTGTGTTGGG AATGGTGGCA ATATTGCAAG AAAAGCATTG CTTGAAGCGG GGCTTAAAGA 10740
 TTCAATACCT GCGTCACAA TCGATCGGCA ATGTGGGTCT GGACTTGAAA GTGTTCAATA 10800
 TGCATGTCGC ATGATCCAAG CCGGAGCTGG CAAGGTATAT ATTGCAGGTG GTGTTGAAAG 10860
 15 TACAAGTCGA GCACCTTGGG AAATCAAACG ACCGCATTCT GTGTACGAAA CAGCATTACC 10920
 TGAGTTTTAT GAGCGTGCAT CATTTGCACC TGAAATGAGC GACCCATCAA TGATTCAAGG 10980
 20 TGCTGAAAAT GTGGCCAAGA TGTATGATGT TTCAAGAGAA TTACAAGATG AATTTGCTTA 11040
 TCGAAGTCAT CAATTGACAG CGGAAAATGT AAAGAATGGA AATATTTCTC AGGAAATATT 11100
 ACCTATAACC GTTAAAGGAG AAATATTCAA CACTGATGAA AGTCTAAAAT CACATATTCC 11160
 25 GAAAGATAAC TTTGGCCGAT TTAAGCCCGT GATCAAAGGT GGGACCGTTA CCGCTGCGAA 11220
 TAGTTGTATG AAAAATGATG GTGCAGTTTT ATTGCTTATT ATGGAAAAAG ATATGGCATA 11280
 CGAATTAGGT TTCGAGCATG GTTTATTATT TAAAGATGGT GTTACGGTAG GTGTTGATTC 11340
 30 TAATTTTCCT GGCATTGGTC CAGTACCAGC CATTTCCAAC TTAATAAAAA GAAATCAATT 11400
 AACGATAGAA AATATTGAAG TCATTGAAAT TAACGAAGCG TTCAGTGCAC AGGTAGTTGC 11460
 CTGCCAACAA GCTTTAAATA TTTCAAATAC GCAATTAAAT ATATGGGGTG GTGCATTAGC 11520
 35 ATCAGGTCAT CCATACGGTG CAAGCGGTGC CCAATTAGTG ACTCGATTAT TTTATATGTT 11580
 TGACAAAGAG ACTATGATTG CATCTATGGG GATAGGGGGA GGTCTAGGAA ATGCAGCATT 11640
 40 ATTTACTCGA TTCTAACCAG CGATTAAATG TGTCATTTTT TAAGGATAGT GTGGCTGCAT 11700
 ATTATCAGTG TTTTAACCAA CCTTATAGAA AAGAAGTACC ACCATTAATG TGTGCGTCAT 11760
 TATGGCCAAA ATTTGATTTA TTTAAAAAAT ATGCAAATAG CGAACTGATT TTAACAAAAT 11820
 45 CAGCAATTAA TCAAACCTCAA AAGATAGAAG TAGACACAAT ATATGTAGGG CATTTAGAAG 11880
 ATATTGAATG CCGACAGACT CGCAATATCA CACGTTATAC AATGGCTTTA ACATTAACCTA 11940
 AAAATGATCA ACATGTCATA ACGGTTACAC AAACCTTTTAT TAAGGCGATG AAGTAGAGAT 12000
 50 GGAGTTTAAT GAGATATGGA TAAATGAATA TTTGGCGCTC GTAAATGATG ATAATCCAAT 12060
 ACATAATGAG ATTGTGCCAG GACAATTAGT GAGTCAAATG ATGCTGATGG CTATGTCATT 12120

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATTCATTGAA CAACACGAAC ACGAAATTAT AGCAATTAAT GACGATGGAG AGATTAAAAT 12240
 AAAAATTTCT TTAGGCACAA AAAAATAACC GATATTAGCT GCATGAACGC ATATTAATTA 12300
 GGAGATGAAA GGACAGCTAA TATCAGTTAT GTATTGTTAT TATTATTGGG AACAGAGATG 12360
 AATATAGGTT ACGTTTCTTT CTTTGACCGG GGATGCATTA ATCTAAAATA ATAATAACAA 12420
 CTATATCAAT GTTTAATAAA TTCTGGATTA TTGGAACGAT TAGTCAATTT AACTAACTTT 12480
 CATATGATCT ATATCGTCTT GTAATAAAGA GAGCAATTTG AATATTTTCAG TATCACTAAA 12540
 TGAATCGTCA CATTTAATTG AAACATGCTG AAACGTTTTG GTTATAATTT CATAAACTGG 12600
 TGCGCCTTCA TGGTGATACT GTCGATAAAT AATCATAACC TATATTACCT CCTTTGCTAC 12660
 TCTATGGTTA TATTATAAAT AACATTTTAA TGTGTGACAT CAACCTTAAG TATCAACTTT 12720
 TTATCAGACA TAGAACGTAT GATTTACTAA GACTATTTAT GTATAAAAGT TCTAAATAAA 12780
 TATATATTTA TAGAGTCGCC TGGCAGTCAT TTGGGAATA TAACATATAT GATTAGAGAG 12840
 GCATCTATCG CAAAAGAATG ATAATGATAG AGGTATTGAG CATATAGATG AGTTTAAGTT 12900
 CATCTTGAAA ATAAAGGCTT ATTTAGTCAT AGATGTAGAT GTATAGGAAA TATTTGTATG 12960
 TATTGTTTCA TATGTATGAA ATTTTCAATA AAAGCTAATA ACGCTTATAT GTAACTTTCA 13020
 AATTTAAATT ATATACAGAG CATGATGATT ATAAAAAAT AACCACATCA CATAAATTGA 13080
 GTTCATACCC AATTTAAGTG GTGTGGCTAA TAATGTTGAT TTATAGATGA ACCGCCTAAT 13140
 CGTTAAACCT CTGTTACTTC AACATCGATA TGTTCAATAC GGTGTGATGC ACCGTGATCC 13200
 ACAGGACCAA CAAATCATT CATTTTCCAA CCGTTTTTAA TAGCAGAAGC GACGAAAGCT 13260
 TTCGCGCTAA TCACAGCTTC TTTGCGTGAC TTACCGTTAG CTAAATATGC AGTTGTTGCC 13320
 GCAGCAAATG TACAACCAGC ACCATGGTTA TAACTTTGTT GGAACATGTC TGTTGTTAGT 13380
 TGATAAAATG TTTGACCATC ATAGTATAAG TCATACGATT TATCTTGATC TAAAGCTTTG 13440
 CCACCTTTAA TGATGACATG CTGTGCGCCT TTATCAAAGA TAATTGTTGC AGCCTTTTTC 13500
 ATATCTTCAA TTGAATTTAA TTTACCTAAT CCTGATAATT GACCCGCTTC AAATAAGTTT 13560
 GGTGTCACCT CCGTTGCTTT AGGTAGTAAA TATTTAATCA TCGCCTCAGT ATTTCCAGGA 13620
 TTAAGCACTT CATCTTCGCC TTTACAAACC ATGACAGGAT CTACTACAAA ATATTGTGCA 13680
 TTAGATGCCT CATATACTTC TCCAGCACGT TTGATTATCT CCTCAGTACC TAACATACCT 13740
 GTTTTAATAG CATCAGGTCC GATTGATAAA GCCGTTTCAA GTTGTTTTTC AAATACATCC 13800
 ATTGGTAATG GTGTAACATC GTGTGACCAT GTATCTTTAT CCATAGTAAC GATGGCAGTT 13860
 AAAGCGACCA TGCCATACGT ATCTAATTCT TGGAACGTTT TCAAATCTGC TTGCATACCT 13920

	CACTCCTACA TAATAATATT GTATTCATCA TATCATTTTT AACCTAATTG AAAAATATTA	14040
	AGCATTCAAT ATTTGATGAT TGTGAAATG AATCATTCAT ACTATTGTAA CTTTGTAAAA	14100
5	TGTCATTAC TTTAGATAAG TGTGATATGT TAAAATATGT CCTGAGGTGA GATTGAATGG	14160
	AATGGTCGCA AATTTTTTCAT GACATAACAA CGAAACATGA CTTTAAAGCT ATGCATGATT	14220
	TTTTAGAAAA AGAATATTCC ACTGCAATCG TATACCCTGA TAGGGAAAAT ATATATCAAG	14280
10	CGTTTGATTT AACACCGTTT GAAAATATCA AAGTTGTTAT ATTAGGACAA GACCCGTATC	14340
	ATGGTCCAAA CCAAGCACAT GGATTAGCAT TTTCAGTGCA ACCTAACGCA AAATTCCCTC	14400
15	CATCTTTACG TAATATGTAT AAAGAATTAG CAGATGATAT TGGATGCGTT AGACAAACAC	14460
	CGCATTTACA AGATTGGGCA AGAGAAGGCG TCTTGTTATT GAATACAGTT TTAACCGTAA	14520
	GACAGGTGA AGCAAATTCT CATCGTGATA TTGGTTGGGA AACATTTACT GATGAAATTA	14580
20	TTAAAGCAGT GTCTGATTAT AAAGAACATG TTGTCTTTAT TTTGTGGGGG AAACCTGCAC	14640
	AGCAAAAAAT AAAGCTTATC GATACATCTA AACATTGTAT TATAAAATCA GTGCATCCTA	14700
	GTCCACTGTC TGCATATAGA GGATTCCTTG GATCAAAACC GTATTCCAAA GCGAATGCCT	14760
25	ATTTAGAGTC AGTAGGAAAA TCACCAATTA ATTGGTGTGA AAGTGAGGCG TAGATGTTGA	14820
	ATAGAGAAAC TTTAATAGCA CGAATTGAGC AAGAATTAGT ACAAGCAGAG CAGGCACAGC	14880
	ATGACCATGA CTTTGTAAAA CATATGTATG CCATACATAT ATTAACATCT TTATATGCTT	14940
30	CAACATCAAA TACACCACAT ATTGGTGAAC AACAAATGAA TCGTCGTATT GCTAACCATA	15000
	ATCAAATGCC ACAATCACAA ATAACGCAGC CAACTCATCA AGTGACAGTT GCTGAAATTG	15060
35	AAGCGATGGG TGGTAAAGTA AATACGCATT CAGCACATCA TCATAATAAG TCATATTCAC	15120
	AACCTTCAAA CCAACAACAA AGATTAGCGA CAGATGATGA CATTGGCAAT GGTGAATCCA	15180
	TATTGATTT TTAATAAGCA ACAATGAAAC ATAATTACTT AATAGCTTGT TAAGTATGTA	15240
40	GGTTAATAAT CAAGACGCAT ATACTTTTAT TCGAGTGTTT GGATTTAAAC ATTTATTAAT	15300
	ACTGAATTAT ATAAGGAGAG GTAGCAATGA AATTATTTAT TATTTTAGGT GCATTAAACG	15360
	CGATGATGGC TGTCGGTACA GGTGCATTG GTGCGCATGG TTTACAAGGA AAAATAAGTG	15420
45	ATCACTATTT ATCAGTATGG GAAAAAGCAA CGACGTATCA AATGTACCAT GGCTTAGCAT	15480
	TATTAATTAT AGGTGTAATT AGTGGTACAA CTTCAATCAA TGTTAACTGG GCTGGCTGGT	15540
	TAATATTTGC TGGTATTATT TTCTTTAGTG GATCATTATA TATTTTAGTA TTAACCTCAA	15600
50	TTAAAGTTTT AGGTGCGATT ACGCCAATTG GTGGCGTATT GTTCATCATT GGATGGATAA	15660
	TGTTAATCAT TGCGACATTC AAATTTGCTG GTTAAATTTT AAACTTTAG ATTACCTATG	15720

55

	TGGGTATAGA ATACCTTCGA GGTGAGTTTT TATTTATGGA AAAAAAGAAT AAGCAAATAG	15840
	ATAGAGGCGA TTTAAAACAA AACCTATCTG AAAAGTTTGT ATGGGCGATT GCATATGGTT	15900
5	CATGTATCGG ATGGGGCGCA TTCATCTTAC CAGGAGACTG GATTAAGCAG TCAGGTCCGA	15960
	TTGCAGCATC AATTGGTATA GTTATTGGTG CATTATTAAT GATATTAATT GCGGTTAGTT	16020
	ATGGCGCATT AGTAGAGAGA TTTCCAGTAT CAGGGGGCGC GTTTGCCTTT AGTTTCTTAA	16080
10	GTTTCGGCAG ATATGTGAGT TTCTTCTCAT CATGGTTTTT AACTTTTGGT TATGTCTGTG	16140
	TCGTTGCTTT AAAtGCGACC GCATTCAAGT TACTAGTTAA ATTCTTATTG CCAGATGTCT	16200
	TAAATAATGG GAAACTATAC ACCATTGCGG GCTGGGACGT TTATATTACG GAAATCATT	16260
15	TTGCGACCGT ATTACTACTT GTATTCATGC TAGTAACGAT TCGTGGCGCA AGTGTATCTG	16320
	GATCATTACA ATATTATTTT TGTGTGGCGA TGGTAATCGT CGTATTATTG ATGTTCTTTG	16380
20	GTTCACTCTT TGGTAATAAT TTTGCACTTG AAAATTTACA ACCGTTAGCT GAACCTAGCA	16440
	AAGGATGGTT AGTGTCTATT GTGGTTATTG TATCCGTGGC ACCATGGGCA TATGTTGGAT	16500
	TTGATAATAT TCCACAAACA GCAGAAGAGT TTAACTTTGC ACCAAACAAG ACATTTAAGC	16560
25	TTATCGTGTA CAGTTTATTA GCAGCATCAT TAACTTATGT TGTCATGATT TTATACACTG	16620
	GTTGGTTATC AACAAATCAT CAAAGTTTAA ATGGGCAGTT GTGGTTAACA GGTGCTGtTA	16680
	CACAAACAGC ATTTGGTTAT ATTGGATTAG GTGTATTAGC AATTGCAATT ATGATGGGTA	16740
30	TATTTACTGG TTTAAATGGA TTCTTGATGA GTTCAAGTCG CTTGTTATTT TCTATGGGAC	16800
	GTTCAAGTAT TATGCCAACA ATGTTTAGTA AATTACATAG TAAATACAAA ACACCATATG	16860
	TCGCAATCAT ATTCCTAGTA GGAGTGTGCT TAATTGCACC TTGGCTAGGA AGAACTGCAT	16920
35	TGACTTGGAT TGTAGATATG TCATCTACTG GTGTATCCAT TGCCTACTTT ATTACATGTT	16980
	TGTCTGCAGC GAAATTATTC AGTTATAACA AACAAAGTAA TACGTATGCA CCGGTTTACA	17040
40	AAACGTTTGC TATTATCGGC TCATTTGTAT CATTCATTTT CTTAGCGTTG TTATTAGTGC	17100
	CAGGTTCTCC TGCAGCACTG ACTGCACCGT CTTATATTGC ATTACTTGGG TGGTTAATCA	17160
	TCGGTTTAAT ATTCTTTGTG ATTCGATATC CTAAATTGAA AAATATGGAT AATGATGAAT	17220
45	TAAGTCGCTT GATTTTAAAT AGAAGTGAAA ATGAAGTTGA TGATATGATT GAAGAACCTG	17280
	AAAAAGAAAA AACTAAATAA TAAAAGAATC GCACAATAAA CCTTCTTCAT TCGGAGGCGT	17340
	ATCGTGCGAT TTTTGTATT ATAAATTGAC ATTTAAGACG AGGCAGCTGA ACCTTATATA	17400
50	TAATTGCTAA GAGTTAGGGC TGAGCCATTT CTAACAAATA TTTATAATCG TTTAAAAGAT	17460
	TTCACGAACC CAGAAACAAT TAATTGGGAA ATTTGGTCCG CGAATAATAA ACCTAATGCG	17520

55

AAGACTAAAT TTTTGTAGC ATCGTATGCT AAGCCACCAG GTACTAATGG AATGATACCC 17640
 GTTACCATAA AAATGATGGC AGGTTCTTTT TGTTTACGAG CCATATAATG ACTTAACAAG 17700
 5 CCTAATGCTA AACTACCAA GAAACTAGAG TATATAGTGT GCACATTAAA GCCGTTGAAG 17760
 AATAAGGTGT AAACCATCCA TCCACACGTA CCAACGAAAC CACATGATAG ATATAATTTT 17820
 CTAGGTGCAT CAAAAATGAC GCAGAA 17846

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5544 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

ATTGACACTT GGTGAAAGTA ATATCGCCGC GCTATTTTGG CAAAATGGAC ACTTAGAACC 60
 TGAGTTACAA GATGAACAGC CAATTAATAT ATTAGGATCT GkTCAAATCA ACGAATGGAA 120
 25 TGGAATCAA TCACCGCAA TAATTATTCA AGATATTGCG ATGAATGAAC AGCAAATATT 180
 AGATTATAGA AGTAAGCGAA AAAGTTTACC TTTTACAGAA AATGATGAAA ATATTGTCGT 240
 GCTTATTCAT CCTAAAAGTG ATAAAGTAAA TGCGAATGAA TATTATTATG GTGAAGAAAT 300
 30 TAAACAACAA ACTGATAAAG TAGTATTAAG AGATTTACCA ACGTCAATGG AAGACTTGTC 360
 TAATTCCTTG CAACAACCTGC AATTTTCTCA ACTTTATATA GTTTTGCAAC ATAATCAITC 420
 GATTTACTTC GATGGTATAC CTAATATGGA TATTTTTAAA AAGTGTTATA AAGCATTAA 480
 35 AACTAAACAA GAAACAAATA TCCAGAAAGA GGGTATGTTA TTGTGTCAAC ATTTAAGTGT 540
 GAAACAGAT ACACTTAAAT TCATGTTGAA AGTTTCTTA GACTTAAAT TTGTAACACA 600
 40 AGAAGATGGT TTAATTCGAA TCAATCAACA ACCTGATAAA AGATCGATTG ATTCCAGCAA 660
 AGTATATCAA TTAAGACAAC AACGTATGGA TGTGAAAAG CAATTATTAT ATCAAGATTT 720
 TTCAGAAATA AAAAATTGGA TAAAGTCACA ATTGTCGTGA GCAATTTAGG AGGAAATATT 780
 45 AATGGATTTA AAGCAATACG TATCAGAAGT TCAAGATTGG CCGAAACCAG GTGTTAGTTT 840
 CAAGGATATT ACTACAATTA TGGATAATGG TGAAGCATAT GGCTATGCAA CAGATAAAAT 900
 TGTAGAATAC GCAAAAGACA GAGATGTTGA TATCGTTGTA GGACCTGAAG CGCGTGGCTT 960
 50 TATCATTGGC TGTCTGTAG CTTATTCAAT GGGGATTGGC TTTGCACCTG TTAGAAAAGA 1020
 AGGGAAATTA CCTCGTGMAG TCATTGTTA TGAGTATGAC CTAGAATATG GTACAAATGT 1080

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATTAGCTACT GGTGGTACGA TTGAAGCAGC AATAAAATTA GTTGAAAAAT TAGGCGGTAT 1200
 CGTAGTAGGT ATTGCATTTA TAATTGAATT GAAATATTTA AATGGTATTG AAAAAATTAA 1260
 AGATTACGAT GTTATGAGTT TAATCTCATA CGACGAATAA TAAATAATAT AATTTTATCA 1320
 AATGAAATCC TTCATCAAAT GTATAAGAAC CAATGACTTA ATTAAAAAAG TTGTTTAAGT 1380
 TTTCTTAACA TGAGATGTTA GGATTTTTTA TTTACTGAAA ATGTTAGATG ATTGAGCATT 1440
 ATACCTTAAT AACATCGTTT ATTTATTTCA TAAATTGTAG TATCATAGAA CTAATATTTA 1500
 AAAAATGAAA CAGTAGATTT AGGTCGAATT TTTGTAAAAG TTTTAAAAGT AGGAATAGTA 1560
 TACAAATTAA ACTCGCTCAA GTAAAATTAA TATTACGATT AATGACGACA GGATAAATAT 1620
 TTATCGTCGA CGGACGTATG ATTGGTGTGG GACAAATACT ATTCAACAAG AGTACCTAAA 1680
 TCATTGTTTA AGGCGAAGTA ATAAATATGA ATGGGGTGTA TCATATAATG AACAACGAAT 1740
 ATCCATATAG TGCAGACGAA tCTTCACAA AGCAAATCA TATTTGTCAG CAGATGAATA 1800
 TGAGTATGTT TTA AAAAGCT ATCATATTGC TTATGAAGCA CATAAAGGTC AGTTCCGAAA 1860
 AAACGGATTA CCATACATTA TGCATCCTAT ACAAGTTGCA GGTATTTTAA CAGAAATGCG 1920
 ATTAGACGGA CCGACGATTG TCGCAGGTTT TTTGCATGAT GTAATTGAAG ATACACCGTA 1980
 TACATTTGAA GATGTAAAAG AAATGTTCAA TGAAGAAGTT GCTCGAATTG TTGATGGTGT 2040
 GACGAAGCTT AAAAAAGTAA AATACCGCTC AAAAGAAGAA CAACAAGCTG AAAATCATCG 2100
 CAAGTTATTT ATTGCGATTG CCAAAGATGT ACGCGTAATT TTGGTGAAAT TAGCAGACAG 2160
 ATTACATAAT ATGCGTACCT TGAAAGCCAT GCCGCGCGAA AAACAAATTA GAATTTCTCG 2220
 AGAAACATTA GAAATTTATG CACCATTAGC ACATCGTCTT GGTATTAATA CAATCAAATG 2280
 GGAAC TAGAA GATACGGCTC TTCGTTATAT TGATAATGTG CAATATTTTA GAATAGTCAA 2340
 TTTAATGAAG AAGAAACGTA GTGAACGTGA AGCGTATATC GAAACGGCTA TTGATAGAAT 2400
 ACGTACTGAA ATGGACCGAA TGAATATCGA AGGCGATATA AATGGTAGAC CTAAACATAT 2460
 TTACAGTATT TATCGGAAAA TGATGAAGCA GAAAAACAA TTTGATCAAA TTTTGTATTT 2520
 GTTGGCGATA CGTGTATTG TCAATTCTAT TAATGATTGT TATGCGATAC TTGGGTTGGT 2580
 GCATACGTTA TGGAACCGA TGCCAGGACG TTTTAAAGAT TATATTGCAA TGCCTAAACA 2640
 AAATTTGTAT CAGTCATTGC ATACTACAGT AGTAGGCCCA AATGGAGACC CGCTCGAAAT 2700
 CCAATACGA ACGTTTGATA TGCACGAAAT TGCTGAGCAT GGTGTTGCAG CACACTGGGC 2760
 TTACAAAGAA GGTAAAAAAG TAAGTGAAAA AGATCAAAC TATCAAAATA AGTTAAATTG 2820
 GTTAAAGAA TTAGCTGAAG CGGATCATAC ATCGTCTGAC GCTCAAGAAT TTATGGAAAC 2880

	TGAGTTGCCA TATGGTGCTG TGCCGATTGA TTTTGCTTAT GCGATTCCACA GTGAAGTAGG	3000
	TAATAAGATG ATTGGTGCCA AGGTGAATGG CAAAATTGTA CCAATTGACT ATATTTTACA	3060
5	AACAGGCGAT ATTGTTGAAA TACGTACTAG TAAACATTCA TATGGACCAA GTCGTGATTG	3120
	GTTGAAAATT GTTAAATCGT CTAGTGCCAA AGGTAAAATT AAAAGTTTCT TCAAAAAACA	3180
	AGATCGTTCA TCTAATATTG AAAAAGGCCG AATGATGGTT GAAGCTGAAA TAAAAGAGCA	3240
10	AGGATTTAGA GTCGAAGATA TTTTGACAGA GAAAAATATT CAGGTTGTTA ATGAAAAATA	3300
	TAACTTTGCA AATGAAGATG ATTTATTTCG AGCTGTAGGA TTTGGCGGCG TGACATCCTT	3360
15	ACAGATTGTT AATAAATTAA CTGAAAGACA ACGTATTTTA GATAACAAC GTGCTTTAAA	3420
	TGAAGCACAA GAAGTTACGA AATCATTGCC TATTAAAGAC AACATCATT CTGATAGTGG	3480
	TGTCTATGTA GAAGGTTTAG AAAATGTACT TATCAAGTTG TCAAAATGTT GTAATCCTAT	3540
20	ACCaGGTGAT GATATTGTAG GTTATATCAC CAAAGGTCAC GGTATTAAAG TACATCGCAC	3600
	TGATTGCCCA AATATTAAGA ACGAACTGA ACGACTAATT AATGTTGAAT GGGTAAATC	3660
	AAAAGACGCA ACTCAAAAAT ATCAGGTTGA TTTAGAGGTA AtGCGTATGA CCGAAATGGC	3720
25	TTGTTGAATG AAGTACTACA AGCTGTTAGC TCGACAGCCG GCAATTTAAT TAAAGTTTCA	3780
	GGACGTTTCA ATATTGATAA AAATGCAATA ATAAATATTA GTGTCATGGT GAAAAACGTG	3840
	AATGATGTTT ATCGTGTGGT AGAAAAGATC AAACAATTG GTGATGTTTA TACAGTAACA	3900
30	AGAGTTTGGA ACTAGAGGTG CAAAATATGA AAGTAGTTGT ACAAAGAGTT AAAGAAGCAT	3960
	CGGTGACGAA TGATACATTA AATAATCAAA TCAAAAAGG ATATTGTTTA TTAGTCGGTA	4020
35	TCGGTCAGAA CTCTACAGAG CAAGATGCAG ATGTAATTGC AAAGAAAATT GCTAATGCAA	4080
	GATTATTTGA AGATGACAAT AATAAATTAA ACTTTAATAT CCAACAAATG AATGGTGAAA	4140
	TACTATCAGT TTCACAATTT ACTCTCTATG CAGATGTAAA AAAAGGTAAC CGTCCAGGTT	4200
40	TCTCAAATTC TAAAAATCCT GATCaAGCGG TAAAAATTTA TGAGTATTTT AATGCaTGCG	4260
	CTACGAGCGT ATGGTCTTAC TGTGAAAACA GGTGAATTTG GAACACACAT GAATGTTAGC	4320
	ATAAATAATG ATGGTCCAGT CACTATTATT TATGAAAGTC AGGACGGCAA AATTCAATGA	4380
45	AAAAATAGA GGCATGGTTA TCTAAAAGG GTCTTAAAAA TAAACGTACT CTAATAGTAG	4440
	TGATTGCCTT TGTCTTATTT ATCATCTTTT TATTTTTATT GCTGAATAGC AATAGTGAAG	4500
	ATAGTGGGAA CATCACGATA ACTGAAAATG CTGAATTACG TACAGGTCCA AACGCTGCGT	4560
50	ATCCAGTCAT ATATAAGTT GAAAAAGGTG ACCATTTTAA AAAGATTGGT AAAGTAGGTA	4620
	AATGGATTGA AGTTGAAGAT ACATCCAGTA ATGAAAAGG TTGGATAGCT GGATGGCACA	4680

55

TAGTGCTTGA TCCTGGTCAT GGAGGTAGTG ACCAGGGTGC TTCAAGCAAT ACTAAATATA 4800
 AAAGTTTGA AAAAGATTAT ACGTTGAAAA CAGCAAAAGA ATTGCAGCGT ACTTTAGAAA 4860
 5 AAGAAGGCGC AACTGTTAAG ATGACAAGAA CAGACGATAC ATATGTTTCA CTAGAAAATC 4920
 GTGATATCAA AGGCGATGCC TATTTGAGTA TACATAATGA TCGGTTAGAA TCATCTAATG 4980
 CAAATGGAAT GACaGTTTAT TGGTATCATG ATAATCAAAG AGCTTTAGCA GATACGTTAG 5040
 10 ACGCTACGAT TCAGAAGAAA GGTCTACTTT CTAATCGCGG TTCAAGACAA GAAAATTATC 5100
 AAGTGTTAAG ACAAACAAAA GTTCCTGCTG TTTTATTAGA ATTAGGTTAT ATTAGTAACC 5160
 CAACTGATGA AACGATGATT AAAGATCAAT TACATAGACA AATTTTAGAA CAAGCAATTG 5220
 15 TTGATGGCCT TAAAATTTAT TTTTCTGCGT AGGGCTTGCA AAAATATGTG AAAGTAGTTA 5280
 TCATTGATAT TGAATTTTAT AACTAAAACC GTTAGTATTC TTGAAATGGT AAATGAAATA 5340
 20 GGTAGCAATC TAACTAAGAT TGTGTAGGAA TATAATCCAT AGACTGAAAG ATTATGCTGA 5400
 GTAGTTTATA TACATTGAAC ACAAGAAGAG GTGCTTTATG AAAAGTAAAG CCGTTAAACG 5460
 TACGTTaAAC GTTTTGAGTG GGTTTATTAA ATGCACGCTT ATAAAAAGTA ATGATGATTA 5520
 25 CAATTAGGCA TGTTTTTTTAA ACCA 5544

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1067 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

AAAAGATTGC AAATATAAAT GGCATGTTTA ATATGTTAGA ACAACAAATC ATTCATAGCC 60
 40 AAGATATGGC TCATTTTGA AGTGAATTTT TTTACGTCAA TCATGaGCAT CGAGAAAATC 120
 ATGAaGCACT CCTAATTTAT TACAAAAATA GTATCGACAA TCCTATTGTA GATGGTGCAT 180
 GTTATATTTT AGCCCTACCT GAAATTTTCA ATAGTGTGTA TGTTTTCGAA TCAGAGTTAC 240
 45 CATTTTCATG GGTATATGAT GAAAATGGCA TTACCGAAAC AATGAAATCA CTTAGCATTG 300
 CATTACAATA TTTAGTTGCA GCAGCTTTAG AAGTAACTGA TGTGAATATA TTTAAGCCTT 360
 CAGGATTTAC AATGGGAATG AATAATTGGA ATATTGCTCA AATGCGAATC TTTTGGCAAT 420
 50 ATACAGCAAT TATTAGAAAA GAAGCACTAT AACATTAATA ATTAATTAGC TATAAAGATG 480
 ATTCACAACA ATCATCTTTA TAGCTTTTTT ATGTCTAATT ATTTTGTAGG AAAATmACAA 540

AATTTTATGT TTTCAAAAGT AAACAATCAA AAGATGTTAG AAGATTGCTT CTATATAAGA 660
 AAGAAAGTGT TTGTAGAAGA ACAAGGCGTC CCTGAGGAAA GTGAAATTGA TGAATATGAA 720
 5 TCTGAATCTA TTCACCTCAT TGGATATGAT AATGGACAGC CAGTTGCCAC TGCTCGAATA 780
 CGCCCTATTA ATGAAACAAC TGTCAAATA GAACGAGTAG CTGTGATGAA ATCACATCGT 840
 GGACAAGGAA TGGGTAGAAT GCTTATGCAA GCTGTAGAAT CATTAGCTAA AGATGAAGGT 900
 10 TTTTACGTAG CTAATATGAA TGCCCAATGT CATGCTATCC CATTTTATGA AAGTTTAAAC 960
 TTTAAATGA GAGGTAATAT ATTTCTTGAG GAAGGCATCG AGCATATTGA AATGACAAAA 1020
 AAGTTAACCT CGCTTAATTA AAAAAAGTTG TATCTATTTT AGAAACA 1067
 15

(2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18613 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

AAGACGATG ATAACAACAA TACGTGTAGT GAAAGATTTT AATCTACATA TTACTGACAA 60
 AGAATTCATT GTATTTGTTG GACCATCGGG ATGTGGTAAA TCAACAACAT TACGAATGGT 120
 30 TGCTGGACTA GAGTCTATCA CATCTGGAGA TTTTATATT GATGGGGAAC GCATGAACGA 180
 TGTGAACCA AAGAATAGAG ATATTGCGAT GGTATTTCAA AACTATGCAT TATATCCACA 240
 TATGACTGTT TTTGAAATA TGGCATTG GCTAAAGCTA CGTAAAGTAA ATAAAAAGA 300
 35 GATTGAACAA AAAGTTAATG AAGCAGCTGA AATATTAGGA TTAAGTGT ATCTTGGTCG 360
 TAAACCAAAA GCGTTATCTG GCGGACAGCG TCAACGTGTT GCTTTGGGCA GAGCTATTGT 420
 TAGGGATGCG AAAGTCTTTT TAATGGATGA ACCATTATCG AATCTTGATG CGAAYTtCGA 480
 40 GTACAAATGC GCACAGAAAT ATTGAAATTA CATAAGCGAC TTAATACTAC GACAATTTAT 540
 GTTACACATG ATCAAACTGA AGCATTGACG ATGGCTAGTC GAATTGTTGT TTTGAAAGAT 600
 45 GGCGACATTA TGCAAGTCGG CACACCTAGA GAAATATATG ATGCCCTAA TTGCATATT 660
 GTGGCGCAAT TTATCGGCTC ACCAGCAATG AATATGTTGA ATGCTACAGT TGAAATGGAC 720
 GGATTGAAGG TAGGAACACA CCATTTTAAA TTACATAATA AAAAATTGA AAAGTTAAAA 780
 50 GCTGCTGGCT ACTTAGACAA GGAAATTATT TTAGGTATTC GAGCTGAAGA CATTATGAA 840
 GAACCAATAT TTATTCAAAC TTCTCCAGAG ACACAATTG AATCTGAAGT AGTTGTATCC 900

	AAATTAGATT CAAGAACTCA AGTGATGGCG AACGACAAGA TTACACTAGC ATTTGATATG	1020
	AATAAGTGTC ACTTTTTTGA TGAAAAACA GGAAATCGTA TCGTCTAAGG GGGAGTATTC	1080
5	ATGTCTAAAA TTTTAAATG TATCACGTTA GCCGTGGTAA TGTATTAAAT CGTAACTGCA	1140
	TGTGGCCCTA ATCGTTCGAA AGAAGATATT GATAAAGCAT TGAATAAAGA TAATTCTAAA	1200
10	GACAAGCCTA ACCAACTTAC GATGTGGGTG GATGGCGACA AGCAAATGGC GTTTTATAAA	1260
	AAAATTACGG ATCAATATAC TAAAAAACT GGCATCAAAG TAAAGCTTGT AAATATTGGT	1320
	CAAAATGATC AACTAGAAAA TATTTTCGCTA GACGCTCCTG CAGGAAAAGG TCCAGATATC	1380
15	TTTTTCTTAG CACATGATAA TACTGGAAGT GCCTATCTAC AAGGCTTAGC TGCTGAAATC	1440
	AAATTATCAA AAGATGAGTT GAAAGGTTTC AATAAGCAAG CACTTAAAGC GATGAATTAT	1500
	GACAATAAGC AACTAGCATT GCCAGCTATC GTTGAAACAA CCGCACTTTT TTATAATAAA	1560
20	AAATTAGTGA AAAATGCACC GCAAACGTTA GAAGAAGTTG AAGCTAATGC TGCCAAACTA	1620
	ACTGATAGTA AAAAGAAACA ATACGGTATG TTATTTGATG CTAAAAATT CTATTTTAAT	1680
	TATCCGTTTT TATTCGGCAA TGATGATTAT ATTTTCAAGA AAAATGGCAG TGAATATGAT	1740
25	ATTCATCAGC TAGGACTAAA TTCAAAACAT GTCGTCAAGA ATGCTGAACG ATTACAAAAA	1800
	TGGTACGACA AAGGGTATCT TCCTAAGGCA GCAACACATG ATGTCATGAT TGGTCTTTTT	1860
	AAAGAAGGAA AAGTAGGACA ATTTGTCACT GGACCGTGGG ACATTAATGA ATATCAAGAA	1920
30	ACGTTTGATA AAGATTTAGG AGTAACAACA TTACCTACAG ATGGTGGCAA ACCTATGAAA	1980
	CCATTTCTAG GTGTACGTGG TTGGTATTTA TCTGAATATA GTAAACATAA GTATTGGGCT	2040
35	AAAGATTTAA TGCTGTATAT CACTAGTAAA GATACATTAC AAAAATATAC AGATGAAATG	2100
	AGCGAAATTA CTGGACGTGT TGACGTGAAA TCATCTAATC CAAATTTAAA AGTGTGTTGAA	2160
	AAGCAAGCAC GTCATGCTGA ACCGATGCCT AATATTCCTG AAATGCGACA AGTTTGGGAA	2220
40	CCGATGGGCA ATGCAAGCAT ATTTATTTCA AATGGTAAGA ATCCTAAACA AGCGTTAGAT	2280
	GAGGCGACGA ATGATATAAC GCAAAATATT AAGATTCCTC ATCCATCACA AAATGATAAG	2340
	AAAGGAGATT AGTTATGACG AAACGTAACC CTAAATTAGC GGCATTATTA TCTGTTATAC	2400
45	CTGGTTTGGG ACAGTTTTAT AATAAAAGAC CCATTAAAGG GACGATATTT TTTATCTTTT	2460
	TCATCAGTTT TATTTCTGTT TTTTATAGCT TTTTAAATAT TGGTTTTTGG GGATTGTTCA	2520
	CATTAGGGAC AGTACCTAAG TTAGACGATT CTCGTGCTTT ACTTGACAAA GGTATTATTT	2580
50	CTATCTTACT CGTTGCTTTC GCAATCATGC TATATATCAT TAATATTTTA GATGCATATC	2640
	GTAATGCTGA ACGATTTAAT CGCAATGAGG AAATAAAGGA TCCGAAGCGC GTATGGTGGC	2700

55

	TGTAGTTGTA TTTCCATTAA TAYyTATGTT TGGAGTAGCA TTTACAAATT ACAATTTATA	2820
	CAACGCGCCT CCGAGACACA CATTAGAATG GGTGGTTTA GATAACTTTA AAACGTTATT	2880
5	CACAATTGGC GTTTGGCGTA AAACATTTTT CAGTGTATT ACTTGGACAT TAGTATGGAC	2940
	GCTTGTGCA ACGACACTTC AAATTGCATT AGGGCTGTT TTGGCAATTA TTGTAAATCA	3000
	CCCTGTCGTC AAAGGTAAGA AATTTATCCG TACTGTGTTA ATCCTACCTT GGGCTGTACC	3060
10	ATCATTTGTG ACAATTTTAA TATTTGTAGC GTTATTTAAT GATGAATTTG GTGCGATAAA	3120
	TAATGATATT TTGCAACCTT TATTAGGTGT AGCACCAGCA TGGTTAAGTG ATCCGTTTGT	3180
	GGCAAAAGTG GCATTAATCG GCATTCAAGT ATGGCTTGA TCCCATTTG TCTTGTCACT	3240
15	GTTCACTGGA GTACTGCAAA GTATTTATC AGATTGGTAC GAAGCAGCAG ATATGGATGG	3300
	TGCGTCTAGT TGGCAAAAGT TTAGAAACAT CACATCCCG CATGTCAATT ACGCCACAGC	3360
20	GCCATTGTTA ATTATGCAAT ATGCAGGTAA TTTCAATAAT TTTAATCTTA TTTATCTATT	3420
	TAATAAAGGC GGTCCACCAG TGTCAAGGCA GAATGCTGGT AGTACAGATA TCTTGATATC	3480
	TTGGGTGTAT AATCTGACAT TTGAGTTTAA CAACTTCAAC ATGGGTGCAG TTGTGTCAAT	3540
25	AATTATTGGA TTTATTGTG CTATTGTGCG ATTTATTCAA TTCAGACGTA CAAGTACGTT	3600
	TAAAGATGAG GGAGGTTTAT AAGATGACAA AGAAGAAAAA CATATTAAAA GCAATCGGTA	3660
	TTTACAGTTT TATAGCGATG ATGTTTGTCA TCATTTTATA TCCACTACTG TGGACATTTG	3720
30	GCATTTCCCT TAATCCAGGT ACGAACTTGT ATGGTGCCAA AATGATACCA GACAATGCAA	3780
	CATTTAAAAA TTATGCATTC TTACTATTCTG ATGACAGTAG TCAATACCTG ACTTGGTATA	3840
	AAAATACGCT TATCGTAGCA TCTGCAAATG CACTGTTTAG TGTGATATTT GTCACGTTAA	3900
35	CAGCATATGC TTTTCTAGA TATCGCTTGT TTGGTCGTAA ATACGGGCTG ATTACATTTT	3960
	TGATTTTACA AATGTTCCCT GTATTAATGG CAATGGTCGC AATCTATATT TTGCTAAATA	4020
40	CAATTGGATT ATTAGATTCT TTATTTGGAC TAACACTGGT ATATATTGGT GGATCAATAC	4080
	CGATGAATGC CTTTTTAGTG AAAGGTTACT TCGATACGAT TCCAAAAGAA CTTGATGAAT	4140
	CTGCCAAAAT TGATGGTGCA GGCATATGC GTATTTTCTT ACAAATTATG CTTCCATTAG	4200
45	CTAAGCCGAT TTTAGCAGTT GTTGCTTTGT TCAATTTTAT GGGGCCATTT ATGGACTTTA	4260
	TATTACCTAA AATACTATTA AGAAGTCCTG AAAAATTCAC ATTAGCAGTT GGATTGTTCA	4320
	ACTTTATTAA TGATAAGTAT GCAAATAATT TCACAGTGT TGCAGCAGGG GCAATTATGA	4380
50	TTGCAGTACC TATAGCAATC GTATTCTTGT TCTTGCAACG CTATTTAGTA TCAGGTTTAA	4440
	CAACAGGTGC GACAAAAGGT TAGTTTGAAA TTAGGAGTGG GGCAGAAATG ATAAAGAACC	4500

55

	GGGTGTGGTG GTATTGCGAA TGGCAAGCAC ATGCCAAGTT TACAAAAAGT TGAAAATGTT	4620
	GAAATGATCG CATTTTGTGA CGTAGACATT TCGAAAGCAG CGAGTGC GGC AGAAGCATAC	4680
5	GGAACTGACA ATGCAAAGGT TTATGATGAT TACAAAGCAT TGTAAAAAGA TGACACGATT	4740
	GATGTTATCC ATGTTTGTAC GCCAAATGAC TCGCATTGTG AAATTACTGT AGCAGGGTTG	4800
	CATGCTGGTA AACATGTGAT GTGTGAAAAA CCAATGGCTA AAACGACAGC AGAAGCTCAA	4860
10	AAAATGATAG ATACAGCTAA ATCAACAGGT AAAAAATTAA CAATAGGTTA TCAAAATCGT	4920
	TTCCGAGCAG ATAGTCAATT TTTACATCAA GCAGCGCAAC GTGGCGACTT AGGAGACATT	4980
	TACTTCGGAA AGGCACATGC CATTTCGTCGT CGAGCAGTAC CAACATGGGG TGTCTTTCTA	5040
15	GACGAAGAAG CTCAAGGTGG AGGACCATT ATCGATATCG GTACACACGC TTTAGATTTA	5100
	ACGTTATGGA TGATGGATAA TTATGAACCA GAATCAGTGA TGGGTTCAAC ATTCCATAAA	5160
20	TTAAATAAAC AGCATCATGC GGCAAACGCT TGGGGTTCAT GGAATCCAGA TGAATTTACA	5220
	GTTGAAGATT CTGCGTTTGG ATTTATTAAA ATGAAGAATG GAGCGACGAT CATTTTAGAA	5280
	TCCGCTTGGG CGATTAATTC TTTAGAAGTG GATGAGGCAA AATGTTTATT ATCAGGAACT	5340
25	AAAGCAGGTG CTGATATGAA AGATGGTCTA CGTATTCATG GTGAAGACAT GGGTACACTT	5400
	TATACCAAAC ACGTTGAATT GGAAAACAAA GCGCTCGACT TTTATGAAGG TAATGAAGTG	5460
	GATGAAGCTG AAGAAGAAGC AAAAGCTTGG ATTGATGCAG TTGTAAATGA TACTGAACCA	5520
30	GTTGTGAAAC CGGAACAAGC AATGGTAGTT AAAAAATTC TTGAAGCGAT TTATCAGTCT	5580
	GCAAAATCAG GCAAAGCAAT TTACTTTGAA TAACATCATA CGGTAAGGAG GCACATCATG	5640
	ACAAAATTAA AAGTTGGTGT GATAGGTGTT GGTGGTATTG CACAAGACCG TCATATTCCA	5700
35	GCATTGCTGA AACTCAAAGA CACAGTCTCA TTAGTTGCAG TACAAGATAT TAATACAGTG	5760
	CAGATGATTG ATGTTGCGAA GCGCTTTAAT ATACCTCATG CAGTTGAGAC ACCTAGCGAG	5820
40	CTGTTTAAAC TTGTTGATGC GGTGGTCATT TGTACACCTA ATAAATTCCA TGCTGATCTT	5880
	TCTATAGAAG CATTGAACCA TGGTGTCCAT GTATTGTGTG AAAAGCCAAT GGCGATGACG	5940
	ACGGAAGAGT GTGATCGCAT GATTGAAGCG GCTAATAAAA ATCACAATT ATTAAGTGTG	6000
45	GCAATCATCATT ATCGTCACAC AGATGTGGCA ATTACTGCTA AAAAAGCAAT TGAATCAGGT	6060
	GTGGTTGGTA AACCTTTAGT AGCACGTGTA CAAGCGATGC GTAGGCGTAA AGTGCCTGGC	6120
	TGGGGTGT TTACCAATAA AGCGTTGCAA GGTGGCGGTA GTTTAATCGA TTATGGTTGC	6180
50	CAC TTGTTAG ACTTATCTTT GTGGCTACTA GGTAAAGATA TGGTGCCGCA TGAAGTGCTA	6240
	GGAAAAACAT ATAATCAATT GAGCAAACAA CCGAATCAAA TTAATGATTG GGAACATTT	6300

55

	GCAAGCATGC AGTTTGAATG TTCGTGGTCT GCAAATATCA AAGAAGATAA GGTTCACGTT	6420
	AGTTTATCAG GAGAAGATGG CGGTATCAAT TTATTTCCAT TTGAAATATA TGAGCCCCGC	6480
5	TTTGGAAC TA TTTTGAAAG CAAAGCTAAT GTTGAGCATA ACGAAGACAT TGCTGGTGAG	6540
	AGACAGGCGC GTAAC TTTGT CAATGCGTGT TTAGGGATAG AAGAGATTGT GGTGAAACCG	6600
	GAAGAAGCAC GCAATGTAAA TGCCCTTATA GAAGCGATTT ATCGTAGCGA TCTTGATAAC	6660
10	AAGAGCATAC AACTTTAATG ATTATCATAT ATGATACAAA ATTCTCAATA TAAAAAGAAG	6720
	GAGTGCTTTT CAATGAAAAT AGGTGTATTT TCAGTATTAT TTTACGATAA AAATTTTGAA	6780
	GATATGTTAG ATTATGTCTC AGAATCTGGA TTGGATATGA TTGAAGTTGG AACAGGTGGT	6840
15	AACCCAGGAG ATAAATTTTG TAAGTTAGAT GAGTTGTTAG AAAATGAAGA CAAGCGCCAA	6900
	GCATTTATGA AGTCAATCAC AGACAGAGGC TTACAAATAA GTGGTTTCAG TTGTCATAAC	6960
20	AATCCAATTT CTCCAGATCC GATAGAAGCG AAAGAAGCCG ATGAAACGTT ACGTAAAACA	7020
	ATCCGTTTAG CAAATCTATT AGACGTGCCA GTTGTTAATA CATTTTCTGG CATTGCAGGA	7080
	TCAGATGATA CCGCTAAAAA GCCTAATTGG CCTGTTACAC CTTGGCCAAC AGCCTACTCT	7140
25	GAAATTTATG ATTATCAGTG GAATGAAAAG TTGATACCAT ATTGGCAAGA TTTAGCTGAG	7200
	TTTGCAAAAAG AGCAAGATGT AAAAATTGCC ATAGAGTTGC ATGCAGGATT TTTAGTGCAT	7260
	ACACCATATA CAATGTTGAA GTTACGTGAG GCTACAAATG AATATATCGG TGCTAACTTA	7320
30	GATCCTAGTC ATCTATGGTG GCAAGGTATT GACCCAATTG CTGCGATTCT CATATTAGGC	7380
	CAAGCAAATG CAATTCATCA CTTCCATGCT AAAGATACGT ATATTAATCA AGAAAATGTA	7440
	AATATGTATG GTCTAACTGA TATGCAACCA TATGGTAACG TTGCGACAAG AGCATGGACA	7500
35	TTCCGTACAG TTGGTTATGG ACATAGTCCA TATGTATGGG CAGATATCAT AAGTCAACTT	7560
	ATTATTAATG GATATGATTA TGTATTAAGT ATTGAACATG AAGATCCTAT TATGTCAGTA	7620
40	GAAGAAGGTT TCCAAAAAGC TTGTCAAAC TTTGAAATCTG TTAATATTTA CGACAAGCCA	7680
	GCAGACATGT GGTGGGCATA ATACGAACTC GAGGTTAGTC TGAAGTTTGT CTGAAGTAAG	7740
	ACTGGTGGCA GTGTTGAATA AATGCATATG TCGCCAAGCC ATTGCCAAAA ATTTACACCC	7800
45	TTAAATCAAG TCATTGTTTG TAAAGAAGGT GTACTTTATA TAAGTATATA GCGATGGTCA	7860
	TACCCATTCA CAGTAACAAT CCTCACCATT GAAAAGAGTA TATAACCTTT TCAATAGTGA	7920
	GGTATATGAT AATAAAAAA GCCTGTTGTC ACAATGGTCA TAGACACGAC ATACTTTAAA	7980
50	GGTTTCTGAA TATAATATTT CAGAATGCAC TTTAAAGATG GACGTCGATG TAGACTAAAG	8040
	TGATGACAGG CTTTCATCTT TTTAAATATT CATTAATTTT TCTTCTTGTT TAATACGTAC	8100

55

	TAATACACCG ATTAATTCAG GAATGATGTT TAAGAAGTAA TTTGGGTGTT TTGTAATTTT	8220
	ATATAATCCA GATTTAATAA TAGGATGGTT AGGTAAAATG AATAATTTTA ATGTCCAAAT	8280
5	ACCACCTAAA GTTTTAATAA CCATAAATAA CATGATATAA GCAAAGATTA ATATAACTAA	8340
	GCCAATACCA TTTGCAAAGC TAAATGTATC TTTATTAATA AATGCCTCTA CACCAGCCAA	8400
	TACATAAATT AAAACGTGTG TTATTGCTAA AAACCTCGAA TTTTAAACGC CATATTCAAC	8460
10	TGCACCGTCT GCTTTTAATT GTTTTGAGTG ATTAATAGAT ATCTTTAAGC TGACAAGTCT	8520
	GATACAGAAA AAGATAAGTA ATATAGATAG AATCATGATG TCCTCCGTCA TTATGTCATA	8580
	TGTATAAGCG TTGATTTTGA CAACATAAAG TATTTTATAG ATAAAGCTTG TCAAATACTA	8640
15	TTAACTATTT ATTAATTTTA GTACATAAAT ATGTTTCTAA GTATGTGTTT ATGTTTCAGTA	8700
	TTTTGGATAA TTTAATAATT TTAAGGATAT TAAGCGCTTA CACCGACGTG ATATATTTGG	8760
20	CTTAACGAAA ATGATTGAGG TGACAGAGAT GAACTTTTTT GATATCCATA AGATTCCGAA	8820
	CAAAGGCATT CCATTATCGG TACAACGTAA ATTATGGCTT AGAAACTTCA TGCAAGCTTT	8880
	CTTCGTAGTG TTCTTTGTTT ATATGGCTAT GTATTTAATT CGAAACAACT TTAAGGCGGC	8940
25	ACAACCGTTT TTAAAGAGG AAATTGGATT ATCTACATTA GAACTTGGTT ATATCGGATT	9000
	AGCATTTAGT ATCACGTACG GTTTAGGAAA AACATTACTT GGATATTTTG TCGATGGACG	9060
	TAACACAAAA CGTATTATCT CGTTCCTACT TATCTTATCT GCGATTACAG TTTTAATTAT	9120
30	GGGATTTGTT TTAAGTTACT TTGGTTCTGT AATGGGATTA TTAATTGTAC TTTGGGGACT	9180
	TAACGGGGTG TTCCAATCAG TTGGTGGACC TGCAAGTTAT TCAACGATTT CAAGATGGGC	9240
	GCCAAGAACG AAACGTGGCC GATACTTAGG ATTCTGGAAT ACATCACATA ATATCGGTGG	9300
35	TGCCATAGCA GGTGGTGTG CACTTTGGGG TGCTAATGTA TTCTTCCATG GAAATGTTAT	9360
	AGGGATGTTT ATTTTCCCAT CGGTGATTGC ATTACTTATT GGTATCGCAA CATTATTTAT	9420
40	CGGAAAAGAT GATCCGGAAG AATTAGGATG GAATCGTGCT GAAGAAATTT GGGAAGAGCC	9480
	GGTCGATAAA GAAAATATTG ATTCTCAAGG TATGACGAAA TGGGAGATCT TTAAAAAATA	9540
	TATECTGGGA AATCCTGTTA TATGGATTCT ATGTGTTTCA AACGTCTTTG TATACATTGT	9600
45	ACGAATCGGT ATTGATAACT GGGCACCGTT ATATGTGTCA GAGCATTTAC ACTTTAGTAA	9660
	AGGCGATGCA GTTAATACGA TATTCTACTT TGAAATTGGT GCATTAGTTG CAAGTTTATT	9720
	ATGGGGCTAC GTATCAGACT TATTAAGAGG TCGTCGTGCA ATTGTAGCTA TTGGCTGTAT	9780
50	GTTTATGATT ACATTTGTTG TCTTATTCTA CACAAATGCT ACAAGTGTCA TGATGGTTAA	9840
	CATTTCAATTG TTTGCATTAG GTGCGTTAAT CTTTGGTCCG CAATTATTAA TTGGTGTATC	9900

55

	CGCGTATCTA TTCGGTGA CT CAATGGCGAA AGTTGGTTTG GCGGCTATTG CTGATCCAAC	10020
	ACGTAACGGT TTAAACATCT TTGGATATAC ATTAAGTGA TGGACAGATG TTTTCATCGT	10080
5	CTTCTATGTT GCATTATTCC TAGGCATGAT TCTATTAGGA ATCGTTGCTT TCTATGAAGA	10140
	AAAGAAAATT AGAAGTTTAA AAATTTAATA TAAATCGGAT TAAAAGTATC GCCAATCTAT	10200
	TGCAATATAG TTGGCAATCC TGCCCCGACG GCATGTGCGT GAAGAGATGA AAGATACTGC	10260
10	TTCTACCCCTT GCAAATATAT CATCTCTATG TCTCGGGGCA GATCATAATT CCCTGTTATG	10320
	AAGTATCCTT ATTTGCCCGA CTTAGGGTGA CTCAATGAAT TTAATCCTTA CAATAAAGAC	10380
15	ATATAGCGGT GTCAATATTG TAGGGAGTAT TGTTTTATAT TTAAACTCTC TAAAAAGCGG	10440
	ACTGAAAGAA AAGTGA AAC TTCTCTATCA GTCCGCTTTT TCATAGAACA AAATGGAGGC	10500
	GCCATAATCA TTAGTTATGT GCTAATCTAT TTTGCTTGCT TACAATAATC ACTTGGCGAC	10560
20	ATTTGTAAAT ATTTTTTAAA ATGATAGCTA AACATTTTAT ACTCTGAAAA GCCTACTTTG	10620
	TCTGCAATTT CATAGTGTTT GTAATGTCGA TCTAACAAAT GCAGAGATTG TAAAATACGA	10680
	TAGCGATTTA AATAATCGAC AATTGTAATA CCAACATGAT CTTTAAATGT TCGCATCGCA	10740
25	TACGATTAC TAACATCGAT ATGTTGAATT AAATCTGAAA CAGTCACTTT CGTTTGATAA	10800
	GATTGCTTAA TTTGATCCAC AATCTGGTTT ACATAATAAT CATCGTATTC TACTTTTAAT	10860
	AGTGGTTGGA AGGCATCATG ACAAGATGCT AAGCTACGGC CGTTCTGTGA TTGTTGCTCT	10920
30	AATAAGGTAC GGACAAGTCT TCCTAAAATA ACTTCTAATT GTGCATGGTC TACTGGTTTT	10980
	AATAATAAT CAAGAACATG ATGTTGAATG CCGGCTTTCA TATATTCAAA GTCATCGTAA	11040
35	CTCGATAATA TGATGACATT ACAATCTAGA TCGGCAATAT CATTGAGTAA ATCGACGCCA	11100
	TTTTTACGTG GCATACGAAT ATCAGTAATT ACTAATTCTG GCTGATGTTG TTGAATTAGT	11160
	GATATGCTT CAACACCATC TTTAGCAGTG TATATTGTAT TGAAATGATA GTCTCCCCAA	11220
40	GGAATGATT GCTTTAATCC TTCTCGAATA ATTCGTTTAT CATCACAAAT AACTACCTTA	11280
	AACATCTACA TTCCCCCTTG AAAGTGGTAT TTTATAACAA ATTAACGTAC CTTGATTACG	11340
	CTTTGAAAA ATATGGAGTC GTGCATGTGA ACCATATTGA ATCATGCTT TATTGTGTAA	11400
45	ATGATTTAAT CCCAATGCT TAGTATCAAA TACATCATT TTAAGAGATT GCGGTACATA	11460
	TTGCAGGCGA GATGACGACA TCCCGATACC ATTGTCGCAA ACTAAAACAT GTAAATTCTG	11520
	ACGTGCCAAT GTCAGGCGTA TAGTAATGTC CAATGACTCA GTATCTCTAC CATGTTTAAT	11580
50	AGCATTTTCT ATGAGTGGCT GAAGCATCAT TTTACCAATT GTCTGGTGAC GCGCTTCTTC	11640
	AGAACTTTCA ATATGGAGCT TAATCATGTC ATCAAAACGG ATGTTTTGTA TTGCAACATA	11700
55		

EP 0 786 519 A2

	GTAACGTAAC	ATTTGCGATA	ATTGTTGGAC	CACAGTTtGT	GCTAATTTTCG	GAGATAACGT	11820
	AATTAAATAT	TGTATTGTTT	GCATCGTATT	GAATAGGAAA	TGAGGCTGGA	ATTGGCGTTC	11880
5	TATTTCCCTT	AACTGAATAT	CACGCAAGCG	ACGTTCTGTA	TGCTCGATAG	AATGGATCAG	11940
	TTGCTCATTT	GATTCAAATA	AATCGTAAAT	ATAATTATTA	ATTTCTTCTA	GTTCACTGTT	12000
	GTTTTTTAAA	GGCGTATATG	TACCTAGATG	ACGATTTTTG	GCATAGTAAA	TTTTTTGAAT	12060
10	AATCGTTTTG	ATATCTTTTG	TTTGTCGTTT	AGCCATATTA	TCTGCGCTAA	TGAAACCAAA	12120
	TATTACTAGT	AAAACAAGAA	CTACGGCCAT	AACAATTAAC	AACGTGATAC	CATCTTCAAT	12180
	GTTTTCATGT	ATATCTTTAT	AAATAATGAG	ACGATGGTCA	GCATGGTTTA	ATTTTACAGA	12240
15	TTCATTCATA	AATCCGAATT	GTTGTGGTcT	ATACTTTTCA	CCTATAGTAA	AACGGTCATC	12300
	GTTGGCGTAT	AAAATATTGT	CATATTGATC	AmCGATAAGT	GCGAATTGTC	GGTTATCTTT	12360
20	CtTAATTTCA	CTTAAACGTG	GGGTGtAGC	CATATAAAAt	TTaAGCATAT	ATGTACTATT	12420
	TTTGAATTTA	AGCTGATGCG	TTGAAAATAA	ATACATATTT	TTAGTGTTTA	AATGTTTATA	12480
	ATTATTGGTT	ATAAACTGAT	TTGGTCCAGA	TAATTCATAA	TAAAGTGTG	CGGGCTGTTG	12540
25	GKGTATTAAT	TTTAATAATT	CACGTTTTGT	AGCGGTCACA	TCATGATGAT	TTGyTAAATC	12600
	GAGCTCTTGA	AACGAATTAT	TATGCTGTGT	AATAAATGTC	TGAATCTGCT	TTTCAGTATG	12660
	ATGTAAAGAT	GACTGACTTT	CATCAACATG	TTGATGAATC	GTACGATGCT	CAATCCAAAT	12720
30	ATAGATGGCA	TAGAAGCTTA	CTAGTCCAAT	AATAATGACT	AAAAATACTG	GAAAAATAGT	12780
	AGACnCAAAT	AACGATCGTC	TTAATTGATG	TCTATAAGGT	TTGTATGCCn	TCATTGAATC	12840
	ATCTCCAAAA	ATTTATGATG	TGGAATATCC	GGTAATTTAG	ATTTCCGGTAT	TAAAGGTATG	12900
35	TTCTTAAGAT	TTTCGATAGA	CTGATCGCTT	TGTTCACTAA	CATCCTTTTCG	AATTGACTTG	12960
	GCATcGAAC	CTGCAACTAA	TCGTtGTTGT	ACTGAGCGGC	TTGTTAAATA	TTGCACTAAC	13020
40	TTTTTACGCT	TAGGATGAGG	GTGTGCATTT	TTAACTAAAG	CAATtCCATC	AACATTTAAC	13080
	ATTGTTCCCT	CAATTGGATA	AACGATTGAT	ACAGGATAAC	CTTTGTTTTT	CCATGTGCGT	13140
	GCATCTTGTT	CGTAGCTTAG	ACCTGCGTAA	TATTTACCTT	TTGCAACATC	TTCAATGACT	13200
45	TTAGACGTCT	TTGACAGTTG	CATCGCATGG	TTTTGGAATT	GATGCACATC	ACTTACTCGA	13260
	TGATGCATGC	TATAAATAGC	ACGCATATGT	TGATAGCCTG	TCGTTGTTGT	ATTTGGATTT	13320
	GAGTACGCAA	TTTTACCTTT	AAGTATAGGT	TGTAATAAAT	CTTGATAACC	TCGAATCTTA	13380
50	ATATCTCCTT	GTAAATCTGA	ATTCACTACT	ATAACTGTTG	GCATTAATAG	AAAAC TAGTA	13440
	ACATATTTAT	TGTTGAGCG	ATAATCCTCT	AATTGCTGTG	TTACAGATGT	ATCTTGATAG	13500

55

	CCACGCTCCG AAAAATCTTC GTTATGCAAG TTTGAAAGCA GTACTTGAGT AGATCCGTGT	13620
	TTAATTTCAA TTTTGACATG CTCTTGTTTT TCAAATTCAT TTAAAAATTGG ACGAATCAAG	13680
5	TTTGATTGAT ACGGAGAATA AACTGTTAAT ACATTTTTAT CGGATTCAGA GTGACGCGTA	13740
	TTAGCGCATG CTGaTAAAAA AATGAGAAAT AATAGCAAGA TATAAATTTT TGATTTTCATG	13800
10	ATATCCCATC AATTCTATGT ATATTTTAAT ACAATAATTT TAGCAATAAA TGACGCATAA	13860
	GTAATGTTAA ATATTTAGAA ATGTTTATAG ATGACTTGTT AAGACGTGTC AAATGTTGTG	13920
	ATAGCACAAA ATTTTTGTTT GTCAAGACGA TTTACCGAGG CTGTAAAATC AACTGTTAT	13980
15	ATTTTATTG TAGCTGTTAT ATAAAAATCG GCAAGATATT GAACGGTTCA AAAGTGAATT	14040
	TTTACGTCAA TAAAAGTATT TAATCCAGTC TCTTCATATA TAAAAGTAAA TCTTTCTAAG	14100
	TGTTGATTTA ACGCTTATCA ACAATCATT TTTATAAACA AATATATACT CCTAAATTAA	14160
20	CTTTTAAAGC AATGAAAATA GTGAACATTA TAACTGTTGT GTAACAGAAT GCAATTAGCA	14220
	TATTACTGTT ACACAAATTA GTACAGTTTC TATGTTTGA CATACATTG ATGAAAATTG	14280
	TACATAATTT ATGTGAAAAA AATCACAACA AACATGCTAC AATGACTATG AAAACGTTAA	14340
25	CATAGCATT CAAATTCACA ACATTATACA GATGGAGGCG TTTAGTATGT TAGAAACAAA	14400
	TaAAAATCAT GCAACAGCTT GGCAAGGATT TAAAAATGGA AGATGGAACA GACACGTAGA	14460
	TGTAAGAGAG TTTATCCAAT TAACTACAC TCTTTATGAA GGTAAATGATT CATTTTTAGC	14520
30	AGGACCAACA GAAGCAACTT CTAACTTTG GGAACAAGTA ATGCAGTTAT CGAAAGAAGA	14580
	ACGTGAACGT GCGGCATGT GGGATATGGA CACGAAAGTA GCTTCAACAA TCACATCTCA	14640
35	TGATGCTGGT TATTTAGACA AAGATTTAGA AACAATTGTA GGTGTACAAA CTGAAAAGCC	14700
	ATTCAAACGT TCAATGCAAC CATTCGGTGG TATTCGTATG GCGAAAgCAG CTTGTGAAGC	14760
	TTAÇGGTTAC GAATTAGACG AAGAACTGA AAAAATCTTT ACAGATTATC GTAAACACA	14820
40	TAACCAAGGT GTATTCGATG CATATTCTAG AGAAATGTTG AACTGCCGTA AAGCAGGTGT	14880
	AATCACTGGT TTACCTGATG CATACGGACG TGGACGTATT ATCGGTGACT ATCGTCGTGT	14940
	AGCTTTATAT GGTGTAGATT TCTTAATGGA AGAAAAATG CACGACTTCA ACACGATGTC	15000
45	TACAGAAATG TCAGAAGATG TAATTCGTTT ACGTGaAGAA TTATCAGAAC AATATCGTGC	15060
	ATTAAAAGAA TTAAAAGAAC TTGGACAAAA ATATGGTTTC GATTTAAGCC GTCCAGCAGA	15120
	AAACTTCAAA GAAGCAGTTC AATGGTTATA CTTAGCATAC CTTGCTGCAA TTAAAGAACA	15180
50	AAACGGTGCA GCAATGAGTT TAGGTCGTAC ATCAACATTC TTAGATATCT ATGCTGAACG	15240
	TGACCTTAAA GCAGGCGTTA TTA CTGAAAG CGAAGTTCAA GAAATTATTG ACCACTTCAT	15300
55		

EP 0 786 519 A2

	AGACCCAACT TGGGTAAC TG AATCTATCGG TGGTGTAGGT ATTGACGGAC GTCCACTTGT	15420
	TACGAAAAAC TCATTCCGTT TCTTACACTC ATTAGATAAC TTAGGTCCAG CTCCAGAACC	15480
5	AAACTTAACA GTATTATGGT CAGTACGTTT ACCTGACAAC TTCAAAACAT ACTGTGCAAA	15540
	AATGAGTATT AAAACAAGTT CTATCCAATA TGAAATGAT GACATTATGC GTGAAAGCTA	15600
	TGGCGATGAC TATGGTATCG CATGTTGTGT ATCAGCGATG ACAATTGGTA AACAAATGCA	15660
10	ATTCTTCGGT GCACGTGCGA ACTTAGCTAA AACATTACTT TACGCTATCA ATGGTGGTAA	15720
	AGATGAAAAA TCTGGTGAC AAGTTGGTCC AAACCTCGAA GGTATTAACA GCGAAGTATT	15780
	AGAATATGAC GAAGTATTCA AGAAATTTGA TCAAATGATG GATTGGCTAG CAGGTGTTTA	15840
15	CATTAECTCA TTAAATGTTA TTCACTACAT GCACGATAAA TACAGCTATG AACGTATTGA	15900
	AATGGCATT A CATGATACAG AAATTGTACG TACAATGGCA ACAGGTATCG CTGGTTTATC	15960
20	AGTAGCAGCT GACTCATTAT CTGCAATTAA ATATGCACAA GTTAAACCAA TTCGTAACGA	16020
	AGAAGGTCTT GTAGTAGACT TTGAAATCGA AGGCGACTTC CCTAAATACG GTAACAATGA	16080
	CGACCGTGTA GATGATATTG CAGTTGATTT AGTAGAACGC TTCATGACTA AATTACGTAG	16140
25	TCATAAAACA TATCGTGATT CAGAACATAC AATGAGTGTA TTAACAATTA CTTCAAACGT	16200
	TGTATACGGT AAGAAAAC TG GTAACACACC AGACGGACGT AAAGCTGGCG AACCATTTCG	16260
	TCCAGGTGCA AACCCAATGC ATGGCCGTGA CCAAAAAGGT GCATTATCTT CATTAGTTC	16320
30	TGTAGCTAAG ATCCCTTACG ATTGCTGTAA AGATGGTATT TCAAATACAT TCAGTATCGT	16380
	ACCAAAATCA TTAGGTAAAG AACCAGAAGA TCAAACCGT AACTTAACTA GTATGTTAGA	16440
	TGGTTACGCA ATGCAATGTG GTCACCACTT AAATATTAAAC GTATTTAACC GTGAAACATT	16500
35	AATAGATGCA ATGGAACATC CAGAAGAATA TCCACAGTTA ACAATCCGTG TATCTGGTTA	16560
	CGCTGTTAAC TTCATTAAAT TAACACGTGA ACAACAATTA GATGTAATTT CTCGTACATT	16620
40	CCATGAAAGT ATGTAACAAA ATTTAAGGTG GGAGCACTAT GCTTAAGGGA CACTTACATT	16680
	CTGTCGAAAG TTTAGGTACT GTCGATGGAC CGGGATTAAAG ATATATATTA TTTACACAAG	16740
	GATGCTTACT TAGATGCTTG TATTGCCACA ATCCAGATAC TTGGAAAATT AGTGAGCCAT	16800
45	CAAGAGAAGT CACAGTTGAT GAAATGGTGA ATGAAATATT ACCATACAAA CCATACTTTG	16860
	ATGCATCGGG TGGCGGTGTA ACAGTCAGTG GTGGCGAACC ATTGTTACAA ATGCCATTCT	16920
	TAGAAAAATT ATTTGCAGAA TTAAAAGAAA ATGGTGTGCA CACTTGCTTA GACACATCGG	16980
50	CTGGATGTGC TAATGATACA AAAGCATTTC AAAGGCATTT TGAAGAATTA CAAAAACATA	17040
	CAGACTTGAT ATTATTAGAT ATAAAACATA TTGATAATGA CAAACATATT AGATTGACAG	17100

55

TATGGATTCTG ACATGTCCTT GTGCCTGGTT ATTCTGATGA TAAAGACGAT TTAATTAAAC 17220
 TAGGGGAATT TATTAACTCT CTTGATAACG TCGAAAAGTT TGAATTCTG CCATATCATC 17280
 5 AGTTAGGTGT TCATAAGTGG AAAACATTGG GCATTGCATA TGAATTAGAA GATGTCGAAG 17340
 CGCCCGATGA TGAAGCTGTT AAAGCAGCCT ACCGTTATGT TAACTTCAA GGGAAAATTC 17400
 CCGTTGAATT ATAAATACAA TTCAGACCGA AAAGAAAGCA TATGCAACTT CAAGAGTGAA 17460
 10 GGGGCATATG CTTCTTTTTC AATTGAGTAT TGAGTATTAG CAAGACGTAG TAAGTATATG 17520
 AGACAACTTC TACAATGGTT GAAGGAAGAC GTTTTGTAA GTAGCTATGC TGATAAAGAA 17580
 TGTGATGTCT TGTAAAGGT GGGGTCCAA TATCATCATT TAGCTGATGT TGAATGGGT 17640
 15 ATTATTGCT ACTTGCATAT GAATATGAGT CTTTCAAAT TTTTATTGAC CCTGAGTAAT 17700
 GAAAAATATT AAGATGAAAC TTAATATTAA AgCAATGCGG AGCGTGATTA TGAAGAGAAT 17760
 20 TAGTAAAGAT ATATGGGCAG TATTTAAATT ACTGTATCaA AATAAAGGGC GTTTTAGCAT 17820
 TAATGCCTTA CTATTGCAGT TAATCATGAT TTTTATTAGT AGTACATACT TAATTTTACT 17880
 ATTTAATATG ATGTTAAAAG TAGCTGGCAA AGCCAACTTA CGATTACAA TTGGACGGAA 17940
 25 ATCGTTAGTC ATCCCGCCAG TGTGATACTT CTTATTATAT TCATATTAAG TGTGCTTT 18000
 CTGATTTATG TAGAGTTTTC ATTGTTAGTT TATATGGTTT ATGCCGGCTT TGATCGACAG 18060
 ATTATTACAT TTAAATCCAT TTTTAAAAAT GCCTTTGTAA ATGTGCGTAA ACTCATAGGT 18120
 30 GTACCAGTTA TTTTCTTTGT CATTATTTA ATGTTAATGA TACCCATTGC CAACCTAGGA 18180
 CTAAGTTCAG TATTAACAAA AAATATTAC ATACCTAAAT TTTTAACGGA AGAAGTTATG 18240
 AAAACGACGA AAGGTATAAT CATTACGGT ACCTTTATGA TTGCTGTATT TATATTAAAT 18300
 35 TTTAAATTAA TATTTACTCT ACCGTTAAG ATTTTAAACC GCCAGTCGTT ATTTAAAAAT 18360
 ATGAGACTAA GTTGGCAAAT TACGAAGCGA AATAAGTTTC GGCTTGTTAT AGAAATAGTT 18420
 40 ATATTAGAAC TCATCATGG TCGATTTA ACATTAAAT TTTCAGGAGC AACATATCTT 18480
 GCTATTTGTG TAGATGAAGA AGGAGATAAG TTTTGTCT CATCAATTTT ATTTGTTGTA 18540
 TTGAAAAGCG CATTGTTCTT CTATTATkTA TTACGAAAT TATCATTAAT CAGTGTGTTA 18600
 45 GTACTGCACT TAA 18613

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1214 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

	AAAGTTTTAA AAGGGGTGAG ATACTTGGCG AATAATCCAT TCCAGCTTTG CGTTTAAAAG	60
5	GAATTATACT TGCCATTGTC GGTGCTTGTT TATGGGGATT AGGTGGTACT GTTTCTGATT	120
	TCTTGTTCAA ATATAAGAAT ATTAATGTCG ATTGGTACGT CACTGCTCGA CTTGTAGTCA	180
	GTGGTGTITT CTTACTTATT ATGTACAAA TGATGCAACC CAAACGTTCA ATATTTAGCG	240
10	TATTCCAAGA TCGACGTATG TTAGGCAAAT TACTTATCTT CAGTATACTG GGCATGTTAG	300
	TAGTACAATA TGCTTATATG GCATCTATTA ATACAGGTAA TGCTGCGATT GCAACATTAC	360
	TACAATACAT TGCGCCAGTT TATATTATTA TTTGGTTTGT CATAAGAGGC GTTGCAAAAC	420
15	TAACATTATT TGATGTGCTT GCTATTATCA TGACACTATT AGGAACATTT TTATTATTAA	480
	CAAAATGGTTC ATTTTCTAAT TTAGTCGTCA ATCCTGCAAG TTTATTCTGG GGTATTTTAG	540
20	CTGGTGTAGC ACTCGCTTTT TACACAATTT ATCCTTCAGA CCTACTTAAC CGCTTCGGTT	600
	CGATTCTAAT TGTCGGGTGG GCAATGCTTA TTTCTGGTGT TGCATGAAT TTACGCCATC	660
	CAATTTGGCA CATTGATATC ACTAAATGGG ACATATCAAT TATATTATTT TTAATCTTTG	720
25	GTATTATCGG TGGTACCGCA CTCGCATTTT ATTTCTTTAT CGACAGTTTA CAATACATAT	780
	CAGCGAAAGA AACACATTA TTCGGAAGT TGAACCTGT CGTAGCCGTT ATCGCAAGCA	840
	GTCTATGGTT ACATGTGGCA TTCAAACCAT TTCAAATCGT AGGCATCATT CTTATTATGA	900
30	TTTTAATTTT ATTACTATCA CTTAAAAGAC AACCTGAAAC ATTAGATGAA TAAGAAAAC	960
	CTGATAATCA CTTTAGCAAG TAACTATTAT TTAACAACGT AGTTACCTTA TAGGTGATAT	1020
	CAGAGTTTTT TATTTTAGTT AATAATATTT TTCATTGGT ATAAAAAGC GTCGTCGCTC	1080
35	TGGTAATCGG AAATACTGGA ATAAATATG GAATTGGGTA ATAATCCCAG GTANTAAAAG	1140
	TCCATGTTCC GATANTCTT CCGCANCTCC AACCAAATTT GCCGATAAGG TTCCAAAAGG	1200
40	CATCCTGGGG GTAC	1214

(2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9458 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

	ATTTTGGTTT CATTACGAT GGGGTnATAC AGCAAACACA nCTAAAATAA CTATCAATAG	60
--	--	----

	CTTAGACAAT AAAAAATATG CCACTACAAT CGCTAATATT ACGATTAAAA AAGAAGCGTT	180
	AACGATTACT TTCATCGTTG TTCTATCTCT GAACATCATA TTAAAGACAA CTAGACTAAT	240
5	TGATAATGAA ACAGCAAAAA AAGTAATAGC TAACACTAAT TTCATCATAA ATAGACAGAC	300
	TAAACCTATG ACTAATAATG TATTAGAAAT TACAGCTGAC GTTTTTAACA TTCTCGaATT	360
	AATATGCACT CACCCTTTTT ATTTAAATAA CTTACATAAT CATAATAATA CATGATGTTT	420
10	CATAGGCCTG TCGATGATTG ATTCACAATA GCACGTGATT TTTTGTGTTT TCAATATTAT	480
	TCATTTATTG CATCAAAAAC ACCCTTTTTA ATTTTTACAA AAATTAAAAA AAGTGCTCCT	540
	AACTGCTTG CATGTAGAAA CACTTTTTCA TTGTAATGTT ATTCTTCTCG AGACATACCT	600
15	TTTAGCATAT TAAGCATGTA TGTAAACTA CGGTTGATGT CGTCATCTTT CAATACGCCC	660
	AATAGACTTC TTATAGTTGT CTTAGCATTT GGAAGCGTT GATTGGCAAC GTGTAATCCT	720
20	TTATTAACTT TATTTAGGAA GTCGCTTAAA TCTGATACAT TGAGTTCACC TAATAAAAAT	780
	ACCATTGAAG CCATATTAGA TAATAGCCCT GTATAAATAT CTTTATTAAG TTCAACTGCA	840
	AAATTTATTTA TGATGACTTG ACGTCCTCGA ATTGCACCAT TTAAAGCATC TAATAGTTTT	900
25	GCATCATCTA ATGTTTTAAT AAGCTTGATT GCTTTTAATA TACTATCTTT ATTCGCTGCA	960
	ATTGCCTCTG TAACTTCATT TAACTTTCT AACTTAATTT GTTCTTCTGA TTTTCTAAG	1020
	CGTCTAATTT TAGAAGATAT TCTCTCAGCC ATTATTTATC CACCTGATTT CCCGGGAAAA	1080
30	CATAATCTGA ACGTTCCCAT TTTTCTGTA CTTGAACACT GTACTGCGGT TGACGTTTTT	1140
	TATTGACACG GAAATTATTA GGGTTCAACG GTGACTTACC ACGTTTCGTA ATTACCTCCA	1200
	AACGACAGCT AGTACGTTTA TAAGATGGTG TATCCGTGTA TTGATCAACA TCACTaTTAG	1260
35	TTAATAAGTT AATTGCACCT AGATCTCCAT TTTCCATCGC aTCaTTATTT AATGGAATAT	1320
	AGATTTCTTT ACCTTTAACA CGATCTGTCA CGTGAAGTTG TAATACCGCT TCTCCTGTyT	1380
40	CAGAAATCAG CTTAACTTCT GCACCTTCAT GAATGCCTCT ATCTTCAGCA AGCTCTGGAG	1440
	AAATTTCAAC AAATGCACGT GGCACTTTGT ATTTAATCAT TGGTGTGTTGA TAAGTCATAT	1500
	TACCTTCATG GAAGTGCTCT AACAATCGAC CATTGTTTAC ATGAATATCA TAAATTTCAT	1560
45	CTTGCTTAAA GTAATTATCA AATGATAATG GGAATAATTT TGCTTTACCA TTATCAAAAT	1620
	TGAATCCTTC TAAGTATAGA ATAGGCTCAT CAGTACCATC AGGTTGTACT GGCCATTGTA	1680
	AACTATTGAA TCCTTCTAAA CGATCATAAC TTACCCAGC ATATAGAGGT GTTAAGCGTG	1740
50	CTACTTCATC CATAATTTCA CTAGGATGCT TGTAATTCCA ATCAAATCCT AATCTATTAG	1800
	CAATTGCTTG GAAAATTTTC CAGTCAGGTT TTKAATCACC AAGAGGTTCT AATGCTTGGT	1860

55

	TTGCTGGCAA TACAACATCT GCGTATGTTG CTGTGAATGT TAAAAATTCA TCTTGGA	1980
	CCATGAAATC TAATTTTTC AAGCGAGCTT GTACAAAATT AATATTTGAA TCCACAATAC	2040
5	CCGTATCTTC ACCATATAAG TACAATGAGT GTACTTCTCC GTCATGTATA CCTTCTACCA	2100
	TTTCATGATT ATCTTTACCA GCTTTTGGAT TCAATTTAAC GCCATATTCT TTTTCAAATT	2160
10	TAGCGCGAAT ATCATCCGCT TCAATACTTT GATAACCAAGT AATCTTATCA GGCATACTTC	2220
	CCATATCACT ACATCCTTGA ACATTATTAT GTCCACGTAA TGGATACGCA CCAGTACCAG	2280
	GACGACGATA ATTACCTGTT ACTAATAATA AGTTTGAAAT CGCTGTACTT GAGTCACTAC	2340
15	CAATGTCTTG TTGTGTAATA CCCATTGCCC AACAAATTAC AACAGATTCA GCTTTAGCAC	2400
	ATTCTTCAGC AAATTTAATC AATTCTGATT CAGGAATACC TGTGCTTCT TCAGCAAAG	2460
	CCATTGTAAA TGTTTCTAAT GATTTGTAAT ATTCATCAAA ATCATCTACC CACTCATCAA	2520
20	TAAATGCTTT ATCGTGTAAT TCATGATCAA TAATATACTT AGTCACTGCA CTTAACCACG	2580
	CTAAATCCGT ACCTGGTTTA GGTGATAAA AACGATCCGC ACGTTCTGCC ATTTCATGTT	2640
	TTCTAATATC AAATACATGT ATTTTTTGAC CAAATAATTT TTGTGCACGT TTCATGCGTG	2700
25	ATGCGATAAC TGGATGAGCT TCGGCTGTAT TAGTACCTAT CAATACAGAC ATTGCCGCTT	2760
	TTTCTAAATC TTCAATACTA CCTGAGTCAC CGCCGTGTCC AACCGTTCTA AATAAGCCTT	2820
	TTGTTGCAGG TGCTTGCAA TATCTGAAC AGTTATCAAC GTTATTTGTG CCAATAACTT	2880
30	GTCTTGCTAA TTTTTCATT AAATACGATT CTTCAATCGT CGCTTTAGAA GAAGAAATGA	2940
	ATGATAGTGC ATCTGGGCCA TGCTTTCTT TAATAGCTGT AAAATTATCT GCAATGACGT	3000
35	TTAAAGCTTC ATCCCATCT ACTTCATGGA ACTCACCAT TTTCTTACT AGTGGTTTAG	3060
	TTAATCGTTG ATCTGAATTA ATATGTCCCC ATGAAAACTT ACCTTTAACA CAAGTCGCAA	3120
	TTTTATTGTC TGGAGAATCA TGTGATGGTT GTACTTTTAA AATTTCTCTA TCTTTAGTCC	3180
40	AAACTTCAAA TGAACAACCC ACACCACAAT AAGTACACAC TGTTTTAGTT TTCTTAATAC	3240
	GCTCTTTACG CATTTCTGCT TCTGAATCTG AGATTGCAA TAGTGGACCA TAACCAGGTT	3300
	CTGCTTTTTT AGTTAAATCA ATCATGCTG CTAATGAACC AGGTTCCGTA TCAGTCATAT	3360
45	AACCCGATT ACCTTCCATA TTCATTCCA TCATGGCATT ACATGGACAT ACCGTCCGAC	3420
	ATTGACCACA AGATACACAT GAAGACTCAT TAATCGGTAC ATCATTATCC CAAATAACAC	3480
	GTGGATGTTT ACGATCCCAA TCAATTCTAA TAGTTTCATT CACTTCGATA TCTTGACATG	3540
50	CTTCTACACA ACGCCACAT AAGATACATT GATTTGGATC ATAACGATAA AATGGGCCGT	3600
	AATCTTTTTT GTATGGCTTC TCTTTATATT CATACGTTTG ATGCTGAAGC CCCCATGCAT	3660

55

	TATGCTTTTC TAAAATTCTGA TCAAGCGCTT CTTTTTGAGC ATCTTTCACA TCATTGTTCA	3780
	CAGTATTTAC AGTCATTGGA CGATCAATCA CCGTACTACA TGAACGTTCA ATTTTACCGT	3840
5	CAATCTCAAC AGTACATGTA TCACATGTTT GAATTGGTCC CATCGACTCG TTATAACAAA	3900
	TTGAAGGTAC AAAAGTATCT TGTGATTTAA TAAATTCAAG TAAATTCGTA CCTGGTTCTA	3960
	CAAGATAATC TTTTCCATCA AGTGTAACCA CCAAATGTTT TTGCATATTA CTCACCCCGT	4020
10	CTATATATAT TTTCCGTAAA TGACTTTTAA TAAATTGCTC ATATCCACCT AAAATAACGA	4080
	TGCCCCACAC ATCTTTCAGA TAGAATTAAT TTAATTGTAT TACTTTATGT ACTAGTTGTT	4140
	AAGTAAATTT TGTATTTTGG CCTTTTTTACA ATCATTTTTTA TTTGAAATAT TTTGCGCGAA	4200
15	ATTAAATCAT CTTTTTGTTT AATTGAAAAT AATTATCATT ATTAGTTTTC CAATTATCTG	4260
	TTTCACGCTT TTTGCCATAT CTTTCACAAC CTTATTAATG ACAATATTTA ATAATCACCT	4320
20	CACCTAAAAA TCGTTATACT ATTTATAAAT ACCCTTTTTTC TGAAAATTAA TAACCCAAGT	4380
	TTGATAAATA TCTACTATCA TTTAGAAGGT AATATTTATC TTTAAATTAA ATTTGTAATG	4440
	GATTAATTTA TAAAAATCAA ATCAGGCATT AAATAAAATA GCCCATAAAT ACAAAGTGTT	4500
25	ATCACCTTCT ATTTACGGGC TATTAGTTCT ATTCGTTATT CTATTTACAG ATCATTCTAT	4560
	CTAATTAATT TGTGTACAAT TTTGATAACT TATTTTCCCT TAGTTTACTA CTCTAGATTA	4620
	TCTTTTAATA ACTTAGTACT TTCAGCTTTT GACTGCTCAC TAGGAATGAA GTAGTACAAT	4680
30	CCGTCACCTT GAATGCCGCC TTGACCACTC AATTGATGTT TATTAATCGT GTCAITAGCA	4740
	TCTTTATAAT TGCTTCTAAT CGTATTCAAA TCACCTAATG TTAAATCTGT TTTAACATTA	4800
	TTTTGAATTT CATTCAATTAG ACTATTAAAA TGTGTAATCG ATGATGGGCT TGCAATCTTA	4860
35	TTGGCCATCG CTTCAAGCAC AATTGCTGA CGTTGTTGTC GACCAAAGTC ACCACCAGCA	4920
	CCTTCTTCTT TACGACTTCT AATAAACTTC AATGCTTGAT CACCATTTAC ATGTGTCTGC	4980
40	TGTCCTTTTG TAAAACGAAC ACCATCAACA GTGAATGTAT CATTACTTAC TACATCAACA	5040
	CCGCCGATGC TATCTATCAT ATTATGCAAA CCATCCATAT CGATTGTGCG ATAATGATCA	5100
	ATTGGCACAT TCATTAATTT TTCAAGTGAT TTAACAGCCA TATTTGGTCC ACCATATGCA	5160
45	TAGGCATGTG CAATTTTTTTC AGTAGTACCA CGGCCAACAA TTTCCGCTCT TGTATCACGC	5220
	GGTATACTTA CTATTTTCAGT TTTCTTCGTT TTAGGGTTGA TAGATAAAAT CATAATACTA	5280
	tCACTACGCT CTCGCCACC CTTTTTCTTA CGATCAGCAT CTGAATCGAC ACCAAATAAA	5340
50	GCGATTGTGA ATGGATCACC ATCGTTTAAA CTCACTTTTT TATCTCTTAA TTCTGAATGA	5400
	TTGCGATCTA ACGGATTGTG TATCTTATTA CCAGTAATAA AAATTTTAGC AGCTACATAC	5460

55

	GGTAGGCTCA TTTTACTTTT AGACGAACGT TTCAATCCCA CCACTCCTTT ACTATTCCTT	5580
	ACATACTTTG TCTGTTTTCT CTATTTATTA TATAGTAAAA TAATTTTTTT ACTATACTTC	5640
5	TGTAGACGTA TAACTATTTT TTATCATTTT TTATCTCTAG AGAATATCTA TCTGTATTTT	5700
	TGATAACCAC CATTGTCATT TAAAATTTTA AGTACCGTTT CATGACATGC TTTATTACTT	5760
	ATAATAAAAG GTGCACCCTT TAAATGATCA ATTGCCTTAC CATCTAAAGT CGTCATTTTT	5820
10	AGATTCAATA GTTCTGCAAA TAAAACTGT GCAGCAATGT CCCAAGGTTT AGGATTTGTA	5880
	TTAATATGTG CCCCAAATTG ACCTTTTGCC ACTCGCATAG AATCTAATCC GCAAGCACCA	5940
	ACTAAACGAT AACTAAATGA GCGTCAAAT AAATCTTGCA CCGTATCTAG ATTCATCACT	6000
15	TGTGCATTAA ACGATATAAT AGCGTCTTCC AATTTTAACG ATGGTGGTTC TTCCATCTTA	6060
	ATTCCATTAC AAAAAGCACC TTCTCCTCGT ATTGCTTTAT AAAGCTTTTT ATGCGGATAA	6120
20	TCATATACGT ACGATAACAT TGGTTTACCT TCATAAAAT ACGCCAATAT AATACAATAA	6180
	TCTTCTTGCT GTTTTACTAA ATTGGCAGTT CCATCAATGG GATCCATAAT CCATAAATGA	6240
	TTAATTTTCAT TCGTAATCAT TTCATTACTT TTTTCTTCCG CTAATAGTTG GTGTTCCGGA	6300
25	AAATGTGTTG CTAAAAATTG TTGGAATTGT TGTGTAATCT GTTTATCTAC ATTTGTAAC	6360
	AAATCAAATC GATGACGCTT AGTTTCTGTA GTCATTTCCTA TAATTAATTG CGGAATAACA	6420
	TTGTCTATTT GTTTCAACCA CGAACATATT AACTTATCTA TTTGCTGTAA TGTMTTATCT	6480
30	GTCATTTTCGT CCACCACTTC TCATATCATT ATCATTTTAT TATTACCCTA TATTAAAAGA	6540
	ATCAACAATA CAACTGAAGA CTTCTTCATT TTATGCATAA AAAAATCGGC TAGTCACGTG	6600
	CTAGCCGACA AATAGAAAGG AAAGTAAGTA ATAAATATTG AAGATGTTGT GATGTAACCT	6660
35	GAACGATTAA AAGCTATCTG TTATATAGCT CTACCCCTTT GTTTAATCGC TCCCCCTGTT	6720
	ACAAATAATA TCATAGCACA ATCTTTTTTA AAATGTAAGC GTTTTCCACA AAATTTTTAC	6780
40	GATTTTTTTA AAAAGATATT GAAAATGTCC TCATTGTCAC TCTTATGTTA TACTTTGTGT	6840
	AATATATCAT CTTTTAGGAG GTGGCTGTCA TGAATAAAGC TGAAAGGCAA AATTTAATAA	6900
	TTACTGCAAT TCAACAAAAT AAAAAAATGA CCGCTTTAGA ATTAGCTAAA TATTGCAACG	6960
45	TATCCAAACG CACAATTTTA AGAGATATTG ATGATTTAGA AAATCAAGGT GTTAAAATTT	7020
	ATGCGCATTG TGGGAAAAAT GGTGGTTACC AAATACAACA AGCACAATCT AAAATTGCAT	7080
	TAACTTATC TGAAACACAA TTATCAGCCT TATTTTGTAGT GCTTAATGAA AGTCAGTCGT	7140
50	ACTCGACATT ACCATATAAA AGCGAAATCA ACGCAATTAT AAAACAATGT TTAAGTCTTC	7200
	CACAAACACG CTTAAGAAAA TTGCTTAAAC GCATGGACTT TTATATTAAA TTTGATGACA	7260

55

	ATGTGATGTT AGTAGATCAT AGGGTTGATG ATAATATTAA AGCTGAAAAC GTTATATTTA	7380
	TTGGCCTTTT GTGTAAACAT GGACATTGGC ATGCAGTCAT TTATGACATT GCTCAAGACA	7440
5	AAACTGCCGA ACTCGAAATT GAAAATATTA TAGATATTTT GTATTTCATTC GGTAAGACGA	7500
	TTCAAACCAG AGACATATCC ATTGATAACT ATCATCAATT TTTAAACCCC ATCGATTCCCT	7560
	AAAAACAGC AGTAAGATGA TTTTCAATTA GAAAATATCT TGCTGCTGTT CTCTATTTAT	7620
10	ACAATACTTC GTATTGAATG GATTTCGCTTT CCTAGGGTGC CGTCTCAGCC TTGGTCTTCG	7680
	ACTGGCACTG CTCCCTCAGG AGTCTCGCCA TTAATACTAC GTATTAAACAT GTAATTTTAC	7740
	TTTGAAATAC TTAACAAAAT AAAACACTTT GCCCAACTTA CACTACCAAT AGAACTGCT	7800
15	GTTAGAATTC CTCAAAATGA TATTTCCGCGA TATGTTAATG AAATTGTTAA AAAGATAGCT	7860
	GATAGCGAAT TCGATGAATT CAGACATCAT CGTGGCGCAA CATCCTATCA TCTAAAAATG	7920
20	ATGTTAAAAA TCACCTCATA TTCATATACT CAATCTGAAT TTTCTGGCCG TAGAATAGAA	7980
	AAATTACTTC ATAACAGTAT TCGAATGATG TGGTTAGCTC AAGATCAAAC ACCTTCTTAT	8040
	AAAATATTA ATCTTTTATG AGTGAATCCT AATACTGATG CGCTAATTGA ATCTTTATTT	8100
25	ATTCAGTTTC ATAATAAAT GCATATCAAA AAAGCTGATT TCTATCAAAT AATTAATAGA	8160
	AATCAGCTTT TTTCAATTGCC TAAAACTTA ATGTCCCGAC CTCTTTATCT ACGCATAAAT	8220
	ACTTATTACT GATATAACGA AAGAAACAAA ATTATTTGCT ATATGTAATG CAATTGTTGA	8280
30	ACCTAGGTTT CTTCCAGATT TTAAATAAGT GAAAACTAAT ATGATGGATA GTATGAGATA	8340
	TGGACCAAAC TCAAACGGCG ACTTTGCATC AGTCACATGA ATAAATGCAA ATAAGAACAC	8400
	CGAAACAATA CTCATAGCTA TAAATTTAAA CTTCTTACCT AATTCTCCAA TTAAATATG	8460
35	TCTAAATACG ATTTCTTCAA CTATTGGACC TACAATCACA ATTAATAAGA ATGCTACAGG	8520
	TAAAAATGCA GGCACCTCAA ACATTTTATT TAGCTCAAGT TCATTGGCTG TTCACTATA	8580
40	TTGCAAATGT TTAGGTAGAA ACTGTGTCAT ATATTCATAT GTATAAATTA AGATGAGAGC	8640
	AATAATATAC GTTATTGACA ATCTAAGCCA ATATTTTTTG ATATACGCAA AACCAGCTCG	8700
	AAGCCTTGAT GGCATCACTT TTAAATGAAA TAAATAAAAT GCGCCAATCC CAATCGTATA	8760
45	TGCTAAAGCT TGTGTGATAG TCGCTACAAA TATCAGATTA CTATCGATTT CATAATAACC	8820
	AAACAAAATT GGTCTATGT AAGCTGCAAT TGTGAGTGCA TAAATATAA CACCTATAAT	8880
	TGGAATTATA AGCAAATCTC TCCATGCTAT ATCTTTAAAC GTGTATTTCT TTTTTCATT	8940
50	TTCCaCTGTT ATATCCCTTC CTGTTTAAATA ATTGATTTTT GGAGGTACTT CTACATGATA	9000
	AACGAACTA AGTATATGAG ACAACAAAT ACTAATTTGA TTCAAATCAT TGATACGATT	9060

55

ATAGTTACTA ATGAATTGAA TAAGTTCAAA GGCTTTGAAA CATCATATAT AATAAACGAA 9180
 AATCAAGTTT CCTATTATGA AATTATAACA CTACTTAATA AACGTCCCCT CgACAAGTCG 9240
 5 ACTATGGTAA CAAAATTCAA TATCTTAATT TTTATCATAC AGAACTATCT AACGCATTAT 9300
 TTGCAATTAA ATTTGCCCAT TAACCTATTT TTCATAAAAT GTCATTTAAA CAAGTTATTT 9360
 ATTAAAATTC ACTTTATTAC ATAAATTATA CAATTArAAA GTTCTTCAA ATTGTAAAGA 9420
 10 TGCATTAATC GAGTTATAAT CATAATGATT AAGATGGT 9458

(2) INFORMATION FOR SEQ ID NO: 115:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 910 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

AnGCGTATCA TGTCACGCAT TTAACTACT TCTTTACCAC AAGATTATAC AGTCACATTA 60
 25 GTTGATCGTA TGCCATTTCA TGGATTGAAA CCAGAATTTT ATGCTTTAGC TGCGGGCAGC 120
 AAATCAGATA AAGATGTTCTG TATGAAATTC CCTAATCATC CACAAGTGAA TACAGTTTAT 180
 30 GGTGAAATTA ACGACATAGA TTTAGATGCT CAAATTGTCT CAGTCGGTAA TTCTAAAATT 240
 GATTATGATG AGCTAATCAT TGGTTTAGGA TGTGAAGATA AATATCATAA CGTTCCAGGA 300
 GCCGAAGAAT ATACACATAG TATTCAAACA CTCTCAAAGG CTCGGGATAC TTTCCATAGT 360
 35 ATTAGTGAAC TACCAGAAGG TGCTAAAGTC GGTATCGTTG GTGCTGGATT AAGCGGCATA 420
 GAACTTGCCA GCGAATTAAG AGAAAGTAGA TCAGACTTGG AAATATATCT TTATGACCGT 480
 GGGCĒGCGAA TTTTAAGAAA TTTTCCAGAA AAATTAAGTA AGTATGTTGC GAAATGGTTC 540
 40 GCCAAAAATA ATGTTACCGT TGTTCCAAAT TCAAATATTA ATAAAGTTGA ACCTGGTAAA 600
 ATATATAACT GTGATGAACC TAAAGATATT GATTTAGTTG TATGGACAGC AGGAATTCAA 660
 CCTGTTGAAG TTGTTGTA CTTGCCGATT GATATAAATA GTAATGGACG CGTGATAGTT 720
 45 AACCAGTATC ATCAAGTACC AACATATCGT AACGTCTATG TAGTTGGTGA TTGTGCTGAT 780
 TTACCACATG CGCCAAGTGC TCAGTTAGCC GAAGTTCAAG GTGATCAAAT TGCCGATGTG 840
 50 CTTAAAAAGC AATGGCTAAA TGAACCATTA CCTGACAAAA TGCCGGAACCT AAAGGTACAA 900
 GGTATCGTTG 910

(2) INFORMATION FOR SEQ ID NO: 116:

55

(A) LENGTH: 10182 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

10	TTTTTGATTG AAAGTGGTGA TTAAACAAGC ATTTTAAATA GCAATGATTT GAAAGTCACA	60
	CATGATCCTA CCACTGATTA TTATAATTTA TCTGGTAAGT TGTCGAACGA TAATCCAAAC	120
	GTAAACAAT TAAAACGTAG ATATAATATT CCTAAAAACG CATCAACAAA GGTGGAATTA	180
15	AAGGGAATGA GTGATTTAAA AGGCAATAAT CATCAAGATC AGAAACTTTA TTTTATTTT	240
	TCAAGTCCTG GAAAAGACCA AATCATTTTAT AAAGAAAGCC TTACTTATAA TAAAATAAGT	300
	GAACATTAAT ACTTATGCTG TAATTATAGA AACATCCAAA TCATCTATTA AATCCTATA	360
20	TTATAAAAC ACCTCACATA ACTCGTTCAA CTGTACCAA CCACATTACA TTAGATTTTA	420
	GGCTAACTAT TGTGATGTAC ATCAAAAACG AATTTGTGAG GCGTTGTATA TTTTACAAAG	480
25	GTGACTAGCG TTTCGTATAG CATTTCCAAC ATTACTACAC TCAAGCGTCA CGCTAAAGTT	540
	CGAAATCGAA TCCTTTCATT CAACAAAAGC TCATATCCAC TACAACTTC ATATCAAGCG	600
	TATAAACTAT CTTGTGATAC TATCTCGATC ATATCTATAG TATGCATTG TGTCCGTTT	660
30	CACTGAAGTA TATGTATCAT CAGTTAAGTA TAAACCGTCA TCCTTCAATG TTACTTGATA	720
	AGCATATTC CGTGCTAACC AGGCAATATC TATATAATTT TCTCCTGCGT TTCATAACT	780
	TCTTAAATCT TCAATATGTG CACTAACTTC AGGGGAAATG ATTCTAACAA CACTTTCATC	840
35	AACCCAATAT TTGTCATGCA TCCATCGCAC TTGATCTGCC AATAAAGGTA ACTGCACATC	900
	ATTGAAATAT AGACGAAAGC CGTCACTATC ATACATTTGC CGATATGGTA ATGGCTGTTT	960
	TCTAATCACT AACACCTCGC CACCCATTAC GGTGCCTTCT CTAGTATCAT CACTTCCACC	1020
40	CGAAGCTTCA TACGTTGTTG GGTCAACCTG TAGTCCATGT ACATCTCCAA TATAAGCATC	1080
	TGGTTTATGT TCCATTGCAT GTCCATGTGC AATCAATGCT AATATTGTAG ATTGTGAAAA	1140
45	TTGAGGCTCC CATTCAATGC GATTAGGATG GCTACTATAA ATTCTAGGTT CATCTATAGC	1200
	CTGCTGAATA TCCATGCCAA AACTAATAC ATTGATTAAT GTTTGCGCAA CACTAGCAAT	1260
	GATACTTATG GCACCAGGTG CACCTACTGT TAATATTGGC TTCCCGTGAT ACATCACAAT	1320
50	CGTTGGAGCC ATGTTACTTA GTGGTCGTTT ATATGGTGCA ATTTGTTAA TACCACCATC	1380
	TACTACATCA AAGCCATCCA TTGTCGTATT CAATAACACA CCGTAGCCTG GAATCGTGAT	1440
55	ACCTGAACCA TAAATCATAC CAATTGATGT CGTAAATGAA GCAATATTAC CTTCTTATC	1500

	ATCAGACACA ACACCATGCT CTATATCAAT ATTTGCTTTA TTGCTATCAA TGAGCGTACT	1620
	GCGTGCTTTT AAATAATCAT CATCAATTAA TGACTGTACA GGCACCTCAT GAAAATTATC	1680
5	ATCCGCCAAG TATTGCGCAC GATCACTATA TGCTAAATGC ATCGCTTGTA TCAAATGATG	1740
	CAAGTAATCA ACAGATCTTG GACCCATAGA TGGTAAATCG ACATGTTCTA ATAACCTCAA	1800
10	TATTTGAATT ACCGTGATAC CGCCAGAACT AGATGGTCCC ATTGaATAAA TGTCATAGTC	1860
	TTTAAATGTT GCACTGATTG GCGCTTTAAT CTGAATGTCA TATTTGGCTA GATCCTCTAA	1920
	AGTGATTGTC CCACCACATG CTTTGACAAC ATTGACTAAT TGTTCGCAA TGTCACCTTT	1980
15	ATAAAATGCA TTAAACCCCTT GTTCTCTTAA TATTTGAAAT GTCTTACCTA ATTGCGGGTTG	2040
	TACAATCCAA TCACCTTCAC GCCAATATTG ATTTTCATGC GTAAATACTT GTGCCGTTTC	2100
	ATGATACTTT GTCAATCGTG CGTGTGCTG GCGCGAATAT TTTTCAGTAG CCCAATTGGC	2160
20	TGCATGACCT TCAATGGCTA GTTCAATTGC AGGATTAATT AAATCTTCCA ATGACAATTT	2220
	AGCATAACGC TTGTGAATAT AATCAAACAG CTTTGGAATT GCTGGCACAG CGACAGTTTT	2280
	ACCATGTGTA GTCATATCAA AAAATGATTT ATATTCGCCT GAATCATCTA GATAAAATTG	2340
25	TTTGTCTACA TGTTCAGGTG CTGTCTCAGG TGCATCAAAC GCAGTTATAC TGCCAGTACT	2400
	TTGCTCATAA TATAGCAAAT ACCCGCCACC ACCAATACCT GATGCAAATG GTTCTACCAC	2460
30	ATTCAATGCC AGTTGAATTG CAATCACTGC ATCCATGGCG TTGCCACCTT GATCTAATAC	2520
	ATCCTTACCA ATTTTAGCCG CAAGAGGATG TGATACGGAA ATTAACCCCTT CTTTAGATGT	2580
	TTTTGTCTGT TTGTCATTTA AGTTAATGAC CATACTATAT CCTCCTACTT TCTGTTAAAT	2640
35	ATTTAAACA TTATTGATTA ATGGCTTTTT CTACTTTTTT TAAATCTTGA CGTTGCTCGT	2700
	TACCAGTATC GACAAGTGGT GTAATCGGTG ATGCAATTTT AAATTTATCG CCACGATAAA	2760
	ACTTAATAAA TTGATCCTGA TCTATCGCAT TAACTACTGC TTGTCTCAAG TTTGGATGCG	2820
40	TCTTAAATAT ACCTTTTTTTA ATATTTAGCA TTAAAAAGAC TGACTTGCGT CCATTTTTTG	2880
	GAATAATGCT TAAATTTTTTA TCCGACTTAA TTAAATCAAA ATGTTTTTGA TTCACATCTG	2940
	CCAACATATC AATTGAATGA TTTCTAAGTT CTGACAATGC ATTATTCGGG TCACCATTAA	3000
45	ACTTCAATGT AATATTTTTTA ATTTTAGCTG GTCCATAACT ACCTTTTTTCT GTTTCGTTGA	3060
	ATCCTGGATT ACGTTGAAAC GTTGCTTGAT ATGCATTTTT CTGTGTCATA ATGTATGCGC	3120
50	CACTTGCATA CAGCGCATTT TTCCCATCTG AATTTGCAGG AATTGTACTG CTATCCCAT	3180
	ATCCTTTTGG ATATTCTTGA TTTACTTGAT TAACAAATTT TTTAGATAAA ATGCCTGCCG	3240
	AAGAGTGTGT TAAGTAATTT ACCTCTCGAG GCATCGATTG ATCTGTCGTA ATTTTAACAA	3300

55

	TATAAGCTTT AATCAACTTA TCATAGATTG ATTTATCGTC CTTGTCTTTC TCTTTACGCA	3420
	ACTGATCGAT GTCCTCATCT TTTAATATCT TGATGTCATT TATATGTTTG TGCATATTGT	3480
5	AAGTATTATT GTTAGGCACA GACTTTTTTAT CACGTGCTCT ATCTAAAGAA AACTTAACAT	3540
	CTTCAGCCGA TACACGCTCT CCAGTATTAC GTGCTTGTCCT ATTGACCACT TTCGCAAAAT	3600
	AATCATCATC TCTTAACAAG AAATAAAATG CTTTATTGTC CTTATTCCACA GCATAATCAT	3660
10	GACTTAACGA ACCTTTCGTT GTTAAATGAT CATTTTCATC TAATAATAAT AACCTTGTGT	3720
	ACATATTTCAT ATTAATTGAA TATACTGACG GCGCAATTGA ACGTATTGGA TCCAATGTAG	3780
	GAATTTCCACC ATCTTGTTGT GTCATCACAA GTGGCCGCGT ATCTCGTTCT CTACTATTGT	3840
15	TGTAATCAAA TTGTTGCCAT ATTAATGCAC GTGAATTTGG CAATCCAACA CTATTTTTTAT	3900
	CTAACACTTT ATTGTCATAT ACTAAATTCT TTTTGTATCC ATATAAAGGC GCCATATACC	3960
20	CTTTATCAAA TACAACTTCA TCTTCAATTT GCTTATATGT TTGTTTAAACA TCTGCTTCAT	4020
	TTTGAGTAGA AGCTTTATTT AACAACTGGT CTACATGTTT ATCTTTCAAT AACTATTGTG	4080
	ATCCTGTAGA ACTAAATAAT GCCGTCATAG CATAGTTCGG GTCACCAAAC ACTGTCATCC	4140
25	AGTCATCAAT TTGGATATCA TAATTGCCGG CTTGACGTTG TGTACGATAG CTACCATAAT	4200
	CTGGTTGGAT ATTCATCTTC ACGTTAAATC CTGCATTTTC CAATTGATCT TTAACGATAT	4260
	TCATATCATT TTCATAACTT GCTTGTCTTA GGAAATGTAT TGTGCTCGC TCGCCTTTCA	4320
30	CTTCAACTTT CGATGACTTT TGAGCCACTT CTGATTTTCGT AGGGACACCA CAACCACTTA	4380
	ATACCAACGC TAAAACTATA ATTGCGATAC TAATGATTTT CTTACATCT ATCCCTACCT	4440
	TTTTAATGAA TTCTTGATC TAGTGCATCA CGCACTGCAT CACCTATAAA ATTAAATGCT	4500
35	AAAACGACGA ACATAATACA AACACCAGGT ACAATAGCTA AATTACTGTG CGTTTCCAAG	4560
	TAGTACTAC CGGTACGTAA AATGTTGCCC CATTGAGCTA CATCAGGTGC AACACCAAGT	4620
40	CCTAGGAAAC TTAAACTACT TGTGTTAAT ACAACCACAC CTATATTAA TGAAAAACGT	4680
	ACAATCATAG GCGCAATCGC ATTCGGTAAA ATATAACGCC ATATGATATT CCAAGTGTTT	4740
	TCACCAAGTGA TACGTGCTGC ATCTACATAT TCCATGCGTT TAATTTCTAA AACACTGGCA	4800
45	CGCATTGTCC GTGCAAATGA TGGTATATTA CCGATACTTA AAGCAATAAT TAAATTTGGA	4860
	ATACTTGCTC CAAATGATGC AATAATTGCC ACCGCTAACA ATAATGATGG AATTGCAAAC	4920
	ACTACATCTA AAATTGCGAT TATTAAATTA TCAATATGAT TAAAATAACC TGCGATAGTG	4980
50	CCTAGTAACA CACCAAAAAT AACTGCAATA ACTACTGAAA TAATTGAAAT TGAAAATGTC	5040
	AGCTTCGTTT CTACAACTAC GCGTGTAAT AAGTCTCTAC CGAAATCATC AGTACCAAAC	5100

55

	GTATCAAATG TAAATTGTGA CACAATTGAT AATGTCAGCA TGTAGACTAA AATAAGTAAC	5220
	CCGATAATCG CAATACGATG TCTAGTAGTT TTTCGTATAA ACGATTCCCA CCCGTTATAA	5280
5	CTATGTATTT GCGATGTACG TTGGTAACGT CTAATACTTA CAAACATTAA TAATGTAAAT	5340
	ACGTTGCCTG TTAATGTCAT CAACAATAAC AACACTTCGA CGATACGTCG CCATAGGTCA	5400
	TGATGCTTCC ATGTTTGTTC CGTTGTAAA ATAATAATTA AAATGATGGT TAAAACGATT	5460
10	AGCAATGTTT CAGCAATATA GAACGTATCG GCCACATAAC CTTTAAAAAG ATTTAATGCA	5520
	CTCGTTAATA TAACTAAAAT ATAAGTTGCT ATGGCGTAAC TTGCGAATAA TTTTAAGGAA	5580
	GCTATCTTTG AATTAAGTTG TGCCATATGC CTCACTTCCT TCGTTGATT TCACTACGTA	5640
15	ATTTTGGATC GATTAAAGCA TAAAATATAT CAATAATTAA GTTTGCTAAA GATATTACAA	5700
	TTGATATATA TACGACCCCA CCCATGACTG CTGGAATATC AGGTATTAGT TGTMTTGGGA	5760
20	CGATATAACG CCCGATACCA TTAATGTAA ATACTTGTTC CGTCACTGCT GAACCGCCTA	5820
	GTAACCTGTC CACTAGAAGA CCAACTAACG TTACAATTGG AATAATGGCA TTTTCAAAA	5880
	TATGTTTAAT AACAACTTGT GTCGTCGATA ATCCTTTTGC ATAAGCAGTT AAAACATAAT	5940
25	CGCTGCGCAT TACTTCAAGT ACAGAAGACC TTGTCATACG CGTGATAGAA GCAGCAATAC	6000
	TTGTTCCAAT GACAAGTACA GGTAATCA ACGATATTGG ATGTTCTGGC ATATAAGATG	6060
	GTGGCAAAT ATCCAATTTT AATGAGAACG CTAAAATGAA TAATAGCCCT TGCCAGAAAC	6120
30	TTGGAATAGA TAAACCAATT AATGCAATTA TCATTACGT GATATCAAGC CAACTATTTT	6180
	GCTTCATCGC ACTGATAATA CCAATTGGTA TTGCAATAAT TAATGCCACC ATTAGCGCTA	6240
	ATACTGCGAC AATTATTGTA ATTGGAATTC TTTGCGCAAC TGCTTTAGTC ACAACCTCAT	6300
35	TCCCTTTGTA AGTCGTACCT AAGTCAAAGG TAAAAACACC CTTGATGGTA TCCCACAATT	6360
	GAATZAAAATA AGGTTGTTA AGATGATGTA ATACATTGAA TTGATGTATC TGTGCCTTTG	6420
40	TTGCATTTTG TCCCAGTATG CTATAAGCCG CATCAAGCGG TGAAAATAC AGAATGGTAA	6480
	ACACACTGAC AATAACACCA ATGATGACAA TCACAGCCAT GACAATTCGT TCAAAAATAT	6540
	ATCTAACTAA TGGCTGTAAA TAAAAAGTCA ATAAGATGAA CATCGGCAAG GCCAATATCA	6600
45	CTTTGATCAT GATGAACCTA TGAAATAATA CATTTTCAAA GTATGTTGAA AAATGTGCTT	6660
	GTTCAATATT CTTTGAACTC GTATTAGAAC TTTGTGCCTT GAATATTTTT AATGCTTCTT	6720
	TATGTATTTG TGTGGATGAC TTTTGCTGCG ATAAATATTT ATATTTTTGA TGTACGCCT	6780
50	GTTCAATTTT TGAAATTTCA GAATTATTAG CGTAAAAATT TTTCTCTTA GCAGAAAAGA	6840
	AAAACTTTAT CACTGCATAT AAAAATATTG GCAAGCTTAA TACCGATAAT ACAAACCTTG	6900
55		

	CTTGTA	AAAT	AATCTTGAGT	AGATTACTAT	GATATACAAA	AGTATAGAAT	AAATTTACAC	7020
	ATTTGTG	aAT	AGGGAGGCAC	AACATCATGT	CAAATTTATT	AGAAGTCAAC	AGTCTGAATG	7080
5	TACAATTCAA	TTATGATGAA	ACTACAGTTC	AAGCGGTAAA	AAACGTCTCT	TTCGAATTAC		7140
	GAAAAAACA	TATCCTAGGT	ATTGTTGGTG	AATCAGGATC	AGGAAAAAGT	ATTACCGCTA		7200
	AATCTATTTT	AGGGCTACTA	CCAGATTATC	CAGATCACAC	ATTAACAGGA	GAAATTATTT		7260
10	TTAATGGGCA	ATCGTTAAAT	AATTTATCAA	CTTCAGCGTT	ACAACAAATT	CGAGGTAAGG		7320
	ATATTTCAAT	GATTTTTTCAA	GATCCACTCT	CTTCGTTGAA	TCCAAGATTA	ACGATTGGCA		7380
	AACAAATTAC	AGAAGTAATA	TTTCAACATA	AACGTGTATC	TAAATCTGAA	GCAAAGTCGA		7440
15	TGACAATAGA	CATTTTAGAA	AAAGTAGGTA	TAAAACATGC	AACTCGACAA	TTTGATGCTT		7500
	ATCCACATGA	ACTTTCTGGT	GGTATGCGTC	AACGTGTCAT	GATAGCAATG	GCATTGATTT		7560
20	TAAAGCCACA	AATTTTAAATC	GCAGATGAaC	CAACAACGGC	ATTAGATGCC	AGTACACAAA		7620
	ATCAATTACT	GCAGTTAATG	AAGTCCCTTT	ATGAGTACAC	AGAAACATCT	ATTATTTTTTA		7680
	TCACTCACGA	TTTAGGCGCT	GTGTATCAAT	TTTGCGACGA	TGTGATTGTA	ATGAAAGATG		7740
25	GAAGTGTCTG	TGAAAGTGGC	ACGGTTGAAA	GTATTTTTTAA	ATCGCCACAA	CATACCTATA		7800
	CAAAACGCTT	AATAGATGCG	ATTCCTGATA	TTCATCAAAC	GCGTCCGCCA	AGACCGTTAA		7860
	ACAATGATAT	TTTATTAAAA	TTCGATCGCG	TGAGyGgGAT	TACACATCAC	CGAGTGGCAG		7920
30	CCTATACCGA	GCAGTTAATG	ATATTAACCT	GGCTATTAGA	AAAGGCGAAA	CATTAGGCAT		7980
	TGTCGGTGAA	TCAGGGTCAG	GGAAATCGAC	ATTAGCTAAG	ACGGTCGTCG	GTCTAAAGGA		8040
	AGTGTGAGAA	GGCTTTATTT	GGTATAACGA	ATTACCATTA	AGTTTATTTA	AAGATGATGA		8100
35	ATTGAAATCT	TTACGACAAG	AGATACAAAT	GATTTTTTCAA	GATCCATTCTG	CATCTATTAA		8160
	TCCAAGATTT	AAAGTCATTG	ATGTGATTAA	ACGACCACTA	ATCATTCTATG	GGAAAGTCAA		8220
40	AGATAATGAT	GACATTATTA	AAACTGTCGT	ATCGTTGTTA	GAAAAGGTTG	GCCTAGATCA		8280
	AACTTTCTTA	TATCGCTATC	CAACGAATT	ATCTGGTGGG	CAACGTCAGC	GTGTAAGTAT		8340
	CGCGAGAGCA	CTTGCTGTTG	AACCTAAAGT	GATTGTTTGC	GACGAGGCAG	TGTCCGCTTT		8400
45	AGACGTTTCA	ATTCAAAAAG	ATATCATCGA	GTTATTAAAA	CAATTACAGT	TAGACTTCGG		8460
	CATCACTTAT	TTATTCATCA	CACATGACAT	GGGTGTTATC	AATGAAATAT	GTGATCGCGT		8520
	TGCAGTTATG	AAAAATGGCG	AAATCGTTGA	ACTGAATAAC	ACAGAAGATA	TTATCAAACA		8580
50	TCCGCAGTCA	GACTATGCAA	AGCAACTTAT	TTCAGAAGTA	GCAGTTATTG	CTAAATAAAA		8640
	GTCATGCGTT	GTGCAACTTT	ATCACTGTAT	GGTCTGAAAT	AAATTGCGCG	ACTTCTGATG		8700

55

5 TATCAAGTTT TAGGTGCTTT GCCATGATTT AAGAGTCACC CCCATACTTT GGGCATTTTA 8820
 ACGCCAGAAT AAATCCCCCG CCACTATGTG AAGTGTGGGG GATTATTTAT ATTTTATTAG 8880
 AATATTGAGA TTTTGTAGTG TGTCAACTTA GCTTAGTCAA TGTATATTTA ACGTCACTTA 8940
 CTCTTTTCTT TTCATAATTA ACACATTCAA ATAACTTTG ATCAAAAAC ACAAAGTTAA 9000
 10 AAGTACCATC TTGTAATATG CTCTCATACA TTATCCCGTC ATATTTAAGG CTTCGAATAT 9060
 AATCAGCTAA ATATTGAAAT GGCAAATAAT CTATTCCTTG TTCATCGCTT GGATTTGTTA 9120
 TTCCTTTATG AATCTTTTTT AATGTTTGGT AATTTACAAA ATACTTTCTA AATCCATCAT 9180
 15 CGCCAGCTTT GATTGCATTA CTAGTTAAAT TAGTTAAATT CGCAATTTTC AATTTCTCTT 9240
 TTGTCACGTT TTTTGTAAAC TTAACCTTAC CTATATAAAT AATGTCATTA TGCTTAGGTT 9300
 TAACTTCTTC TATACTGACC TGTCTTTTTG TACTAAGGTA TAATACGCTT ATCCATTTAG 9360
 20 AATTCAATCT TCCTGCCGTT GCAAATCCCT TTGGTGGTGA CATTAGTTCA CTTTTCTCTG 9420
 TAATGAAGTT AACTATTCTA GATCTATATA ATGGTTCAAA TCTTTCTCTA AATTCCTCAA 9480
 TACTATAGTA ATTAGTAGTG ATATCGAGAA AGAACGCTAA ATTCTCTAAA TTGATCATAT 9540
 25 TTTTATGAAA TCTATTTTAA TACTTCAAGC TCTCACAAA TCCATCCCAG TCATTATTTG 9600
 CTACAATTAG ATTTTATTTT GTATATTTTAT TATCGTTTAT GATTTTAGCG CCTACTAAAT 9660
 CTTCACACAC TCGTCTATCT AAATTTTCAT CATCTTTAAA AAGTTCATT AAAATACAAC 9720
 30 TTATTTGAGC TTCCTCAACA TTAAATATAC TCCAGTCGTC TTTTAATGCT ATTTCAATCT 9780
 TTTTACCTTC TTTTGGGCTA AAAGTATCTG GTAAATTTAT ACTAATATCA TATAATTCTA 9840
 35 ATGCTGGTCT TAAATAATCT CTAATAAGTT CTAATTTATC TATGTCCTTA GTCGTATCAA 9900
 ATATTTTAAAC ACCAAGATGA TTGTTATCAA TATCACAATT GTCAAATTTG CTATTTATCA 9960
 TTTGCAATGA TTTCTACGAT TTCAGTATTA TTAAAACATT TTTCACATAT TTTCATTTTG 10020
 40 AGACTCCAAG TATCTATTCA TAATTTCTAG GTGATGCATG ATAGATAACC TTTTAATTAA 10080
 ACCTAATCCT GGATACTTAT TATTTTCATT TAATTCCTCA AATTGTCCCA AGCGCATAAG 10140
 ATCTATTTT AATATCTAAG TTTTGTGACC ATGTTACTAA TT 10182

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3491 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

	AACTCAGGCA ATTGAAACAG CATTAGGTGC TTCATTACAA CATGTCATTG TAGATTCAGA	60
	AAAAGATGGA CGCCAGGCTA TTCAATTTTT AAAAGAACGT AATTTAGGTC GTGCGACGTT	120
5	TTTACCATTA AATGTTATAC AGAGTAGAGT GGTAGCGACT GATATTAAAT CTATTGCTAA	180
	AGAGGCAAAC GGATTTATTA GTATCGCTTC GGAAGCAGTT AAAGTAGCAC CAGAATATCA	240
	AAATATTATC GGAATTTAT TAGGTAATAC GATTATCGTT GATCATTTAA AGCATGCAAA	300
10	TGAATTGGCA CGTGCGATTA AATATCGAAC TCGTATTGTT ACTTTGGAAG GTGATATTGT	360
	AAATCCTGGT GGtTCTATGA CTGGTGGTGG CGCTCGTAAG TCAAAAAGTA TTCTGTCTCA	420
15	AAAAGACGAG TTGACAACAA TGAGACACCA ATTAGAAGAT TACTTGCGTC AAACAGAATC	480
	ATTTGAACAA CAATTTAAAG AGTTGAAGAT AAAAGTGAT CAATTAAGTG AACTGTATTT	540
	TGAAAAAAGT CAAAAGCATA ATACACTTAA AGAGCAAGTG CATCATTTTG AAATGGAGCT	600
20	CGATAGATTA ACTACACAAG AAACACAAAT AAAAAATGAT CATGaAGAAT TCGAATTGTA	660
	AAAAAATGAT GGTTATACGA GTGACAAAAG TCGACAAACT TTGAGTGAAA AAGAACTTA	720
	TCTAGAAAGT ATTAAAGCAT CTTTAAACG ACTAGAAGAT GAAATTGAAC GCTACACAAA	780
25	ACTTTCTAAA GAAGGTAAGG AAAGCGTTAC TAAAACACAA CAAACCTTAC ATCAGAAACA	840
	ATCTGATCTT GCTGTGGTTA AAGAGCGTAT TAAAACACAA CAACAGACAA TAGATCGATT	900
	AAATAATCAA AATCAACAAA CTAAACATCA ATTAAAAGAT GTTAAAGAAA AAATTGCATT	960
30	CTTTAATTCG GATGAAGTGA TGGGCGAACA AGCTTTTCAA AATATTAAAG ATCAAATTAA	1020
	TGGTCAACAA GAAACGAGAA CACGCTTATC AGATGAATTA GATAAATTGA AACAACAACG	1080
35	TATTGAGTTG AATGAACAAA TCGATGCGCA AGAAGCTAAA CTACAAGTTT GTCACCAAGA	1140
	TATTTTAGCT ATCGAAAATC ACTACCAAGA TATTAAAGCT GAACAATCAA AGCTAGATGT	1200
	ATTAAATTCAT CATGCGATAG ATCATTaAAT GATGrATATC AATTGACTGT TGAACGTGCG	1260
40	AxATCTGAAT ATACGaGTGA TGxATCGATg ACGCATTACG TAAAAAAGTT AAGTTAATGr	1320
	AGaTGyCGAT TGATGrACTA GGTCCTGTAA ACTTAAATGC AATTGAACAA TTTGAAGAGT	1380
	TAAATGAACG TTATACATTT TTAAGTGAAC AACGTACAGA TCTTCGTAAA GCTAAAGAAA	1440
45	CATTAGAGCA AATTATAAGT GAAATGGATC AAGAGGTTAC TGAAAGATTT AAAGAACTT	1500
	TCCATGCTAT TCAAGGACAT TTTACAGCTG TGTTCAAACA ATTGTTTGGT GGAGGCGATG	1560
	CAGAATTGCA ATTAAGTGAA GCCGATTATT TAACAGCTGG TATTGATATT GTGGtACAAC	1620
50	CACCGGGTAA AAAGTTGCAA CATTTATCGT TACTGAGTGG TGGTGAGCGT GCATTAACTG	1680
	CTATTGCTTT ACTATTTGCA ATTTTAAAAG TAAGATCTGC ACCTTTTGTT ATATTAGrTG	1740

55

	TATCAGACGA AACACAATTC ATTGTTATTA CACACCGTAA AGGAACAATG GAATTTGCAG	1860
	ATAGGTTATA CGGTGTAACA ATGCAAGAAT CAGGTGTTAC TAAACTTG TG AGTGTGAATT	1920
5	TAAATACAAT AGATGATGTG TTGAAGGAGG AGCAATAATG AGCTTTTTTA AACGCTTAA	1980
	AGATAAGTTT GCAACAAATA AAGAAAATGA AGAAGTTAAA TCCTTAACAG AAGAACAAGG	2040
10	TCAAGACAAA TTAGAAGATA CACATTCTGA AGGTTCAACG CAGGACGCAA ATGATTTAGC	2100
	AGAAAATGCT GAAGTGAAAA AGAAGCCACG CAAGTTGAGT GAAGCGGATT TTGATGACGA	2160
	TGGCTTAATA TCAATTGAAG ATTTTGAAGA AATTGAAGCT CAAAAAATGG GTGCTAAATT	2220
15	TAAAGCAGGA CTCGAAAAAT CTCGTCAAAA TTTCCAAGAA CAATTAAATA ATTTGATAGC	2280
	GAGATATCGT AAAGTAGATG AAGACTTTTT TGAAGCTTTA GAAGAAATGT TAATCACTGC	2340
	AGACGTCGGT TTTAATACAG TGATGACGTT AACTGAAGAA TTACGTATGG AAGCACAACG	2400
20	ACGTAATATT CAAGATACTG AAGATTTGCG TGAAGTCATT GTTGAAAAGA TCGTAGAGAT	2460
	TTACCATCAA GAAGATKATA ATTCAGAAGC TATGAACCTA GAAGATGGTC GTTTAAATGT	2520
	CATTTTAATG GTTGGTGTGA ATGGTGTGG TAAAACAACA ACAATTGGAA AATTAGCTTA	2580
25	CCGATATAAA ATGGAAGGTA AAAAAGTAAT GTTAGCTGCG GCGGATACTT TTAGAGCGGG	2640
	TGCTATTGAT CAATTGAAAG TTTGGGGCGA ACGTGTGGT GTAGACGTAA TTAGCCAAAG	2700
30	TGAAGGTTCT GATCCAGCTG CTGTTATGTA TGATGCGATT AATGCCGCTA AAAACAAAGG	2760
	TGTTGATATT TTAATCTGTG ATACCGCTGG ACGTTTACAA AATAAACAA ATCTAATGCM	2820
	AGAATTAGAA AAAGTTAAGC GTGTAATTAA TCGAGCAGTG CCAGATGCGC CTCATGAAGC	2880
35	ATTACTATGT TTAGATGCTA CAACTGGTCA GAATGCGTTG TCACAAGCTA GAAACTTTAA	2940
	AGAAGTAACA AATGTTACAG GTATTGTATT AACGAAATTA GATGGTACAG CCAAAGGTGG	3000
	TATCGTATTA GCCATTGCTA ATGAATTGCA CATCCAGTT AAATATGTAG GTTTAGGTGA	3060
40	GCAATTAGAT GACTTACAAC CATTTAACCC TGAAAGTTAT GTCTACGGCT TATTGCTGA	3120
	TATGATTGAA CAAATGAAG AAATAACAAC AGTTGAAAAT GATCAAATTG TAACAGAAGA	3180
	AAAGGACGAT AATCATGGGT CAAATGATT TAGTTAAAAC GTTACGAATG AATTATTTGT	3240
45	TTGATTTTAT CAATCCTTAT TGACGAATAA ACAACGTaAT TATTTGGAAT TATTTTATCT	3300
	TGAAGATTAT TCTTTAAGTG AAATCGCAGa TACTTTTAAAT GTGAGTAGaC AAGCAGTTTA	3360
50	TGATAATATA AGAAGAACTG GCGATTTAGT TGAAGATTAT GAAAAGAAAT TGGAATTATA	3420
	CCAGAAATTT GAGCAACGCC GAGAAATATA TGATGAAATG AAACCACATT TAAGTAATCC	3480
	AGAACAAATA C	3491

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4253 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

10	AGTACGTTTT ATAATTATAA GTACGTAATT AACATATTAA CATATCGCAA GTATGTATTT	60
	AAATAAgATT GTTATAATTT CAAAGTTCAT CCAAGaTTAT GGCgTTTGCA TTTACCTATT	120
15	AAAAACGTTA TTATATCAAA GATGCGAAAG ATAATACGGG TTTATTTTAT GAAAGTGAGA	180
	AGGATAAAAT GGATAATGAG CAACGCTTAA AAAGAAGAGA GAATATAAGG AATTTCTCGA	240
	TTATAGCACA TATTGACCAC GGAAAATCTA CATTGGCTGA TAGAATTTTA GAAAATACCA	300
20	AATCAGTTGA AACAAGAGAT ATGCAAGATC AGTTACTAGA TTCAATGGAT TTAGAAAGAG	360
	AACGTGGTAT TACAATCAAA TTAACGCgT ACGTTTAAAG TACGAAGCTA AAGATGGAAA	420
	TACTTATACA TTCCATTTAA TCGATACGCC TGGACACGTC GATTTTACAT ATGAAGTGTC	480
25	ACGTTcTTTG GCAGCTTGTG AGGGCGCGAT TTTAGTAGTA GATGCGGCTC AAGGTATCGA	540
	AGCACAACA TTAGCAAATG TTTATTTAGC ATTAGATAAT GAGTTAGAGT TATTGCCTGT	600
	TATTAACAAA ATTGATTTAC CTGCTGCAGA ACCTGAACGC GTGAAACAAG AAATTGAAGA	660
30	TATGATAGGT TTAGACCAAG ACGATGTTGT TTTAGCAAGT GCTAAATCTA ACATTGGAAT	720
	TGAAGAGATA CTAGAGAAAA TAGTTGAAGT TGTGCCAGCT CCAGATGGTG ACCCAGAAGC	780
35	ACCACTAAAA GCGTTAATAT TTGATTCTGA GTATGATCCA TATAGAGGGG TAATTTTCATC	840
	GATAAGAATT GTGGACGGTG TTGTTAAAGC CGGAGATAAA ATTCGAATGA TGGCCACTGG	900
	TAAAGAGTTC GAAGTAACAG AAGTTGGAAT TAATACACCT AAGCAGCTTC CAGTTGATGA	960
40	ATTAACAGTT GGTGATGTTG GTTATATTAT TGCAAGTATT AAAAATGTTG ATGATTCTAG	1020
	GGTTGGTGAC ACCATCACAT TAGCTAGTAG ACCTGCATCA GAACCATTC AAGGTTATAA	1080
	GAAAATGAAT CCAATGGTAT ATTGCGGACT GTTCCCAATA GATAACAAAA ATTATAATGA	1140
45	TTTAAGAGAA GCATTAGAAA AATTACAATT GAATGATGCA TCATTAGAAT TTGAGCCTGA	1200
	ATCGTCACAA GCATTAGGTT TTGGTTATAG AACTGGTTTC TTAGGTATGT TACACATGGA	1260
	AATAATTCAA GAAAGAATTG AAAGAGAATT TGGTATTGAA TTAATTGCAA CTGCACCATC	1320
50	TGTAATTTAT CAATGTGTTT TAAGGGACGG TTCAGAAGTG ACGGTTGATA ACCCAGCACA	1380
	AATGCCAGAT CGTGATAAAA TTGATAAAAT ATTTGAGCCA TATGTTCTGT CAaCTATGAT	1440

	TATAAATATG GACTATTTAG ATGATATTCG TGTAAATATT GTTTATGAAT TACCTTTAGC	1560
	TGAAGTTGTA TTTGATTTCT TCGATCAACT TAAATCTAAT ACTAAAGGAT ATGCATCATT	1620
5	5 TGATTATGAA TTCATCGAAA ATAAAGAAAG TAATTTAGTC AAGATGGATA TTTTATTAAA	1680
	TGGTGATAAA GTGGATGCGC TAAGCTTCAT AGTTCATAGA GATTTTGCAT ATGAACGTGG	1740
10	10 TAAAGCATTG GTTGAAAAAC TTAAAACGTT AATTCCAAGA CAGCAATTTG AAGTACCTGT	1800
	ACAGGCTGCA ATAGGACAAA AAATTGTAGC GCGTACAAAT ATTAAATCAA TGGGTAAAAA	1860
	CGTTTTAGCT AAATGTTATG GCGGTGACAT AAGCCGTAAA CGTAAATTAC TTGAAAAACA	1920
15	15 AAAAGCAGGT AAAGCTAAGA TGAAAGCAGT TGGTAATGTT GAAATCCAC AAGATGCTTT	1980
	CTTGGCTGTA TTGAAAATGG ATGATGAATA ATTTTAAAAA ATCAATTAAC AATTTACAAT	2040
	GAATAAAGTT TAATAACTAA AAAGAGGGAG CCTAGGATAA ATTAACGTCC TGGGCTTTAC	2100
20	20 AATGTTATAT TGGCAGCCAT CGACAGAGTT AAAATGAGCT TATAACAATG GGGCCCCAAC	2160
	ACAGAAGCTG ACGAAAAGTC AGCTTACTAT AATGTGCAAG TTGGGGTGGG GCCCCAACAT	2220
	AGAGAATTTT GAAAAGAAAT TCTACAGGCA ATGCAAGTTG GGGTGGGACG ACGAAATAAA	2280
25	25 TTTTGCGAAA ATATCATTTT TGTCCCACTC CCTTATGCAT GAGTTTTACT CATGTAATTT	2340
	TATTTTTAAG GACATATTAC ATCTGGCTAA TGTGTAAGAG CCACTACATA ATAAATCATT	2400
30	30 AGTGGTTCTT TATTATTTCT ATCTCACTCC CTCTAAACAA GAATAAATAT TAAATGAAT	2460
	CGATATATTA GACAATCATT GATTAAACGT TAAAGTTAAA AGTAAGAATA ATTGCAGATA	2520
	GTCCAACAGG ATATAGCCGA TTGGATAAAA AGTCTGAGAA GCGGGGCATT AAAATGACGG	2580
35	35 TACAAAGTGC ATATATACAT ATTCCATTTT GTGTAAGAAT ATGTACATAT TGTGATTTCA	2640
	ATAAATATTT TATACAGAAT CAACCTGTAG ATGAGTACTT AGATGCACTA ATCACAGAAA	2700
	TGTCTACAGC AAAATATAGG ATCTTAAAGA CCATGTATGT AGGTGGCGGC ACACCAACGG	2760
40	40 CCCTTTCTAT TAATCAGTTG GAAAGATTAC TTAAAGCAAT ACGTGATACG TTTACAATCA	2820
	CAGGCGAGTA TACATTTGAA GCAAATCCTG ATGAGTTAAC TAAAGAGAAA GTCCAACTAT	2880
	TAGAGAAATA TGGAGTAAAA AGGATTTCAA TGGGCGTTCA AACATTCAAG CCGGAGTTAT	2940
45	45 TGTCTGTTTT AGGTAGAACG CACAATACTG AAGATATTTA CACTTCGGTG TTAAATGCTA	3000
	AAAACGCAGG TATTAAATCA ATCAGTTTAG ATTTAATGTA TCATTTACCG AAACAGACGA	3060
50	50 TTGAAGATTT TGAACAAAGT TTAGATCTAG CTTTAGATAT GGATATTCAA CATATTTCGA	3120
	GTTACGGCTT AATACTTGAA CCTAAAACCC AATTTTATAA TATGTATAGA AAAGGCTTGC	3180
	TCAAACCTGC TAATGAGGAT TTAGGTGCTG ACATGTATCA GTTGCTGATG TCTAAGATAG	3240

55

AACATAATAA GGTTTACTGG TTTAATGAGG AATATTATGG ATTTGGAGCA GGTGCAAGTG 3360
 GTTATGTAGA TGGTGTGCGT TATACGAATA TCAATCCAGT GAATCATTAT ATCAAAGCTA 3420
 5 TAAATAAAGA AAGTAAAGCA ATTTTAGTAT CAAATAAACC TTCTTTGACT GAGAGAATGG 3480
 AAGAAGAAAT GTTTCTTGGG TTGCGTTTAA ATGAAGGTGT GAGTAGTAGT AGGTTCAAAA 3540
 10 AGAAGTTTGA CCAATCTATT GAAAGTGTCT TTGGTCAAAC AATAAATAAT TTAAAAGAGA 3600
 AGGAATTAAT TGTAGAAAAG AACGATGTGA TTGCACCTAC AAATAGAGGG AAAGTCATAG 3660
 GTAATGAGGT TTTTGAAGCT TTCCTAATAA ATGATTAAAA AAAATTGAAA TTTCGAGTCT 3720
 15 TTAACATTGA CTTACTTTGA CCAATTTGAT AAATTATAAT TAGCACTTGA GATAAGTGAG 3780
 TGCTAATGAG GTGAAAACAT GATTACAGAT AGGCAATTGA GTATATTAAA CGCAATTGTT 3840
 GAGGATTATG TTGATTTTGG ACAACCCGTT GGTTCATAAA CACTAATTGA GCGACATAAC 3900
 20 TTGAATGTTA GTCCTGCTAC AATTAGAAAT GAGATGAAAC AGCTTGAAGA TTAAACTAT 3960
 ATCGAGAAGA CACATAGTTC TTCAGGGCGT TCGCCATCAC AATTAGGTTT TAGGTATTAT 4020
 GTCAATCGTT TACTTGAACA AACATCTCAT CAAAAACAA ATAAATTAAG ACGATTAAAT 4080
 25 CAATTGTTAG TTGAGAATCA ATATGATGTA TCATCAGCAT TGACATATTT TGCAGATGAA 4140
 TTATCAAATA TATCTCAATA TACAACCTTA GTTGTTTCATC CTAATCATAA ACAAGATATT 4200
 ATCAATAATG TACACTTGAT TCGTGCTAAT CCTAATTTAG TTATAATGGT TAT 4253
 30

(2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3395 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

TCCCTAATCG AACAAAATTA TGCGCATAAA CAAAGTAGAT TGATATAAAA TTCTTAATTA 60
 TCAGAAATATA TTTACAAATC TGAATTTTAT TAGTATATTG GrTAGTrTTC ATAGAGGCAT 120
 45 GACGGTaTTT GAGCAGGATT TTAAATCGGg ATTTTATAAT CGATTTAAGA GAGGCCACtT 180
 TGCTTGcACA TTAATACTGT cAATGGGAGG GGAATGTATA TGAGTrAAGC ACATCAATTA 240
 ATTCAAGAGG ATGAACATTA TTTTGCGAAA TCAGGACGTA TTAAATATTA TCCGTTAGTG 300
 50 ATTGATCATG GATATGGAGC AACATTGGTT GATATTGAGG GGAAGACATA TATCGATTTG 360
 TTATCGAGTG CGAGTTCTCA AAACGTAGGT CATGCACCTA GAGAAGTAAC AGAAGCGATA 420

	GTACGTTTAG	CTAAGAAGCT	TTGTGAGATT	GCACCTGGAG	ATTTTGAAAA	AAGAGTGACC	540
	TTCGGATTAA	CCGGATCAGA	CGCAAATGAT	GGCATCATT	AATTTGCCAG	AGCATATACA	600
5	GGGCGTCCTT	ATATCATTAG	TTTCACTAAT	GCATATCATG	GTTCAACTTT	TGGCTCATTG	660
	TCTATGTCAG	CTATTAGTTT	AAATATGCGC	AAACATTATG	GTCCGTTATT	GAATGGTTTT	720
10	TATCATATTC	CGTTTCCAGA	TAAATATCGT	GGTATGTACG	AGCAGCCACA	AGCTAATTCA	780
	GTAGAAGAAT	ATTTAGCACC	CTTAAAAGAA	ATGTTTGCGA	AGTATGTACC	TGCTGACGAA	840
	GTAGCATGTA	TTGTTATTGA	AACGATACAA	GGCGATGGTG	GACTTTTAGA	ACCAGTTCCA	900
15	GGGTATTTTG	AAGCGTTAGA	AAAGATTTGT	CGTGAACATG	GTATTTTAAT	CGCTGTCGAT	960
	GATATTCAAC	AAGGTTTTTG	GAGAACAGGT	ACATGGAGTT	CAGTCTCGCA	TTTTAATTTT	1020
	ACGCCTGATT	TAATCACTTT	CGGAAAATCC	TTAGCAGGTG	GTATGCCTAT	GTCAGCAATT	1080
20	GTTGGACGCA	AAGAGATTAT	GAATTGTTTA	GAAGCACCAG	CACATTTATT	TACAACAGGT	1140
	GCTAATCCAG	TTAGTTGTGA	AGCTGCATTA	GCCACAATTC	AAATGATTGA	AGATCAGTCG	1200
	CTTCTTCAGG	CTAGTGCGGA	AAAAGGGGAA	TATGTTAGGA	AACGAATGGA	TCAATGGGTA	1260
25	TCTAAATACA	ATAGTGTAGG	CGATGTTAGA	GGTAAAGGTC	TGAGCATTGG	TATTGATATT	1320
	GTTTCCGACA	AAAACTCAA	AACACGTGAT	GCCAGTGCGG	CACTTAAAAT	TTGTAATTAC	1380
30	TGCTTTGAGC	ATGGCGTAGT	TATTATAGCT	GTAGCAGGAA	ATGTGTTGCG	ATTCCAACCG	1440
	CCATTGGTAA	TAACATATGA	GCAATTAGAC	ACGGCGTTAA	ACACTATAGA	AGATGCACTG	1500
	ACTGCTTTGG	AAGCAGGTAA	CTTAGATCAA	TATGACATAT	CTGGACAAGG	TTGGTAATAG	1560
35	CGATTATCTT	AATATAAAAT	AAAAAATCAT	TTCCACATCT	GGATGTTAAT	CAGATGGGAA	1620
	ATGATTTTTT	TTATTTTTTA	TTTTGGTGGG	TGGTATTTCAG	CTACGTCATT	TTTCTTAGAA	1680
	TGTGTAAGTC	CATAACTTAA	ATATAGGATG	ATACCAACAA	TAAACCAAAT	TAAAGTGTAT	1740
40	AAATTCGCTT	CGAATCCTAA	TCCCAGAAAT	ACTAGCAATA	CTAAAACAAA	TGTAATTGCT	1800
	GGTAACACAG	GATATAAAGG	TAATTTAAAT	GCAGGAATTG	GATAGTCTTT	ACCTTCACGC	1860
	TTTCTCAAAC	GATACATTGC	TAATGAAACG	AACATAAATG	CAACAAGTGT	ACCTGCTGAA	1920
45	ATTAATTGTG	CTAAAAATGC	GAATGGGAAC	ATAGAACCAA	TTAAAACACC	AATAATAGTA	1980
	AGTATAACTA	GTGCGCGATT	AGGTAAATGT	TTGTCGTTTA	AGTGGCTTAA	CCATGAAGGT	2040
50	AATAAGCCGT	CACGTCCAAA	TGAATAAAGT	AAACGTGAGC	CTGCTAACAT	CATACCAATT	2100
	AATGCTGTAA	ACATACCGAT	AACAGAGATA	GCTTGAACAA	TAGCTGCTAC	AACACCATGA	2160
	CCACTTTGAC	GTAAAGCCCA	ACCAACAGGT	TCAGCATTGT	TTGCGTATTG	TGAGTAATGG	2220

55

	CCAAGAATAC CTCTAGGCAT TGTCTTTTGA GGATCAAGTG CTTCTGCTGA GTTTGCTGCG	2340
	ATAGAATCGA AACCGATATA CGCTAAGAAA ATCATTGAAA CACCAGCATA TATGCCTTGC	2400
5	CATCCACCAA AGTCACCTGT AGCAGTTACT TTGTGTTCTG GAATAAATGG CACATAGTTA	2460
	CTAACATTTA TTGCTGTAA ACCTACGATG ACAAATAAAA TAATAGCTAA TACTTTTAAA	2520
	ATAACTAAAA TATTTTCCAT ACGAGCTGCT TCCGACATAC CACGTGATAG TAATAATGCA	2580
10	GTTAATAAAA TAACGATAGC AGCAATAATA TCGATAAAAC CGCCATTTGT ACCAAATGGA	2640
	TTTGATAATG CTGCAGGTAA TTCGATGCCA ATTGGTTTCA CAAGTCCGCG TAAATTCGCT	2700
	GAGAATCCTG ATGCAACAAA GGCTACGGCG ATAAAATATT CAGCTAATAG AGCCCAACCG	2760
15	GCAACCCATC CAAAAAATTC ACCAAATAAT ACATTGACCC AAGAATAGGC TGAACCTGCA	2820
	AATGGCATAG CGGCAGCCAT TTCTGCATAA GTAAATGCAA CTAAACCAGC AACAAATAGCA	2880
20	GCGAGTAAGA ATGATAACGC AACGGCCGGT CCTGCATGTT CTGCAGCAAC AATGCCAGGT	2940
	AGCGTAAAGA TAGATGTCGA TACAATTGTT CCTACACCTA AAGCTAAGAA ATCACGCACC	3000
	CGAAGTGTA GCTTTAAATG ACCATCTTTA TTTTGATAGA TAGCCGGATC CTCTTTTCGT	3060
25	GCTATTTTAT TGAAAAAACT TCCCATAAAC TTTCTCCCA AACATTCATA AACAAATCTA	3120
	TACGGTGT TTTAATATGT TATATCATAG CACAAATAAT CAATATTTTG TCTAAAAATT	3180
	CTGAAAAATC ACAACTTTAT GTTACGTATT AATGACTTGT CTTGATAACA TCCATAGATT	3240
30	TTTTAAATGA TAAAACTGAT TATAACAGAT ATTAAATGAA TAAGTACTAT TTTTTCnAA	3300
	TTTTCTAACA ATTTTGCACA TTATATGTTT AAAATCAATT TCATGTTTAT GGTCTGATTG	3360
35	GCTAGTGTGT ATGAAATGTA AntCTTTGAC TnnGA	3395

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13508 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

	ATCAGGTAAT GCCATGCGTT TAGCTGAAAA TTTTTCAGA ACGTTTAAGT GATATCGGAC	60
	ATCAAGTTGT TTTGATGTCA ATGGATGAAT ATGATACGAC AAACATCGCG CAGTTAGAAG	120
50	ATTTATTTAT TATTACGTCT ACTCATGGTG AAGGAGAACC GCCTGATAAT GCATGGGATT	180
	TCTTTGAATT TTTAGAAGAC GATAACGCAC CTAATTTAAA TCATGTGAGA TATTCAGTAC	240

	TACTAGAAAA TCTAGGCGCT GAGCGTATAT GTAAGCGTGT AGATTGTGAT ATTGATTATG	360
	AAGAAGACGC AGAAAAGTGG ATGGCAGACA TCATTAATAT TATTGATACC ACATCAGAAG	420
5	GTATTCAAAG TGAATCGGTG ATAAGTGAAT CAATTAAGTC TGCCAAAGAA AAGAAATATT	480
	CTAAATCAAA TCCATACCAA GCAGAAGTAT TAGCGAATAT CAATTTAAAT GGTACCGATT	540
10	CAAATAAAGA AACACGACAT ATAGAATTTT TACTTGATGA TTTTAGTGAA TCATATGAAC	600
	CAGGAGATTG TATAGTAGCA TTACCGCAAA ACGACCCTGA ATTGGTTGAA AAATAATAT	660
	CCATGTTAGG TTGGGATCCG CAATCTCCGG TGCCAATTAA TGATCATGGT GATACAGTTC	720
15	CTATTGTTGA AGCACTAACA TCACATTTTG AATTTACTAA ATTAACATTG CCATTATTGA	780
	AAAATGCAGA TATCTATTTT GACAATGAAG AATTATCTGA ACGTATTCAA GATGAGTCAT	840
	GGGCGCGTGA ATATGTTATA AATCGGGACT TTATAGATTT AATAACAGAT TTTCCAACTA	900
20	TAGAATTACA ACCTGAGAAT ATGTATCAAA TCCTTAGAAA ATTACCACCA AGAGAGTATT	960
	CGATTTCIAG TAGTTTATG GCAACGCCAG ATGAAGTGCA TATTACCGTT GGTACGGTTC	1020
	GTTATCAAGC ACATGGACGT GAGAGAAAAG GTGTATGCTC GGTTCATTTT GCTGAGCGAA	1080
25	TTAAACCAGG CGATATAGTA CCAATTTATT TGAAGAAAAA TCCGAACTTC AAATTTCCGA	1140
	TGAAGCAAGA TATACCGGTT ATTATGATTG GACCAGGTAC TGAAATTGCT CCTTTTAGAG	1200
	CATATTTACA AGAACGTGAA GAACCTGGTA TGAAGGAAA AACATGGTTG TTCTTTGGTG	1260
30	ATCAACACCG TAGTTCTGAC TTTTATATG AAGAAGAAAT AGAAGAATGG CTTGAAAATG	1320
	GAACTTAAC ACGCGTAGAT TTAGCATTTT CAAGAGACCA AGAACACAAA GAATATGTAC	1380
35	AGCATCGTAT AATGGAAGAA AGTAAACGTT TCAATGAATG GATTGAGCAA GGCGCACAA	1440
	CTATATTTGT GGCATGAAA AATGTATGGC GAAAGATGTC CATCAAGCCA TTAAAGATGT	1500
	ATTGGTAAAA GAACGTCATA TTTCTCAAGA AGAAGCAGAG TTATTATTGC GACAAATGAA	1560
40	ACAACAACAA CGCTATCAAC GTGATGTTTA TTAGCGATTG GTGTTAAATA TTTTAAGGTG	1620
	TAATGATGTA AAAAGATATA AAGGATGTTG CTCAACATGA ATATGCCATT AATGATAGAT	1680
	TTAACAAATA AAAATGTCGT CATAGTTGGT GGAGGCGTCG TTGCAAGTCG TCGGGCACAA	1740
45	ACATTAAATC AATACGTTGA ACATATGACG GTCATCAGTC CGACAATCAC TGAAAACTT	1800
	CAAAATATGG TAGATAACGG TGTCGTCATA TGGAAAGAAA AAGAATTTGA ACCAAGCGAT	1860
	ATTGTAGACG CGTATCTAGT TATTGCAGCA ACCAATGAGC CACGTGTCAA TGAAGCGGTA	1920
50	AAAAAAGCCT TACCTGAGCA TGCCCTTTTT AATAATGTTG GAGATGCATC AAATGGCAAT	1980
	GTTGTATTTT CAAGTGCACT ACACCGCGAC AAGCTAACTA TCAGTGTATC AACTGATGGT	2040
55		

	TACAGTTCGT ATATCGACTT TTTATATACT TGCCGACAGA AAATAAAAGT ACTTGATATA	2160
	ACATATAACG AAAAGCAACA GTTACTGTCA CAAATTGTGT CACAAGAATA TTTAAATCAT	2220
5	GACAAACAAG CTCAATTTTT AGCGTGGTTG GATGTAAGAT AATAATAGCG GACCGTCTAA	2280
	CCGTCTAAGG TAAGTCTTCT TATTTTAACT TTAACGCTTA ATCATTGAAA TTAAGACATG	2340
	GGCGGCTTTG TGAATAGTCT AATAATGAAG GATTTAAGCG ATAATGATAT GCGTTTTTAA	2400
10	TATGAATATT ACAATAGAGA AAAAGATACG TAGAACAAAC TTAATAAAAT AGGTGGATAA	2460
	ATTGAAATCT GGTGGAAGTC GTTACTATCA TAGCGACCTT TAGCCAGATT TTTTGTGCAA	2520
	TAGAAAGCAA TAATAAAAAT GATAGATCAA AATGAAATAC AGGACAGGAT ATACAAGGAT	2580
15	TAGTCATGCC ATGTTATCAA GTAGGAAAAT CAAACTTCAC TATTGATAGT TACGCAAAAA	2640
	AGATTTTTTT GATAAAATGA GATAACTTAA ATATAAAAAA TTATATTAAT TATAATATTT	2700
20	AAGTTAAAGA GGGGGATTAT GTAAATTGTA TTAAGAGTGG AGGGAGAAAA TAATATGAAT	2760
	AGTGATAATA TGTGGTTAAC AGTAATGGGG CTCATTATTA TTATTTCAAT TGTAGGTTTA	2820
	CTCATTGCCA AAAAGATAAA TCCAGTTGTA GGTATGACAA TCATACCTTG CTTAGGGGCA	2880
25	ATGATTTTAG GATATAGTGT GACAGATTG GTTGGATTTT TTGCTAAAGG GTTAGATCAA	2940
	GTCATCAACG TTGTTATTAT GTTTATCTTT GCCATTATTT TCTTTGGCAT CATGAACGAT	3000
	AGTGGTTTAT TCAAGCCGCT TGTCAAACGC TTAATATTAA TGACACGAGG CAATGTCGTC	3060
30	ATTGTCTGTG CAATGACAGC TTTAATTGGC ACAATAGCCC AATTAGATGG GGCCGGTGCG	3120
	GTAACATTTT TGCTTTCTAT TCCTGCATTA TTACCTTTAT ATAAAGCGTT AAATATGAAT	3180
	AAATATTTAT TGATTTTACT ATTAGCATTA AGCGCGGCGA TTATGAACAT GGTACCTTGG	3240
35	GGAGGTCCAA TGGCTCGTGT AGCTGCAGTG TTAAGGCCA AAAGTGTCAA TGAATTATGG	3300
	TATGGATTAA TACCTATTCA AATAATAGGT TTCATTCTTG TTATGTTGTT TGCGGTATAT	3360
40	CTTGGATTTA AAGAACAGAA ACGTATCAAA AAAGCAATAG AGAGAAATGA ATTACCGCAA	3420
	ACACAAGATA TAGATGTACA TAAATTAGTT GAAGTATATG AACGAGATCA AGATGTAAGG	3480
	TTTCCTGTAA AAGGACGTGC AAGAACAAAA TCATGGATAA AATGGGTGAA TACAGCTTTA	3540
45	ACTTTAGCTG TTATTCTATC GATGTTAATA AATATTGCGC CACCTGAATT TGCATTCTATG	3600
	ATAGGTGTTT CGTTGGCACT TGTATTAAAT TTTAAATCAG TGGATGAACA AATGGAACGA	3660
	TTAAGAGCGC ATGCGCCGAA TGCATTAATG ATGGCTGCAG TGATTATTGC AGCAGGTATG	3720
50	TTTTTAGGTG TACTAAATGA AACCGGTATG CTTAAAGCGA TTGCGACCAA TTTAATCAAA	3780
	GTGATTCCTG CAGAAGTAGG ACCATACTTG CATATTATTG TAGGTTTACT TGGCGTACCA	3840

55

	ACAGCAGGGC AATTGTTGT ACCGTCTGTA TCAACAGCTT ATTCAATGGT CATAGGGAAT	3960
	ATTATAGGTA CATTTGTCAG CCCATTTTCA CCAGCCTTAT GGTGGAAT TGGTTTAGCA	4020
5	GAGGCAAACA TGGGCACGTA TATTAAGTAT GCATTCTTTT GGATTGTTGGG ATTCGCTATC	4080
	GTTATGTTAG TAATTGCAAT GTTGATGGGC ATTGTGACGA TTTAAGTATG AAAAAATAGA	4140
	AACTATGGTC ACGTTGCAAA ATGAAATAAT AGTTGCATAA ACATGTCGAA ATGACGGACG	4200
10	AATCTTTTAA CAATTTTAAA AATTAATGAA ATAATTGTGT AGAAATATGA ATTTCACTAA	4260
	ATGTTAATAA CTTTGTGACG TTTTAGTTAA CAGACTAATA AAAATTTGAA AATACTATAT	4320
	ATAGTGGTAT AACGTAATGA GTAGACACAA TATATAGGAA GAAGGGGTAA AATGAATCAA	4380
15	ATCGAAGAAG CATTAAACGGG TTTGATTTCT AAAGATCCTG CTATTGTTAA CGAAAATGCT	4440
	AACAAAGATA GTGATACATT TTCAACAATG AGAGATTTAA CAGCAGGTAT CGTTTCTAAA	4500
20	TCTTACGCAT TAAATCATTT ATTACCAAAG CACGTTGCAG ATGCACATCA AAGAGGGGAC	4560
	ATACATTTTC ACGACTTAGA TTATCATCCA TTCCAACCGT TAACTAACTG TTGTTTAATA	4620
	GATGCTAAAA ATATGCTACA TAATGGATT T GAAATAGGCA ACGCGAATGT AACTTCACCA	4680
25	AAATCAATAC AAATGTCATC AGCGCAGCTT GTACAAATTA TAGCCAATGT TTCTAGCAGT	4740
	CAATATGGTG GCTGTACGGT TGACCGCGTT GACGAATTAC tTAGTACATA TGCACGACcA	4800
	TAATGAAGAA CAACATAGGA ATATsCGCAA AGCAATTTGT CAAAGAATCT GAAATTGATC	4860
30	GTTATGTTGA TCAACAAGTC ACTAAAGACA TCAATGATGC GATTGAAAGT TTAGAATATG	4920
	AAATTAATAC CTTATATACA TCTAATGGAC AGACACCTTT TGTAACATTA GGATTCGGCT	4980
	TAGGTACAGA TCATTTAAGT CGCAAAATTC AACAAGCTAT CTTAAATACT CGTATCAAAG	5040
35	GCTTAGGAAA AGACCGCAGC ACAGCGATTT TCCCGAACT JGTATTTTCA ATTAAAAAAG	5100
	GAACCAACTT TAGTCCGCAA GATCCGAACT ATGACATTAA ACAACTAGCA TTAAAGTGTT	5160
40	CAACGAAACG TATGTATCCA GATATTTTAA ATTATGACAA ACTCGTAGAA ATATTAGGTG	5220
	ATTTCAAAGC GCCAATGGGT TGTCGTTTCA TTTTACCAAG TTGGAAAGAT GCGGAAGGTC	5280
	ATTTTGAAAA TAATGGTCGT TGTAATCTTG GTGTGTTTAC ACTTAATTTA CCTAGAATGG	5340
45	CATTAGAATC TGCCGGTAAT ATGACGAAAT TCTGGGAAAT CTTTATGAA CGTATCGATG	5400
	TGTTACATGA TGCATTACTT TATCGTATAA ATCGTTTGAA AGATGCTGTA CCGAATAACG	5460
	CACCGATTTT ATATAAAAGT GGCGCATTTA ACTATAAATT AAAAGAAACA GATGATGTTG	5520
50	CTGAGTTATT TAAAAATAAA CGTGCAACGA TTTCAATGGG CTATATAGGG TTGTATGAAA	5580
	CAGCTACTGT TTTCTATGGT CCAGACTGGG AACATCTCA AGAAGCAAAA GCATTTACGC	5640

55

	GGTTCAGTAT TTmCAGTACG CCGAGTGAAT CGCTACGGAT CGTTTTTGTC GTTTAGACCA	5760
	AGAGAGATTT GGAGATATTA AAGACATTAC AGATAAAGGA TATTATCAAA ACTCTTTCCA	5820
5	TTATGATGTA CGTAAAGATG TTACACCTTT TGAAAAGTTA GATTTTGAAA AAGATTATCC	5880
	TTATTATGCG AGTGGTGGTT TCATTCACTA TTGTGAGTAT CCGAAATTGC AACACAATTT	5940
10	GAAAGCACTA GAAGCGGTAT GGGACTACTC TTATGACAAA GTTGGTTACT TAGGTACAAA	6000
	TATTCCGATT GATCATTGTT ATGAATGTGA TTACGATGGA GATTTTGAAG CAACTGAAAA	6060
	AGGATTTAAA TGCCCGAACT GTGGCAATGA TAATCCTAAA ACAGTTGATG TCGTTAAACG	6120
15	AACATGTGGT TACCTAGGCA ATCCAGTTCA ACGTCCAGTA ATTAAAGGCC GTCATAAAGA	6180
	AATTGCGCA CGAGTAAAAC ATATGAAAGC GCCTAAAGAA TGATACTTTT AGACATTAAA	6240
	CAAGGACAAG GTTATATTGC TAAAATAGAA TCAAATAGCT TTGTTGACGG TGAAGGAGTA	6300
20	AGATGCAGTG TTTATGTATC AGGATGTCCA TTTAATTGTG TTGGATGTTA TAACAAAGCC	6360
	TCACAAAAGT TCAGATATGG CGAGAAATAC ACTGATGAAA TATTAGCAGA AATATTAGAT	6420
	GATTGCGATC ATGATTATAT ATCTGGGCTA AGTCTATTAG GTGGCGAACC ATTTTGTAAT	6480
25	TTGGATATTA CATTAAATCT TGTCAAAGCA TTTGAGCAC GTTTTGAAA TACAAAGACA	6540
	ATTTGGGTAT GGACTGGATT TTTATATGAA TATTTAGCAA ATGATTGTAC AGAACGTCGA	6600
	GAGTTATTAT CATAATTGA CGTTTTAGTA GATGGTCTAT TTATACAACA CTTATTCAAA	6660
30	CCTGATTTAC CATATAAAGG TTCTTTAAAT CAACGCATTA TAGATGTACA ACAATCACTC	6720
	TCGCATGCGC GTATGATTGA ATATATAGTT AGTTGAATAT GTATTAGAAG TCAAGGTAAC	6780
35	ATTCGTTGCC TTGGCTTCTT TTTAGGTTAG GTACATAATT GAAAGTTAAT AAAAGCAATT	6840
	CTTTATAAAA ATATATTGAT AGAATATGAC CTAACAATCA TTTTGATACC AATACTAAAA	6900
	GTTGCAATC CGTTTTTTAA AAAAGTTGAA AGAGAAAAGT GGTATTTTAG TGGGAAGGAA	6960
40	GTCIAACTTT TTGGTAGCGT TTTACAATAA ATAAATATTC GTTAATAACG TATAAATATT	7020
	CTTAAATGCC ATTCTAGTAA AATTTGTTAA ATTCGTTAAA TCGTAACTTA AACTGTAT	7080
	TTTAGCGCTA TTAAGGTTTT GTTTATTACG GGAAAATTA TATAAATATT CAATAATTGC	7140
45	CAAGTTTCAA ATTGTATGAA ATTTGCATTA TTATTAAATG TTAGTTATTG TCAATTTTGT	7200
	GAATCAATAT AATTATTACA TTTTGAGATA AATCGAAACA GGATTCATAA AATTAATAAT	7260
	TAGGGGGAGC ACAATTGAAA AAAGAGAAAG TTATGGACTG GACGACCTTT ATAGGGACAG	7320
50	TAGCTGTACT TCTTTTTGCA GTTATACCTA TGATGGCTTT TCCAAAAGCA AGTGAAGATA	7380
	TCATCACTGG TATTAATAGT GCCATTTCTG ATTCAATTGG TTCGATATAT TTATTTATGG	7440

55

	TTGGTAAAGC AAGTGATAAA CCAGAATTTA ATACATTTAC ATGGGCGGCA ATGCTGTTTT	7560
	GTGCAGGCAT AGGCTCTGAT ATTTTATACT GGGGCGTTAT TGAATGGGCT TTTTACTATC	7620
5	AAGTTCCACC AAATGGCGCG AAAAGTATGA GTGATGAAGC ACTCCAATAT GCGACGCAAT	7680
	ATGGTATGTT CCACTGGGGG CCAATTGCTT GGGCTATTTA TGTTCCTACCA GCATTACCAA	7740
10	TTGGTTATTT AGTATTTGTT AAAAAACAAC CGGTGTATAA AATTAGTCAA GCTTGTCGTC	7800
	CGATTTTAAA AGGTCAAACA GATAAATTG TAGGTAAAGT TGTAGATATC TTATTTATCT	7860
	TTGGATTGCT AGGTGGTGCG GCAACATCAC TAGCGTTAGG TGTGCCATTA ATTTCTGCAG	7920
15	GCATAGAAAG ATTAAGTGGT TTAGATGGTA AAAATATGAT TTTACGTTCC GCCATTTTAT	7980
	TAACAATCAC GGTTATATTT GCCATTAGTT CATATACAGG ATTGAAAAAA GGTATTCAAA	8040
	AGTTAAGTGA TATCAACGTT TGGCTATCCT TTGTACTTTT AGCCTTTATA TTTATTATTG	8100
20	GACCGACTGT TTTTATTATG GAAACGACAG TGACAGGGTT CGGAAATATG TTGAGAGATT	8160
	TCTTTCATAT GGCAACATGG TTAGAACCAT TCGGTGGTAT TAAAGGTCGA AAAGAAACGA	8220
	ATTTCCACA AGACTGGACA ATATTCTACT GGTGATGGTG GTTAGTATAT GCGCCATTTA	8280
25	TCGGTTTATT TATCGCTAGA ATTTCAAAG GTCGACGCCT TAAAGAAGTC GTGCTAGGAA	8340
	CAATTATTTA TGGAACGCTT GGATGCGTAT TATTCCTTGG TATTTTGGT AACTATGCTG	8400
30	TGTATTTACA AATTTCTGGA CAGTTTAATG TAACACAATA TTTAAATACA CATGGTACAG	8460
	AGGCAACCAT TATTGAAGTG GTGCATCATT TACCATTCCC ATCATTGATG ATTGTACTAT	8520
	TCTTAGTATC TGCTTTCCTA TTCTTAGCAA CAACATTGA TTCGGGTTCA TATATTTTAG	8580
35	CGGCAGCATC TCAGAAAAAA GTGGTAGGCG AACCATTACG TGCCAATCGT TTATCTGGG	8640
	CATTTGCATT GTGCTTATTG CCATTTTCAT TGATGCTAGT TGGTGGTGAA CGTGCATTAG	8700
	AAGTATTGAA AACTGCTTCA AACTGGCAA GTGTGCCATT AATTGTTATT TTTATTTTCA	8760
40	TGATGATATC ATTTTAAATC ATTTTAGGGC GCGATAGAAT TAACTTGAA ACGCGTGCTG	8820
	AAAAATTAAA AGAAGTTGAA CGTCGTTTAT TGCGAATCGT TCAAGTATCA GAAGAAGAAC	8880
	AAGACGATAA TTTATAATTC AAAGCGGGTC TGGGACGACG AAATGaaTTT TGTGAAAATA	8940
45	TCATTTCTGT TCCaTTCCCC TTTTPTTAGT AGCATTGTAG GATGAACTTT TAGGTTTTCA	9000
	TTAATGTTGT ACTAAAAGAT TTAATTTTTT AGTGCTCCAA GTACTTATTT ATTGTATGAA	9060
50	GCATATTCTA AATCGAAGTT TGAAAGACTC TCATTGATTA TTAAATTAAA TAAAGGGTAT	9120
	GCGTATGTAC AATTCAAATT AATCGAAGGA TGAAATAAAA TGACTAATCA ATTTAAAAAT	9180
	AAACAGTCCA AATTACATGA CAGTTTAGAA TCCATCACAA AAACTTATA TGCGACACCT	9240

55

ACAGAATATT GTTATCTATC ATTCCGGACA CTTAGGTGAC TCCCAACAAG ACATTGCATC 9360
 ATTAGGTGGT GTTTCAAAAG TATTGATGAA TCATGATCAT GAATCTATAG GAGGTTCTAA 9420
 5 TCAAGTTGAA GCCCCTTACT TTATACATGA AAATGATGTG GCTGCACTGA AACATAAGAT 9480
 TTCTGTTCAA AAACAATTTA GTAATCGTGT AATGTTGGAT AAGGATTTAG AAGTTATTCC 9540
 CGCGCCTGGA CATAACCAG GGACGACACT ATTTTATG GATGATGGTC ATCACCCTTA 9600
 10 CTTATTTACT GGAGATTTTA TATGTTTGA AGGGAAGAGA TGGCGTACAG TTATATTAGG 9660
 TTCAAGTGAT AGAGAAAAAT CTATTCAAAG TTTAGAGATG GTTAAAGAAT TAGATTTTGA 9720
 TGTACTTGTA CCTTGGGTTA CTATCAAAGA TGAACCGTTA GTTTATTTTG TAGAAAATGA 9780
 ATATGAAAAA CGTGAACAAA TACAAAATAT TATTGATAGA GTACGTGAGG GCGAGAATAG 9840
 CTAATTGAAA TATATTGGCG AAGCAATGTA ACGAATCTAA GAAAGCCCTA GAAAAACCT 9900
 20 CCATAATTGA TTGTCATATA AAACAAAAAC GGTAATTTCT ATTTATTGAG ATAGAAATTA 9960
 CCGTTTATTT CGTGGACCTA TTGCATTGTT TTTATCATGC ATAATCATCA TTGTCGTTGT 10020
 TTGAGTCAAT TTTAATTTTC AGAATCAGAA GGCTGTTCTG GAATTGGGAA ATATTGAAA 10080
 25 ATTTACCGC TTTCAATCGC TTCGGTTAAC TGTCTAACC ATTCGTAATA AACATGTGTA 10140
 TGATCAAGCT GAGCTTTAAT TTTTGTGCC TCTGTGTTT CAGCTTCAGT TAAATCACTG 10200
 CTTTCAAGTA ATGGATTGAT AATAGCTTGA GCATCTTTTA CTGCTTCGAC ATTGATGTCA 10260
 30 ATTTACCGCT GGAATTTTTT AGTGAAAAAG TTTCGGAAA AGATGAAAAA GTCTTTCTCG 10320
 GCGATAAAAT GTTGTGTGCG GCTTCCTCTC GTAAATTGTT GTTTAACAAT ATCAAATTC 10380
 TGCAATTTCT TAACGCCAGC ACTCATACTT GGTTTGCTCA TTTGCAATTG ATGACGCATT 10440
 35 TCATCAAGCG TCATACTGCC TTCAAACACC ATTGTGCCAT ATAAGTTTCC TACACTTCTA 10500
 TTAGTGCCAT ACAAATCCAT TGTCTGTCCA ATTGAATTAA TTACAATATC TTTTGCTTGT 10560
 TCTAATTGTT GCTGTTTGT CTGAGAACGA GTCATCATTG CACCTCCGTA CATCATTTTG 10620
 40 GTCACGTAA AATAAATACT AATACATTAT AAAACCTTTT CTAAAAAAG ACATTAAAAA 10680
 TATTTAAAGC ATTAAAGTTA AATGTTTCGT TAAATAAAAA TCTAACGAAC TTACAAAAC 10740
 45 TAATCTTGA GTTGTTTTGT AAATTGACAC ATTTTTCATT TCTATGCTAA CATAAGTnTG 10800
 TAAATTCGT TAAATAAAAA TTTAACAAAC TTAACGGrGG TTGTTGAakG GrACTTTTAA 10860
 aACATTTATC TCAGCGTCAA TATATTGATG GTGAGTGGGT TGAAAGCGCG AATAAAAATA 10920
 50 CAAGAGATAT TATCAATCCT TACAATCAAG AAGTGATATT TACGGTTTCT GAAGGGACAA 10980
 AAGAGGATGC AGAACGTGCA ATCTTAGCTG CAAGACGTGC GTTTGAGTCT GGTGAATGGT 11040

55

AACATCgCGA AgCgTTAGCA CGATTAGAAA CATTAGATAC TGGAAAAACG TTAGAAGAAT 11160
 CATATGCAGA TATGGATGAT ATTCATAATG TGTtTATGTA TTTTGCTGGA TTAGCAGATA 11220
 5 AAGACGGTGG CGAAATGATT GATTCACCAA TTCCAGATAC AGAAAGCAAA ATTGTtTAAAG 11280
 AACCAGTAGG TGTAGTTACA CAAATTACAC CTTGGAATTA TCCGTTATTA CAAGCATCAT 11340
 10 GGAAAAATTGC GCCAGCGCTT GCTACGGGTT GTTCACTAGT TATGAAACCA AGTGAAATTA 11400
 CACCATTAAAC AACAATACGT GTTTTTGAAT TAATGGAAGA AGTTGGTTTC CCTAAAGGAA 11460
 CAATTAATCT TATTCTAGGT GCAGGTCTG AAGTTGGTGA CGTAATGTCA GGTcATAAAG 11520
 15 AGGTTGACCT TGTATCATT ACAGGTGGCA TTGAGACTGG TAAGCATATT ATGAAAAATG 11580
 CTGCTAATAA TGTTACGAAT ATTGCCTTGG AACTTGCGCG TAAAAATCCA AACATTATCT 11640
 TTGATGATGC TGATTTTGAA TTGGCAGTAG ACCAAGCGTT AAATGGTGA TATTTCCATG 11700
 20 CAGGTCAAGT TTGTTcAGCA GGATCAAGAA TATTAGTACA AAACAGTATT AAAGACAAAT 11760
 TTGAGCAAGC ACTTATTGAT CGCGTGAAA AATCAAATT AGGTAATGGT TTTGATGCTG 11820
 ATACTGAAAT GGGACCAGTG ATTTCAACAG AACATCGTAA TAAGATCGAA TCTTATATGG 11880
 25 ATGTAGcTAA AGCAGAAGGC GCAACAATTG CTGTTGGTGG TAAACGTCCA GATAGAGATG 11940
 ATTTAAAAGA TGGTCTATTC TTCGAGCCAA CAGTCATTAC AAATTGTGAT ACGTCAATGC 12000
 GTATTGTACA AGAAGAGGTT TTCGACCTG TCGTACTGT AGAAGGCTTT GAAACTGAAC 12060
 30 AAGAAGCGAT TCAATTAGCG AATGATTCTA TATATGGTTT AGCAGGTGCT GTATTTTCTA 12120
 AAGATATTGG AAAAGCACAA CGCGTTGCTA ACAAGTTGAA ACTTGAACG GTGTGGATTA 12180
 ATGATTTCCA TCCATATTTT GCACAAGCGC CATGGGGTGG ATACAAACAA TCAGGTATCG 12240
 35 GTAGAGAATT AGGCAAAGAA GGCTTAGAAG AGTACCTTGT TTCAAACAC ATTTTAACAA 12300
 ATACAAATCC ACAATTAGTG AATTGGTTTA GCAAATAAAA ATTAGATAAG GTGAGTGCCA 12360
 40 TTGTAAGAAC ACAAGACACT CACTTTGTTT TGTATAAGTG GCGAAATGTT GATTGATAAT 12420
 TTGGACTAAA CGCAAAATGA ATCATAGATT ATTTcATTAC TGTTAGTAAC AATCGTAAAA 12480
 GGAAAAGCGA GTGTTTTGGT TAGCTAAGTT TAGCAATTCA ACGATAACCA ATCAGCCACT 12540
 45 AACAAATATT TCATGCAATA CTCACTTTGA AATACAACAA ACTTTGGAGG TCATAACGAT 12600
 GAGTAACAAA AACAAATCAT ATGATTATGT CATCATTGGA GGAGGCAGTG CAGTTCTGT 12660
 ACTAGGTAAT CGTCTGAGTG AAGATAAAGA TAAAGAAGTC TTAGTATTAG AAGCGGGTCG 12720
 50 CAGTGATTAT TTTTGGGATT TATTTATCCA AATGCCTGCT GCGTTAATGT TCCCTTCAGG 12780
 CAATAAATTT TACGATTGGA TTTATTCAAC AGATGAAGAA CCACATATGG GCGGTCGTAA 12840

55

TCAACGTGGT AATCCAATGG ACTATGAAGG CTGGGCAGAA CCAGAAGGTA TGGAAACTTG 12960
 GGATTTTGGC CACTGTTTAC CGTATTTTAA AAAATTAGAA AAAACATACG GTGCAGCGCC 13020
 5 TTATGATAAA TTTAGAGGCC ATGATGGACC AATTAAGTTA AAACGAGGGC CAGCAACGAA 13080
 TCCTTTATTC CAGTCATTCT TTGATGCAGG TGTGAAGCA GGCTATCATA AAACACCTGA 13140
 TGTGAATGGA TTTAGACAAG AAGGTTTTGG ACCGTTTCAT AGTCAAGTAC ATCGTGGTCG 13200
 10 CCGAATGTCA GCTTCAAGAG CATATTTACA TCCAGCGATG AAGCGTAAAA ACTTAACCGT 13260
 TGAAACACGT GCCTTTGTAA CTGAAATTCA TTATGAAGGT AGAAGAGCAA CTGGTGTTC 13320
 GTATAAGAAA AATGGCAAAC TACATACCAT CGATGCTAAT GAAGTCATTT TGTCTGGTGG 13380
 15 GGCATTCAAT ACGCCACAAT TACTACAATT ATCTGGTATC GGTGATTGAG AGTTCCTAAA 13440
 ATCAAAAGGC ATTGAGCCAC GTGTTTATTT ACCTGGTGTG GGTGAAAACCT TTGAAGATCA 13500
 20 CTTAGAGG 13508

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7646 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

GTAAGTATTG TCTTGATTTC CTAATAAAGT TATATCTTGT AATTCATCTT GTTGACGGCC 60
 ATGTGCCATA TAAAGCGCTC CTTTAAATTT ATTTTTTTTAT TATTTTGGCG TCTCGGCGTG 120
 35 CTTTTTCAAA CATGTAATAA CTTGCACCGA TAATAACGAC GTAACCTAAT GTTGATAGA 180
 AATCTGGAGA TTCTCCGAAT AGAATAAATC CAAGTATTGC TGTGAAAATT ATAGATGCAT 240
 40 ACGTAAAAAT AGAAATATCT TTTGCTGCTG CAAACTATA TGCTAAAGTA ACACCAATTT 300
 GACCCACAGC GGCAGCTAAG CCAGCCCCTA ATAGATAAAG TATTTGCATC TGAATCATTC 360
 GTTCATAAGT ATATGCAGTG AAAGGTATTA AAACGATGAC AGAAAATAAG GAGAAGTAAA 420
 45 ATACTATAGT ATATGGTGCT TTTCTTGCTG TAAGTGCTCG AACACATGTA TATGCTGATG 480
 CTGCAAAAAT ACCTGAGAAT AAGCCAGCTA ATGATGGAAT CATAGATGAT GAAAATTCAG 540
 GTTTCACAT TAAAGCAaC CTAAAATAGC AATTATCATT GCTGTAATTT GaTACTTCCT 600
 50 TACCTTTTCA TGtAAGaaa CAATGCTTaA TAAAATCGTC CAGAAAGGAT TGAGTTTCAT 660
 TAATGAATCG GCATCACTAA GTACCATATG ATCAATGGCA TAAATATTTA ACAATACACC 720

	TGGCTGATGG TATTTATATA TAAAAAATAA TGGAATAAAC ATTGCTACTA AGTTTCGTGC	840
	TAATGATTTT TGAAAAACAG GAAGGTCACC TGCAAGTCTG AAAAACACTG ACATAAAACT	900
5	GAAACCAATA GCCGAAATTA AAATGGCAAT GATACCTTTT ACTTTAGGAT TCAATTTTAT	960
	CGCCTCTTTT ATATAAAATT AACGTATTTA TATTAGCATA AAACAACATG TTGTGCATAA	1020
10	ATAGTTGAAA TTTACTATAA AAAGACTATA ATAGACTGTA GCGAACAAAC GTTCTGTGTT	1080
	TATTTGTCGG AATAATAGGG CATTACACTT TTATGAATGT TTGTGTTATT ACATAAAACA	1140
	AATATCAATT CAGTATCAAG CTAATAAGCT TTTTCTTGAT TTCTGTTGAT ACAATTGAGA	1200
15	TTGACACAGA TTTAAAAAAA TCAAGTGATA TCTACTAAAA AATTTTTTTA AATTTGTTCA	1260
	AGTTTTTCTA ATTTAGTATT GGTGCCTAGT TGGAACGTTT TACGAACATT CGATTAGAAA	1320
	ATGGCACTTT AAATCATAGT GTGTCTTATG TATAATGAAA CACATAATAT AGTGTGGTG	1380
20	AAACGAAAAA gACACAATAT CTTGTGTTTT STATGCAAAT GCTTTATTTA TGAAGAAATT	1440
	ACATTTAAAA GTAATTTAAC ACAGAAATTT AATAGTTATT ATCAATTAAT AGTCATATTT	1500
	TTAGAAAATG TACTGAGCAA ATGGAAGATA TCCAATGATG TAAACACTAC ATATAGTGAT	1560
25	TTTTATACAT TCAACCCATA TAAGCTACTA TTTTCTCAA TATAAATCTA TGCAATTGGT	1620
	TTACATTTGA GAAAATAAGT AGCTTCATTA TAGTTAATAC AATGCTGAGA TAACCATAGT	1680
	AACCATGTTG TTAAAGCATT TTTTAATTGG AATGACTACT TTATTTAAAA GGGTTGAAGA	1740
30	AAGAAGGTGA TCCAATGAAA ATAATATATT TTTTCATTAC TGGAAATGTC CGTCGTTTTA	1800
	TTAAGAGAAC AGAACTTGAA AATACGCTTG AGATTACAGC AGAAAATTGT ATGGAACCAG	1860
35	TTCATGAACC GTTTATTATC GTTACTGGCA CTATTGGATT TGGAGAAGTA CCAGAACCCG	1920
	TTCAATCTTT TTTAGAAAGT AATCATCAAT ACATCAGAGG TGTGGCAGCT AGCGGTAATC	1980
	GAAATTGGGG ACTAAATTTT GCAAAAGCGG GTCGCACGAT ATCAGAAGAG TATAATGTCC	2040
40	CTTTATTAAT GAAGTTTGAG TTACATGGAA AAAACAAAGA CGTTATTGAA TTTAAGAACA	2100
	AGGTGGGTAA TTTAATGAA AACCATGGAA GAGAAAAAGT ACAATCATAT TGAATTAAAT	2160
	AATGAGGTCA CTAAACGAaG AGAAGATGGA TTCTTTAGTT TAGAAAAAGA CCAAGAAGCT	2220
45	TTAGTAGCTT ATTTAGAAGA AGTAAAAGAC AAAACAACTT TCTTCGACAC TGAATCGAG	2280
	CGTTTACGTT ATTTAGTAGA CAACGATTTT TATTTCAATG TGTTTGATAT TTATAGTGAA	2340
	GCGGATCTAA TTGAAATCAC TGATTATGCA AAATCAATCC CGTTTAATTT TGCAAGTTAT	2400
50	ATGTCAGCTA GTAAATTTTT CAAAGATTAC GCTTTGAAAA CAAATGATAA AAGTCAATAC	2460
	TTAGAAGACT ATAATCAACA CGTTGCCATT GTTGCTTTAT ACCTAGCAAA TGGTAATAAA	2520

55

	ACATTTTAA	ACGCAGGCCG	TGCGCGTCGT	GGTGAGCTAG	TGTCATGTTT	CTTATTAGAA	2640
	GTGGATGACA	GCTTAAATTC	AATTAACTTT	ATTGATTCAA	CTGCAAAACA	ATTAAGTAAA	2700
5	ATTGGGGGCG	GCGTTGCAAT	TAACCTATCT	AAATTGCGTG	CACGTGGTGA	AGCAATTAAA	2760
	GGAATTAAAG	GCGTAGCGAA	AGGCGTTTTA	CCTATTGCTA	AGTCACTTGA	AGGTGGCTTT	2820
10	AGCTATGCAG	ATCAACTTGG	TCAACGCCCT	GGTGCTGGTG	CTGTGTACTT	AAATATCTTC	2880
	CATTATGATG	TAGAAGAATT	TTTAGATACT	AAAAAAGTAA	ATGCGGATGA	AGATTTACGT	2940
	TTATCTACAA	TATCAACTGG	TTTAATTGTT	CCATCTAAAT	TCTTCGATTT	AGCTAAAGAA	3000
15	GGTAAGGACT	TTTATATGTT	TGCACCTCAT	ACAGTTAAAG	AAGAATATGG	TGTGACATTA	3060
	GACGATATCG	ATTTAGAAAA	ATATTATGAT	GACATGGTTG	CAAACCCAAA	TGTTGAGAAA	3120
	AAGAAAAAGA	ATGCGCGTGA	AATGTTGAAT	TTAATTGCGC	AAACACAATT	ACAATCAGGT	3180
20	TATCCATATT	TAATGTTTAA	AGATAATGCT	AACAGAGTGC	ATCCGAATTC	AAACATTGGA	3240
	CAAATTAAAA	TGAGTAACTT	ATGTACGGAA	ATTTTCCAAC	TACAAGAAAC	TTCAATTATT	3300
	AATGACTATG	GTATTGAAGA	CGAAATTAAA	CGTGATATTT	CTTGTAAGTT	GGGCTCATTA	3360
25	AATATTGTGA	ATGTAATGGA	AAGCGGAAAA	TTCAGAGATT	CAGTTCACTC	TGGTATGGAC	3420
	GCATTAAGTG	TTGTGAGTGA	TGTAGCAAAT	ATTCAAATG	CACCAGGAGT	TAGAAAAGCT	3480
	AACAGTGAAT	TACATTCAGT	TGGTCTTGGT	GTGATGAATT	TACACGGTTA	CCTAGCAAAA	3540
30	AATAAAATTG	GTTATGAGTC	AGAAGAAGCA	AAAGATTTTG	CAAATATCTT	CTTTATGATG	3600
	ATGAATTTCT	ACTCAATCGA	ACGTTCAATG	GAAATCGCTA	AAGAGCGTGG	TATCAAATAT	3660
35	CAAGACTTTG	AAAAGTCTGA	TTATGCTAAT	GGCAAATATT	TCGAGTTCTA	TACAACTCAA	3720
	GAATTTGAAC	CTCAATTCGA	AAAAGTACGT	GAATTATTCG	ATGGTATGGC	TATTCCTACT	3780
	TCTGAGGATT	GGAAGAAACT	ACAACAAGAT	GTTGAACAAT	ATGGTTTATA	TCATGCATAT	3840
40	AGATTAGCAA	TTGCTCCAAC	ACAAAGTATT	TCTTATGTTT	AAAATGCAAC	AAGTTCTGTA	3900
	ATGCCAATCG	TTGACCAAAT	TGAACGTCGT	ACTTATGGTA	ATGCGGAAAC	ATTTTACCCT	3960
	ATGCCATTCT	TATCACCACA	AACAATGTGG	TACTACAAAT	CAGCATTCAA	TACTGATCAG	4020
45	ATGAAATTAA	TCGATTTAAT	TGCGACAATT	CAAACGCATA	TTGACCAAGG	TATCTCAACG	4080
	ATCCTTTATG	TTAATTCTGA	AATTTCTACA	CGTGAGTTAG	CAAGATTATA	TGTATATGCG	4140
	CACTATAAAG	GATTAAATC	ACTTTACTAT	ACTAGAAATA	AATTATTAAG	TGTAGAAGAA	4200
50	TGTACAAGTT	GTTCTATCTA	ACAATTAAAT	GTTGAAAATG	ACAAACAGCT	AATCATCTGG	4260
	TCTGAATTAG	CAGATGATTA	GACTGCTATG	TCTGTATTTG	TCAATTATTG	AGTAACATTA	4320
55							

	ATGTTTTGGA GACAAAATAT ATCTCAAATG TGGGTTGAAA CAGAATTTAA AGTATCAAAA	4440
	GACATTGCAA GTTGAAGAC TTTATCTGAA GCTGAACAAG ACACATTTAA AAAAGCATTAA	4500
5	GCTGGTTTTAA CAGGCTTAGA TACACATCAA GCAGATGATG GCATGCCTTT AGTTATGCTA	4560
	CATACGACTG ACTTAAGGAA AAAAGCAGTT TATTCATTTA TGGCGATGAT GGAGCAAATA	4620
10	CACGCGAAAA GCTATTCACA TATTTTCACA ACACTATTAC CATCTAGTGA AaCAAACCTAC	4680
	CTATTAGATG AATGGGTTTT AGAGGAACCC CATTTAAAAT ATAAATCTGA TAAAATTGTT	4740
	GCTAATTATC ACAAACCTTG GGGTAAAGAA GCTTCGATAT ACGACCAATA TATGGCCAGA	4800
15	GTTACGAGTG TATTTTTAGA AACATTCTTA TTCTTCTCAG GTTTCTATTA TCCACTATAT	4860
	CTTGCTGGTC AAGGGAAAAT GACGACATCA GGTGAAATCA TTCGTAAAAT TCTTTTAGAT	4920
	GAATCTATTC ATGGTGTATT TACCGGTTTA GATGCACAGC ATTTACGAAA TGAACCTATCT	4980
20	GAAAGTGAGA AACAAAAAGC AGATCAAGAA ATGTATAAAT TGCTAAATGA CTGTATTTA	5040
	AATGAAGAGT CATAACAAA AATGTTATAC GATGATCTTG GAATCACTGA AGATGTGCTA	5100
	AACTATGTTA AATATAATGG AAACAAAGCA CTTTCAAACCT TAGGCTTTGa ACCTTATTTT	5160
25	GAGGAACGTG AATTTAACCC AATCATTGAG AATGCCTTAG ATACAACAAC TAAAAACCAT	5220
	GACTTCTTCT CAGTAAAAGG TGATGGTTAT GTATTAGCAT TAAACGTAGA AGCATTACAA	5280
30	GATGATGACT TTGTATTTGA CAACAAATAA CAATTAAAT AAAAGACCTT CACATGTAAA	5340
	GGGAAATAGC GATTCGTTTC GTCTTGTCTC CTACATGTTG AAGGTCTTTT TTTATGTGTA	5400
	TCTAACTCAT TATGAGTCTG AGTAAGAAAT CAATGCTCTA AGATGTACAA TGCTATTTAT	5460
35	ATTGGCAGTA GTTGGCGGGG CCCCAACACA GAAGCAGGCG GAAAGTCAGC TAACAATATT	5520
	GTGCAAGTTG GCGGGGCCCC AACATAGAAG CAGGCGGAAA GTCAGCTAAC AATAATGTGC	5580
	AAGTGGCGG GGGCCCAACA TAAAAGCAGG CGGAAAGTCA GCTAACAATA TTGTGCAAGT	5640
40	TCGGGCGGGG CCCCAACATA AAGAAAAACT TTTTCCTTTA GAAATTATCA CTTCCaCaTG	5700
	AGTTTTACTC ATGTATTCCT ATTTTAAAGT ACACATTAGC TGAGGCTAAT GTTAAGAACC	5760
	ACTACTTAAT CAATCATTAG TAGTTTTTAT CATTTCCACT ATTCCCaGAC ATCaAAATCT	5820
45	TAAGTGTTCT ATTTTACTTT AAGTAAACAA AATACACATT CCGAAAAATT AAATTTCACT	5880
	TTAATTGCAA ATATCAATAA AATTGACACT AAATTATTTG AAAGGCTATT GAAATTATGG	5940
50	TCAAAAAACG CTACTATTAA TGAGAAATAT TATCAATGAT AATGATTATC ATTAATTTAA	6000
	AGGGAGAAAA ATTTGTAATG AAGTATTTAT TAAAGGGAAA TATTTTGCTT CTATTACTAA	6060
55	TATTGTTGAC AATTATTTCTG TTGTTCATAG GTGTGAGTGA ACTATCAATT AAAGATTTAC	6120

	GTATTTTAAT TGCTGGAAGT TCGTTGGCTT TAGCAGGCTT GATAATGCAA CAAATGATGC	6240
	AAAATAAGTT TGTTAGTCCG ACTACAGCTG GAACGATGGA ATGGGCTAAA CTAGGTATTT	6300
5	TAATTGCTTT ATGTTCTTT CCAACCGTC ATATTTTATT AAAACTAGTA TTTGCTGTTA	6360
	TTTGCACTAT TTGCGGTACG TTTTATTG TTTAAATCAT TGATTTTATA AAAGTGAAAG	6420
10	ATGTCATTTT TGTACCGCTT TTAGGAATTA TGATGGGTGG GATTGTTGCA AGTTCACAAC	6480
	CTTCATCTCA TTGCGCACGA ATGCTGTTCA AAGCATTGGT AACTGGCTTA ACGGGAACCT	6540
	TGCCATTATC ACAAGTGGAC GCTATGAAAT TTTATATTTA AGTATTCCTC TTTTAGCATT	6600
15	GACATATCTT TTTGCTAATC ATTTACAGAT TGTAGGAATG GGTAAAGACT TTTACTAATAA	6660
	TTTAGGTTTG AGTTACGAAA AATTAATTAA CATCGCATTG TTTATTACTG CAACTATTAC	6720
	AGCATTGGTA GTGGTGA CTG TTGGAACATT ACCGTTCTTA GGACTAGTAA TACCAATAT	6780
20	TATTTCAATT TATCGAGGTG ATCATTGTA AAATGCTATC CCTCATACGA TGATGTTAGG	6840
	TGCCATCTTT GTATTATTTT CTGATATAGT TGGCAGAATT GTTGTTTATC CATATGAAAT	6900
	AAATATTGGT TTAACAATAG GTGTATTTGG AACAATCATT TTCCTTATCT TGCTTATGAA	6960
25	AGGTAGGAAA AATTATGCGC AACAATAATA AAAAAATAAT GCTTTTAATT GCAGTAACGT	7020
	TATTAATTAG TATGCTGTAC TTATTTGTAG GTATTGATTT TGAAATATTT GAATATCAAT	7080
	TTTCAAGTCG TTTAAGAAAG TTCATATTAA TTATTTTAGT AGGTGCTGCC ATTGCAACCT	7140
30	CAGTGGTGAT TTTCAAGCG ATTACAAATA ACCGTCTATT GACACCATCA ATAATGGGGT	7200
	TAGATGCAGT TTATTTATTT ATCAAAGTAT TGCCAGTCTT TTTATTTGGA ATTCAATCGG	7260
35	TATGGGTTAC TAATGTATAT TTGAACCTTA TATTAACACT TATAACGATG GTGTTATTCG	7320
	CACTAATCCT ATTCCAAGGT ATCTTTAAAA TCGGACATTT TTCAATTTAT TTTATCTTAC	7380
	TTATTGGTGT CCTTTTAGGA ACATTTTTTA GAAGCATAAC AGGTTTTTATT CAACTGATTA	7440
40	TGGATCCTGA GTCATTTTTA GCAATACAAA GTAGTATGTT TGCTAATTTT AATGCTTCTA	7500
	ATTGGAATTT AGTTACTTTC TCAGCAGTGC TATTAGTAAT CTTATTAGTC ATTACAATTT	7560
	TACTATTGCC TTATTTAGAT GTATTGCTTT TAGGTCGTGC TGAAGCAATT AATCTTGGGA	7620
45	TATCGTATGA AAAATTAACG CGAATT	7646

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

5 ATGAATATAT TnnAAATAA ATTATTATGG ATTGCACCAA TnGCCACTAT GATTATCTTG 60
 GTAATCTTTT CTTTAGCTTT TTATCCTGCA TATAATCCTA AACCAAAAGA TTTACCAATT 120
 GGTATATTAA ACGAGGATAA AGGTACAACG ATTCAAGATA AAAATGTTAA CATTGGTAAA 180
 10 AAATTAGAGG ATAAATTATT AGATAGTGAT TCTAATAAAA TTAAATGGGT TAAGGTTGAT 240
 AGTGAAAAAG ACCTTGAAAA AGATTTGAAA GATCAAAAAA TCTTTGGAGT AGCTATTATT 300
 GATAAAGACT TTTCAAAAGA TGCTATGAGT AAAACACAAA AAGTAGTTAT GGATAGTAAA 360
 15 AAAGAAGAAA TGCAACAAAA AGTTGCTTCA GGTGAAATTC CGCCACAAGT GGTTCACAA 420
 ATGAAACAAA AAATGGGGAA TCAACAAGTA GAGGTTAAGC AGGCTAAATT TAAAACGATT 480
 GTAAGTGAAG GATCAAGCTT ACAAGGTTCA CAAATTGCAT CAGCTGTGTT AACTGGTATG 540
 20 GGTGATAATA TTAATGCTCA AATTACGAAG CAAAGTTTGG AAACATTAAAC GAGTCAAAAT 600
 GTTAAAGTCA ATGCCGCGGA CATCAATGGT TTGACGAATC CAGTAAAAGT GGATAATGAA 660
 AAACCTAATA AAGTTAAAGA TCACCAAGCA GGTGGTAATG CACCATTCTT AATGTTTATG 720
 25 CCAATTTGGA TAGGTTCAAT CGTAACGTCT ATCTTATTGT TCTTTGCATT TAGAACTAGT 780
 AACAAATATCG TCGTGCAACA TCGTATCATT GCTTCAATTG GACAGATGAT ATTTGCAGTT 840
 GTTGCAGCAT TTGCAGGTAG CTTTGTATTAT ATTTATTTCA TGCAAGGCGT TCAAAGATTT 900
 30 GATTTTGACC ATCCAAATCG TATCGCAATT TTTGTAGCAT TTGCGATTCT TGGTTTCGTG 960
 GGCCTTATTT TAGGTGTTAT GGTATGGCTA GGTATGAAGT CAGTTCCAAT TTTCTTCATT 1020
 TTAATGTTCT TTAGTATGCA ACTTGTAACG TTACCTAAAC AAATGTTGCC TGAAAGTTAT 1080
 35 CAAAAATATG TATATGATTG GAATCCATTC ACACACTATG CAACAAGTGT AAGAGACTAT 1140
 TATACTTGAA TCATCATATT GAATTAAATA GTACAATGTG GATGTTTATA GGGT 1194

(2) INFORMATION FOR SEQ ID NO: 123:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 558 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

50 GACCGACCTA TACATCCGTA TAAGTATTTT TTGATATAAG TCTTCTAAAT CATAATGATT 60
 AAATCCAAAT GTTTTGATGC GTCGAATAAT TAATGGTTGT AGATCCATTA CTAACCTTTTC 120

55

GTATTTCAAA TATTAACTA ACCCCTTCTA TCTAAAATTT AAGGTTAGTT TAATATTGTT 240
 ACATTCAAAA TTTCAAGATG ACGGAAATGT CATTTCTTAT GATGTCCTCT TCGTATTTTT 300
 5 TCAAATTCTG CAAGGATTC AGAAGATAAC GGAATTCGAG TTCTTGGCTT GTTTTCACTT 360
 ATATCATCTA ATGATTTACT CACATCAATT TCATTTTCTT TTAAATCTCT CCACATTTTG 420
 CGAGATGATA TTCTATATGC ACCTGATCCA AAGATAGCAT GTTGcTCACT CaTATCACTT 480
 10 GTTACAACTG TAATATGcTT AGtATGCTTG tCaTAAAGtT CaTAAACCAT AACGGTTCTA 540
 ATGGAAACCA ATCAGCTG 558

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7762 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

25 GCTTCAGACA TnTGATGATA TAATCTCTCA TCATCGATTA ATTCTTTTGC AGCTTGATAC 60
 ACATnTTGCT TATTTGTTCC AATGACTTTT AATGTGCCAG CTTCAACACC TTCAGGACGT 120
 TCTGTAACAC TTCGCCAAAA CTAAACTGG CTTATTAAAT GATGGCGCTT CTTCTGAAT 180
 30 TCCACCTGAA TCTGTCAAAA TAAATAAGA TTTnTAGCA AAATTATGGA AATCTATACG 240
 TCCAAAGGTT CAATCAATTC AATTCTGTCA TGA CTACCTA AAATCTTTTG AGCCACCTCT 300
 CGAACTTTTCG GGTTTTTATG CATTGGATAT ACCAGTGCTA AATCAGTATA CTCATCTATT 360
 35 AAGCGTCTAA CCGCTTTAAA TATAITTTTCC aTGGGTTTCC CGATATTTTC TCGTCGGTGT 420
 GCTGTCATrA GAATGAATTT KtGTCTATGG TATTTATCCA TGATGTTAGA TTTATAATTG 480
 TCATCAACTG TATATTTTCAT AGCATCAATC GCAGTATTAC CAGTGACAAC AACACTTTCT 540
 40 GAATATTTCC CTTCACTTAA CAAATGCGAT GCAGCATTTT TAGTAGGTGC AAAATGTAAG 600
 TCAGCTAATA CACCAACTAA TTGTCTATTC ACCTCTTCTG GAAAAGGTGA ATATTTATCA 660
 45 TAACTTCTAA GCCCTGCTTC AACGTGTCCA ATCGGCACTT GGTATAAAA TGCCGCTAAA 720
 CCACCTGCAA ATGTCGTCAT CGTATCACCA TGTACAAGTA CCATGTCTGG TTTTCTAAT 780
 TGAATCACTT GTTCTAATTG AGTGATTGAT TTAGAAGTIA TCTCAGAAAG TGTCTGTCCT 840
 50 GATTTTCATA TATTCAAATC GTATTTTGGT TTGATTTCAA AGGTACTTAA TACTGAATCA 900
 AGCATTTCTC TATGCTGTGC TGTAACAACA ACAATTGGCT CGAGCATTTT TTCTTGTTCC 960

	ATCTTTTTCA TCAAACACT TATCTCCGAT TCTTCTATTT AGTACCAAAC AATCTATCTC	1080
	CAGCGTCGCC TAACCTGGT GTGATATATG CTTTGTCAAT aGCTTTTCAT CAAGTGCAGC	1140
5	AATATAAATA TCTACATCTG GATGTGCTTC ATGCATCTTT TCTACGCCTT CTGGTGTCTG	1200
	AATTAAACAC ATGAAGCGAA TATTTTTAGC GCCACGTTTC TTCAATGAAG TAATAGCTTC	1260
	AATTGCTGAT GCGCCTGTTG CTAACATAGG ATCAACAACA ATGATTTGTC TTTCAGTAAT	1320
10	ATCTTGAGGT AACTTAGCAA AATACTCTAC AGCCTTTAAT GTTTCGGGAT CTCGATATAA	1380
	ACCGATATGT CCAACTCTGG CTGCAGGTAC TAAACTTAAA ATACCATCAG TCATACCTAA	1440
	ACCAGCTCTT AAAATTGGAA CGATAGCTAA TTTTTTACCA GCTAATCGTT TAGCCGTCAT	1500
15	TTTAGTTACA GCGGTTTCAA TATCAACATC CTGAAGCTCT AAGTCTCTAG TTACTTCATA	1560
	TGCCATCAAC ATACCAACTT CGTCTACAAG TTCTCTAAAT TCTTTAGTAC CTGTATTTAC	1620
20	ATCTCTAATA TAGCTTAGTT TGTGTTGAAT TAATGGATGA TCGAAAACGT GTACTTTACT	1680
	CATAAAATT ACTCCTATCT TTGTGTATGT TTATTGATAT AGAGGATATT CAGCTGTAA	1740
	TTTCGCAACG CGTCTTTAG CTGTGTGTA TTTTCTTCA TCTTTACTAT TTTTCAATGC	1800
25	TAAACTGATG ATTTTTGCAA CTTCCTCAAA AGCTTTTCA TCAAATCCAC GCGTTGTTGC	1860
	AGCAGGTGTA CCTAAACGTA TACCACTCGT TACAAAAGGT TTTTCTTGAT CGAACGGAAT	1920
	GGTATTTTTG TTACATGTGA TACCAACTGA ATCTAAAGTC TCTTCAGCTT CTTTACCAGT	1980
30	AAGTCTATA GACCCTTTTA CATCAACAGC TACTAAGTGA TTATCTGTAC CGCCAGAAAC	2040
	AATTCTAAAT CCTTCATTAA TTAATGCTTC TGCAAGAACT TTTGCGTTTT TAACCACTTG	2100
	TTGTTGATAC GTTTTGAAAT TATTTCTAA CGCTTCTCCA AAAGCAACTG CTTTtGCTgC	2160
35	AATAACATGC TCAAGAGGTC CACCTGAAT ACCAGGGAAA ATTGTTTTAT CTATGTCTTT	2220
	TTTATATTCT TCCTTACATA AAATCATACC ACCACGtGGT CCGcGTAATG TTTTGTGTGT	2280
40	TGTAGTTGTT ACAAATCAG CATATTCTAC TGGATTTGGA TGTAAACCTG CCGCTACTAA	2340
	TCCTGCAATA TGTGCCATGT CTACCATTAA CTTAGCGTTT ACTTCATCTG CGATTTCTTT	2400
	AAACTTTTTG AAGTCAATTG TTCTTGAATA TGCTGATGCT CCTGCCACAA TAAGCTTAGG	2460
45	CTTATGCTCT AACGCTAATT TACGAACTTC ATCATAATTG ATTCGTTCTG TGTCTTTATC	2520
	TACTCCATAT TCAACGAAAT TGTAAGATTT ACCACTAAAA TTAACAGGCG CTCCATGTGT	2580
	CAAGTGACCA CCATGACTCA AATTCATACC TAAACTGTG TCGCCCATTT CTAATGCAAC	2640
50	TAAGTAAACA GCCATGTTG CTTGTGAACC TGAATGTGGT TGAACATTGA CATGTTCAGC	2700
	TCCAAACAAT GCTTTAGCAC GATCAATTGC GATGCTTTCA GTAACATCTA CAAACTCACA	2760

55

	TTGTGCTTCC ATAACCGCTT CCGATACAAA ATTTTCCGAT GCGATTAACT CTATGTTGCT	2880
	ATTTTGTCTC TGAAATTCTC TCTCGATTGC TTCTGCGATA ACTTTATCTT GCTTGGTGAT	2940
5	ATAAGACATA AAATCTCCCC TTCTTTCAAA AAAACTTATT GGTATTTAGC ACGTTCGCCA	3000
	CCAATCTTTT TCGGCCTAGA TGTGGCAATA GTTACAATTG CCTGTCCTAC TTGCTTTACT	3060
	GAGGTCCTTA CAGGTACACA TACATGTTTA ATATGCATGC CTATTAACGT TTGACCAATA	3120
10	TCAATTCCAC AAGGAACAGT AATATGTTTG ACCACGATCG GATCCTTCAT ATGCTGAAAA	3180
	GCGTATGTTG CCAAACCTCC TCCAGCATGT ACATCTGGAA CGACGGAAAC TTCTTCCATT	3240
	GTTAATGGAT TATACTGAGA TTTTCTATT GTTATCGCTC TGTGATATG TTCACATCCT	3300
15	TGAAAAGCAA AAGTAACGCC TGTCTCTTTA CTCACAACAT CTAATGCATT AAAAATAGTT	3360
	TCTGCAACTT CCA TCGAACC GACAGTCCCT ATTTTTCGCG CAATGACTTC CGATGTTGAA	3420
20	CATCCAATTA AACATATATC TCCTTTATTA AAAAAGGACA TATCTTTTAA TTCGTCTAAT	3480
	AACATTGTCA AATCTTTCAT AAAAGCCAC CCTTCCTAAA AATAAAAAAG GAATATAGCA	3540
	AAGTGCTACA CTCCTCTATT ATAACCTATT TAACTGTTAA CATATACTAA TTATACAGAA	3600
25	TTCTACTAG CAAATAATAT CTTTTAATTT TAAAATTAAA CTTACAAGTT CTTCATAGGT	3660
	ATGTACATAC ATTTCTTTTG TTCCACCGTA TGGATCTATA ACTTCTCCTG CTTCTTTTAC	3720
	ATATTCATGC AATGTGAAAA CATGATTTTG CAAACCAAAG TGTGCCTCTA TTAATTCTTT	3780
30	GTGCGAATAC GACATCGTCA AAATAATATC TGCTTTCAAA TCTGCTTCAG TAAATTGTTG	3840
	CGATAAGGTC GTTTCAGCTA AATGATGTTT TTCAACTAAG TCTTCAACAT AATTCGAAAC	3900
	ACCTTGATTG TTCACAGCGA ATATACCTCT TGATTCAAAT TGATGATTTG GCATAACCTC	3960
35	TTTTGCAATA CTTTCCGCTA ATGGGCTACG ACATGTGTTA CCTGTACAAA CGAATAAAAT	4020
	CTTCATAGTT CACATCCTTT AATAATGTGA TTACCTGCAG CTTTAAACAT GCGATTCATA	4080
	ATTGCTTCTG TATTATCATT CAGCTCAAAG CCGTATATAT ACGCCGCTGA AATATTTTCA	4140
40	TTTTCATCAA GTGAATGTAA CACATCATAA AGATTATGAC TTGCTTGTTT AACATCATTG	4200
	TCATCCTGAC ATAATTGAAT GAATTGCGCT TCACTTGSTA TAAACGCCAC CTTATTACTC	4260
45	GGCACAATAA AAGCTATAGA AGACCAATCT TTACCGTCAT TTCCAATTTT GCTCTCAATA	4320
	TCTGTAATAA TTGTAAGTGG TGTATTGGGT GAGTAATGCT TATACTTCAT ACCTGGTGCA	4380
	ATTGGCTGTT CAGTATCATT ATAATCAGCA TGGGCGATAC TATTCGGAAG TATTTCTGTA	4440
50	ATCATTGCTG CTGTTATAGA ACCAGGTCTT GCAATTTTAT AAGGAAAAGA TGTGCAATCT	4500
	AAAACCGTAC TTTCTAATCC TTCTTCACTT TGTTTCAGCTT GAACAATACC ATCGATACGG	4560

55

	GCACTTGGAG CAGCTAGAGG TTCATTTATG ATTTGTAATA ATTGTCTACC TACAGAATGG	4680
	CTTGGCATTC TAACAGCAAC TGATGATAAA CCTCCAGAAA CTTTTCGACA TAGATAGCCT	4740
5	AGCTTTAACG GCAATATAAA CGAAATAGGG CCCGGCCAGA ATGCCTGCAT TAACTTTTCT	4800
	ACGCGTGGAT CCAAAGTATA TGTAAAATCT TTTAATTGAC CTTTACTGTG TATATGAACA	4860
	ATAAGCGGAT TGTCAGATGG ACGGCCTTTA GCTTCATATA TTTTAGCTAC AGCTTCTTCA	4920
10	TCTGTGCGAT TTGCTGCAAG TCCATAAACT GTTTCAGTTG GTAAACCTAT TAAACCACCG	4980
	TTTAAAACAA TGTCTTTTAT TTCATTAATT TTAGGATATT GCTGTAAATC TTCATTATAT	5040
	TCTCTAACAT CCCAAATTTT AGTATCCAAC TTAATCACGC CTTTCTTATT TATCATAATA	5100
15	TAAAGCAAAA AGCTATGCAC TTAACTAATC ATAGCAAAGG CATAACTTCT AATTACCATT	5160
	TAAATGAGAC GATTCGATCG TGGCCATTTA TATCTTTAAT AATGTCGATT TTTTGTGACG	5220
20	GAAATTTATT TAAAATTATT GATTTAAGTG CCTCACCTTG ATTGTAACCA ATTTCAAAAA	5280
	CAACTGGGCT GCCTTTTTCC ATAACGTGAG GTAAATCTTC AATGATTGAT TCATAAATAG	5340
	CATATCCATG GTTATCTGCA AACAAATGCCT GATGTGGTTC GAATCTCGTA ACCGTTGGAG	5400
25	ACATCGTAAC CATATCTTTT TCATCTATAT ATGGTGGATT AGATATCAAG CCGTTCAACT	5460
	TGATACCTTC ATTAATTAAG GGCTTTAATG CATCCCCTGT TAAAAATTGT ATTTGTGATT	5520
	GATGCTTCTC AGCATTATTA CGAGCCATAT TCATTGCTTC AAGTGAAATA TCAGTAGCAA	5580
30	TAACATTTAA ATCCGGCTTT TCACATTTCA AAGTAATTGC AAGTACACCA CTACCCGTTT	5640
	CGATATCTAC GATTGTTGCA TCATCTTCTA ACTGTTGTAA GAAATGCAAC ATTACTTCTT	5700
	CAGTTTCAGG TCTTGGTATC AAACAATTTG AGTTTACATC AAACGTTCTA CCATAAAATG	5760
35	AGGCAAAGCC AACTATATAC TGTATAGGCT CTCCTAATAA CATACGTTGT AATGCTAAGT	5820
	CGAACTTCAT AATCATCGCT TTCGGCATAT CATCATGCAT GTGGACTACA AAGTCCGTAC	5880
	GCGTCCATTG AAATACATCT AACATTAAACC ATTCAGCTCG TGTGTGTTCA AACCCTTTTT	5940
40	GTGTGTAA ATGAATTGCT TCATCTAACT TTTCTTTATA ATTCACCAT ATTAAAGTTCT	6000
	TTCAATTTAT CTGTCTGCTC TGATAAAGTC AGTGCATCTA TAATTTCTTC TAAATGGCCT	6060
	TCCATAATTT GCCCTAATTT TTGAAGCGTT AGACCTATAC GATGGTCTGT TACACGGCTT	6120
45	TGTGGATAAT TATAAGTTTC AATACGTTCT GAACGATCAC CAGTACCGAC TGCTGATTAA	6180
	CGTTGTGACG CATACTTTTG TTGTTCTTCT TGAACTTTCA TATCGTATAA ACGTGCTTTT	6240
50	AACACTTTCA TTGCTTTTTT ACGGTTTTGA ATTTGAGACT TCTCAGAAGA TGTTGCAATG	6300
	ACACCAGTTG GTAAATGGGT AATACGTACT GCAGAGTCAG TTGTGTTTAC GTGCTGACCA	6360

55

ACATCTTCAA CTTCTGGTAA AACTGCCACT GTAGCTGTTG AAGTATGAAT ACGTCCACCT 6480
 GATTCTGTTT CAGGCACACG TTGAACGCGG TCGCACCAGT TTTCAAATTT CAATTTACTA 6540
 5 TACGCGCCAT TACCAGAAAC TGAGAAACTA ATTTCTTTGT AACCACCATG GTCACTTTCA 6600
 GACGCTTCTA CTATTTCACT TTTGAATCCT TGTGATTCAG CATACTTTGA ATACATACGC 6660
 ATTAAATCAC CAGCAAAAAT CGCAGCCTCA TCACCACCTG CTGCTGCTCT TATTTCTACA 6720
 10 ATAACGTCTT TGTATCATT AGGATCTTTA GGAATCAATA ATATTTTAAG CTCTTCTTCA 6780
 AGATTTGGAA GTTCAGCTTT AATACCATTA CTCTCTCTT TTAACATTTT TACTTCTTCT 6840
 TTATCATCAG TCTCACTTAA CATTTCTTCA ATATCAGCTA ATTCTTCTTT TTAGCTTTA 6900
 15 TAGTTACGAT AAACATCTAC AGTTTTTTGT AAATCAGCTT GCTCTTTAGA ATATTTACGT 6960
 AATTTATCTG AATCATTTAC AACATCTGGG TCACCTAACA GTTCATTTAA CTGTTTCGTAT 7020
 20 CTTTCTTCTA CAATATCTAA TTGATCAAAC ACTTATAATT CCTCCTTATT ATTATCACTA 7080
 GGTGCTACGA TATGGTGC GCACAAACGT GGCTCATAAC TTTCAATTGGC ACCTACTAAG 7140
 ATAATCGGAT CATCGATTTT AGCTGGTTTA CCATTTATTA ATCGTTGCGT TCTACTAGAT 7200
 25 GAAGAACCAC AAACAGCACA AACTGCTTGA AGTTTCGTTA CTTGTTCACT GACAGCCATC 7260
 AATTTAGGCA TTGGTTGGA CGGTTGCGCC CTAATATCCA TATCTAATCC AGCAACAATA 7320
 ACACGGTGTC CATCTGCTGA TAGTTTTTCT ACTATACTTA CAATTTTCATC GTCAAAAAAT 7380
 30 TGCACTTCGT CTATTCCTAT AACATCAACA TTAGTTAAGT CGTGCCTCAT AATTTCACTT 7440
 GCTTTAGAAA TATTAATCGC TTCAATGGCA TTACCATTAT GAGAGACCAC TTTTCTTTA 7500
 TGATATCGAT CATCAATCGC CGGTTTAAAT ACAACGACTT TTTGTTTAGC GTATATACCC 7560
 35 CTTCTTAGAC GTCTTATTAG TTCTTCGGAT TTACCGCTAA ACATACTACC TGTAATACAT 7620
 TCTATCCAAC CGGAATGGTA AGTTTCATAC ATTGAGAGT CCACCTTTTT CAAAACATAA 7680
 TCGCTTTATT ATATCATATT TCAAATATTC ATAAATGTCT TTnTCATAAT TATATCGATA 7740
 40 TTGTACATGA ACAATTATTT TA 7762

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2583 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

	TAAAAAAATT ATTATCAATG ATGAACTAGA ATTGACTGAA TTCCACCAAG AACTTACTTA	120
	TATTTTAGAC AACATANAAG GGAATAATAA TTATGGTAAG GAATTTGTTG CAACCGTTGA	180
5	AGAAACATTC GACATTGAAT AaAGCGGGT GgaAGCACTA TGAATCAATG GGATCAGTTC	240
	TTAACACCTT ATAAGCAAGC GGTGATGAG TTGAAAGkGA AcTTaAAGGC ATGCGCAAAC	300
	AATATGAAGT TGGTGAACAA GCGTCGCCAA TAGAATTTGT TACTGGTCGT GTTAAACCAA	360
10	TCGCTAGTAT TATAGATAAG GCAAACAAAC GACAAATACC ATTTGATAGG TTAAGAGAAG	420
	AAATGTACGA TATCGCTGGT TTAAGAATGA TGTGCCAATT TGTGAAGAT ATTGATGTTG	480
	TCGTCAATAT TTTAAGACAA AGAmAAGATT TTAAAGTAAT TGAAGAACGA GATTATATTC	540
15	GTAACACTAA AGAAAGTGGT TACCGCTCGT ATCATGTCAT TATTGAATAT CCAATTGAAA	600
	CATTACAAGG CCAAAAATTT ATATTGGCTG AGATTcAGAT TCGTACATTA GCAATGAATT	660
20	TCTGGGCAAC GATTGAACAT ACTTTACGAT ATAAATATGA TGGTGCTTAT CCGGATGAAA	720
	TTCAACATCG TTTGGAAGA GCGGCAGAAG CAGCGTATTT ACTTGATGAA GAGATGTCTG	780
	AAATTAAAGA TGAAATTcAG GAAGCTCAAA AATATTACAC GCAAAAACGT TCTAAAAAAC	840
25	ATGAAAATGA TTAACGAGGT GTTATAAATC ATGCGTTATA CAATTTTAAC TAAAGGTGAC	900
	TCCAAGTCTA ATGCCTTAAA GCATAAAATG ATGAACTATA TGAAAGrTTT TcGCATGaTT	960
	GaGGATrGTG AAAaTCCTGA AATTGTTATT yCAGTTGGTG GTGACGGTAC ATTACTACAA	1020
30	GCATTCCATC AGTATAGCCA CATGTTATCA AAAGTGGCAT TTGTTGGAGT TCATACAGGT	1080
	CATTTAGGAT TTTATGCGGA TTGGTTACCT CATGAAGTTG AAAAATTAAT CATCGAAATT	1140
	AATAATTcAG AGTTTCAGGT CATTGAATAT CCATTGCTTG AAATTATTAT GAGATACAAC	1200
35	GACAACGGCT ATGAAACAAG GTATTTAGCA TTAAATGAAG CAACGATGAA AACTGAAAAT	1260
	GGCTCAACAC TTGTTGTGGA TGTTAACtTA AGAGGGAAAC ACTTTGAGCG ATTTAGAGGC	1320
40	GATGGATTAT GTGTATCAAC ACCTTCGGGT TCAACGGCTT ATAACAAAGC GCTAGGTGGC	1380
	GCACTGATAC ATCCTTCACT TGAAGCAATG CAAATTACAG AAATTGCCTC GATAAATAAT	1440
	CGTGTGTTTA GAACGGTAGG ATCACCACCT GTATTACCAA AGCATCATAC ATGTTTAATA	1500
45	TCACCAGTTA ATCATGATAC CATTAGAATG ACGATAGATC ATGTTAGTAT CAAACATAAA	1560
	AATGTTAATT CAATACAATA CCGTGTAGCA AATGAAAAAG TGAGGTTTGC ACGTTTTAGA	1620
	CCATTCCCAT TCTGGAAACG TGTGCACGAT TCTTTCATAT CAAGTGATGA AGAACGATGA	1680
50	AATTTAAGTA TCATATATCA CAACAAGAAA CTGTTAAAAC TTTTtTAGCA CGACATGATT	1740
	TTTCTAAGAA GACAGTGAGC GCCATTAAAA ATAATGCGCG TTTAATTGTT AATGATGAAC	1800

55

AAATACCGAG TGTTAATTTA ATACCTTATG CTCGTAAGCT AGAAGTATTG TATGAAGATG 1920
 CTTTTATCAT CATAGTTACT AAACCAACA ATCAAAATTG TACGCCTTCG AGAGAACATC 1980
 5 CTCATGAAAG TTTAATCGAA CAAGTACTAT ATCATTGTCA GGAACATGGT GAAAATATTA 2040
 ACCCACATAT TGTTACGCGT CTAGATCGTA ATACAACCTGG TATTGTGATA TTCGCTAAAT 2100
 ATGGACATAT CCATCATTTA TTTTCTAAAG TAAACTTGAA AAAAAATATAT ACTTGCCTTG 2160
 10 TATATGGTAA AACCCATACA TCTGGTATTA TTGAAGCTAA TATTAGACGG TCAAAGGATA 2220
 GGATTATAAC TAGAGAAGTT GCCTCGGATG GTAAATACGC TAAACATCT TATGAAGTAA 2280
 TAAATCAGAA TGATAAATAC AGTTTATGCA AAGTTCATTT GCATACGGGA CGTACACATC 2340
 15 AAATTCGTGT ACATTTTCAA CATATTGGGC ATCCAATTGT GGGAGATTCT TTGTATGATG 2400
 GTTTTCATGA CAAAATTCAT GGTCAAGTAC TGCAATGTAC GCAAATATAT TTTGTTTCATC 2460
 CAATCAATAA GAACAATATT TATATTACAA TTGATTATAA GCAATTACTT AAATTATnCA 2520
 20 ATCAACTCTA ATnCACACAG GGGGTGTAAG TATGTCAATG AnCACAGATG AAAAAGAGCG 2580
 TGT 2583

25 (2) INFORMATION FOR SEQ ID NO: 126:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1818 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

35 ATCAAGTGAT ACATTTAACT GGTAAAGGAT TAAnAGATGC TCAAGTTAAA AAATCnGGAT 60
 ATATACAATA TGAATTTGTT AAAGAGGATT TnACAGATTT ATnGCAATT ACGGATACAG 120
 TAATAAGTAG AGCTGGATCA AATGCGATTT ATGAGTTCTT AACATTACGT ATACCAATGT 180
 40 TATTAGTACC ATTAGGTTTA GATCAATCCC GAGGCGACCA AATTGACAAT GCAAATCATT 240
 TTGCTGATAA AGGATATGCT AAAGCGATTG ATGAAGAACA ATTAACAGCA CAAATTTTAT 300
 45 TACAAGAACT AAATGAAATG GAACAGGAAA GAACGCAAT TATCAATAAT ATGAAATCGT 360
 ATGAACAAAG TTATACGAAA GAAGCTTTAT TTGATAAGAT GATTAAAGAC GCATTGAATT 420
 AATGGGGGGT AATGCTTTAT GAGTCAATGG AAACGTATCT CTTTGCTCAT CGTTTTTACA 480
 50 TTGGTTTTTG GAATTATCGC GTTTTCCAC GAATCAAGAC TTGGGAAATG GATTGATAAT 540
 GAAGTTTATG AGTTTGATA TTCATCAGAG AGCTTTATTA CGACATCTAT CATGCTTGGG 600

55

CTCATGTTAA AGCGCCACAA AATTGAAGCA TTATTTTGTG CATTACAAT GGCATTATCT 720
 GGAATTTTGA ATCCAGCATT AAAAAATATA TTCGATAGAG AAAGACCTAC ATTGCTGCGT 780
 5 TTAATTGATA TAACAGGATT TAGTTTTCTT AGCGGTCATG CTATGGGATC AACTGCATAT 840
 TTTGGAAGTG GTATCTATCT ATTAAATCGA TTAAATCAAG GTAATTCAA AGGTATTCTT 900
 ATAGGGTTAT GTGCAGCTAT GATTTTATTG ATTTCCATAT CACGTGTATA TCTAGGTGTA 960
 10 CATTATCCAA CAGATATTAT TGCCGGCATT ATTGGTGGAT TATTTTGcAT TATTTTATCA 1020
 ACGTTATTAC TTAGAAATAA ATTAATAAAT TAAATAGTAA AAAAACAAAA GCAGTAAACC 1080
 TAAAGTGTCTG TAAGGGTTTA CTGCTTTTAT AAAACGTTGT TATAACGTAT ATTGTCTTTT 1140
 15 ACGGGCATAT AAnAGGGGAA TATTTGAnAA TGACCAATCC AACAAGAACG AAACGTTGTG 1200
 GGGGGGATGT TCTATGTGGT ATTGATAATC ATTTTCAACT ACTATTATAC ATTAGTGAGA 1260
 20 ATCATTGTCA ATTAGAACT AAAACTTTTT TTGAATATTT TTTAAGAATA GTAAATAAAA 1320
 CGCATGATTA CGCTATTTTA GAAAATAAAA AAATTGTAT TTCTCATTAG AATTAGAATA 1380
 TTTAAAAGTG ATGAGGTTTA AACATTATAT TGTTTACATA CTCCTTTTGA ATTCATACAT 1440
 25 TATGAAATGT tACTTCCAAG TTCAAAATCG CACATTGAAA TGATGTGTGA AATGTTTAAA 1500
 CTACGGTCAT tTTGTGmAAA TAAAGrTAAT AACTATTCAT TTTACAATAG TGAAAAGTCA 1560
 GTATATGACA ACAATTAATA TTGCGGTAAG GCCTTGTTGTT ACAGTATTCT ATATTTAAGT 1620
 30 ACTGCAATCA GAATTAACAG AATGCCATTA ACTGATTATT AAATATTTGA GTTAATAAAT 1680
 AATTAATGAT TGTAGCTTGA AAAATTTAAA ACATGGTTAT TGATTGTGA TAAAATTTAA 1740
 ACGTAAACAA ACTAATTTAA AAAGCAACTA TTGTATAGAA AAATACAAAA TTTAAATAT 1800
 35 ATTACCTTAT TAGAAAAA 1818

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12658 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

TGTTTAAACA ATAGGGGGAA TCTTATGATT GAAAAATTAG TAACCTTTTT AAATGAGGTT 60
 50 GTTTGGAGTA AGCCATTAGT TTATGGTTTG CTAATTACTG GTGTGCTATT TACATTGCGT 120
 ATgCGATTTT TTCAAGTTAG ACATTTTAAA GAAATGATTC GATTAATGTT TCAAGGAGAG 180

	GGTACAGGTA ATATTGTCGG TGTATCTACT GCAATATTTA TAGGAGGACC TGGTGCCAGTA	300
	TTTTGGATGT GGATTACTGC GTTTTATAGGT GCAAGTAGTG CTTTTATTGA ATCTACACTT	360
5	GGTCAAATAT TCAAGAGAGT TGAAAATAAT GAATACCGTG GTGGACCAGC GTATTATATT	420
	GAATATGGTA TTGGTGGTAA ATTTGGTAAA ATTTACGGAA TTATCTTTGC TATTGTTACG	480
10	ATTATCTCAG TAGGTCTATT GCTTCCTGGT GTGCAATCTA ACGCTATAGC AAGTTCTATG	540
	CATAATGCGA TTCATGTTCC ACAATGGTTA ATGGGTGGTA TTGTTGTAGT TATTTTGGGA	600
	TTAATTATTT TTGGTGGTGT ACGTATTATT GCCAATGTTG CAACAGCCGT TGTACCATT	660
15	ATGGCAATT A TTACATACT GATGGCTGTC ATTATCATT GTATCAATAT ACAAGAAGTG	720
	CCAGCGTTAT TTGCATTAAT TTCAAATCA GCATTTGGAT TACAATCTGC TTTTGGTGGT	780
	ATCGTTGGCG CAATGATAGA GATTGGTGTT AAACGTGGAT TATATTCAAA TGAGGCTGGT	840
20	CAAGGTACAG GTCCACACGC AGCAGCGGCa gcAGaAGTAT CACATCCAAG TAAACAAGGT	900
	CTAGTACAAG CATTTTCAGT TTATATTGAT ACATTATTG TATGTACTGC AACTGCTCTG	960
	ATTATACTTA TTTCTGGTAC ATATAATGTG ACTGATGGTA CGGTTAATGC GAATGGCACA	1020
25	CCGCATTTAA TTAAAGATGG CGGTATTTAT GTTgAAAATG CAACAGGTAA AGATTATTCA	1080
	GGTACTGCGA TGTATGCACA AGCCGGCATt GATAAAGCGT TCCATGGCAG TGGTTATCAA	1140
	TTTGATCCTA CTTTCTCTGG CGTAGgTTCG TACTTTATTG cATTGCTTT ATTCTTCTTT	1200
30	GCAATTTACTA CAATTTTGTC GTACTACTAC ATTACAGAAA CAAATGTTGC TTATTTAACG	1260
	CGTAATCAAA ATAATCAAGT TTCATCGATA TTTATTAATA TTGCTCGTGT GATTATTTTG	1320
	TTGCTACAT TTTACGGTGC AGTTAAACA GCTGATGTAG CATGGGCATT CGGTGATTTA	1380
35	GGTGTAGGTC TAATGGCTTG GTTAAATATC ATTGCGATTT GGATTTTACA TAAGCCTGCC	1440
	GTAATGCTT TAAAAGATTA TGAAATTCAA AAGAAACGTT TAGGCAACGG TTATAATGCA	1500
40	GTTTATCAAC CTGATCCGAA TAAATTACCT AATGCTGTCT TTTGGTTGAA GACGTATCCA	1560
	GAACGTTTAA AACAAGCACG TGCCAAAAAG TAATCTACTT TTGTTTATAG TATATGTAGT	1620
	GATCATTTGA TAAAAAGAA AAGTATTGAG AATTTTAGGt GCTCAGAAAT TTGAATTTTA	1680
45	AAAATATAGT GTCTCTTGGT ACAATAACAA TACAATACT AGGGGCACTT TTTTATGTCA	1740
	GAATTTAAAA CTGGTAAGAT TAATAAACAT GTTTTATATA GTAATATTTT AAATAGAGAT	1800
	GTCACGTTAA GTATTTATTT ACCAGAATCT TATAATCAAC TTGTTAAATA TAATGTCATT	1860
50	CTTTGCTTTG ACGGATTAGA TTTTTCACGT TTCGGGAGAA TACAACGTAC ATATGAATCG	1920
	TTAATCAAAG AAGCGCGTAT TGATGATGCG ATCATTGTTG GATTCCATTA TGAAGACGTT	1980

55

	GTCGGTAAAG AAATATTGCC ATTTATTGAC TCGACGTTTT CTACACTGAA AGTAGGTAAT	2100
	GCAAGGTTAT TAGTAGGGGA TAGTTTAGCG GGTAGTATTG CCTTATTAAC GGC GTTGACC	2160
5	TATCCAACGA TTTT TAGTCG TGTAGCAATG TTAAGTCCAC ATTCAGATGA AAAAGTATTA	2220
	GATAAGCTAA ATCAATGTGC AAATAAAGAA CAATTGACAA TTTGGCATGT CATTGGTCTA	2280
	GATGAAAAAG ATTTTACTTT ACCAACAAAT GGTAAGCGTG CCGATTTCCT AACACCGAAT	2340
10	AGAGAATTAG CTGAACAAAT TAAGAAATAT AATATAACTT ATTATTACGA TGAATTTGAT	2400
	GGTGGTCACC AATGGAAAGA TTGGAACCA TTGCTGTCAG ATATATTATT GTATTTTTTA	2460
	AGTAAAAACA CAGATGATCA ACTTTATGAA TAATTTACAT TAGTAGATTT AGTATGAATT	2520
15	GTCTTCATAT AGTCTGGTCT ATAATATAAT TTATAAAGA TTTTACTGTT TAATTTAATT	2580
	TAAATTTGAC GAAATTGCAA AAGATGTATA ATGAATTATT TTTAATGTAA CGGTTTTCAA	2640
20	AGAAATTTGA TATAATAGCA ATAGGTTAA CAAAGGAGGA ATTCAGATGA TTTTAGGATT	2700
	AGCATTAAAT CCATCAAAGT CATTTCAAGA AGCGGTGGAT TCTTACCGTA AAAGATATGA	2760
	TAAACAGTAT TCACGAATTA AACCACATGT GACAATTAAA GCGCCATTG AAATTAAAGA	2820
25	TGGTGATTTA GATTCTGTCA TTGAACAGGT TAGAGCTCGT ATTAATGGTA TACCAGCAGT	2880
	AGAAGTTCAT GCTACAAAAG CTTCTAGCTT CAAACCAACG AACAAATGTGA TTTACTTTAA	2940
	AGTTGCGAAG ACGGACGACT TAGAAGAATT GTTTAATCGC TTTAATGGAG AAGATTTCTA	3000
30	TGGAGAAGCT GAACATGTTT TTGTGCCACA CTTTACAATA GCACAAGGAC TATCTAGCCA	3060
	AGAATTCGAA GATATTTTTG GTCaAGTAGC ATTAGCTGGG GTAGACCATA AAGAAATTAT	3120
	CGATGAATTA ACTTTGTTAC GTTTTGACGA TGACGAAGAT AAATGGAAAG TTATTGAAAC	3180
35	GTTTAAATTA GCTTAAGTAA CATAATAGTA TTGTTAATCG TAGTATGTTT GAATTAATAA	3240
	GAAATGGTC ATTTTATTG AATGTAATAA AAATGACCAT TTTCTTTATT TTAAATACG	3300
	TTTTAACCTT ACTTAGCTTT TTCTCTATTT ACTATAAAGT TGCTTCCATA AAATACAGCT	3360
40	AAGACTAAAA AGATTAATGC CGAGAAATAA AATGTATTGT TTAAATTGTT GGTAAATTGT	3420
	GTAATTAATC CGCCAAATAA TGGCCCTATC ATTGAGCCGA ATCCTTGGAT ACTATTAAAA	3480
45	ACACCCCAAG TTTCTTCTTG TTCATCTGAT TTGATAAATC GTGCCATAAA GGTATTCCAT	3540
	GCTGGTAATA AGATGCCATA CATTAGACCG ATAGCTAAAG CGATAATCCA CAAGATGTGA	3600
	ATATTAACAA TCATAGATAG AGTAAAAATT AATATCATGT ATAAATAAAA TCCGCTTAGA	3660
50	ATAACACCAT ACATAAGTT TCTGCTGCGG TTATCTATTA GTTTCGATAA AAATAGCATC	3720
	GAAACTGCAC AGCCGATACC ACCAATAATG ATTGCAACAG TATATTCAAT TGTGCTTACG	3780

55

	TGTAAGAGAA TACCAGGGAA CaACAATAAA TGGcGCTTTG TCACATCAAC AATTTGTCTC	3900
	AATTGAGCTT TAACTGGACG AGTATTATAA TTTGTAACT TTACATCGAC AAAATAATAT	3960
5	AATATCCATG CAATTAAAAC GACTAAAGAC ATCATGAAGG CAAAGCGTGT TGGGTGCACT	4020
	TTGATAAGTA GATTCAATAA AACCATACCT ACCAATAGGC CTAACAACCA TGAAAAATAA	4080
	ACATAGCCCA TTTGTTTGCC ACGTTTATCT TCTTCAACAC TGGATAACAT AATGACCCAA	4140
10	ATAGGACTAA CTGCAATACC GAGCATCATA GCACTAAATA TGATTACAAA AGGTGATGCT	4200
	GGAAACCAAA TAACTAAAAA TAAACTTGTA AATGCTAAAA TAAATCCAGT CGTTAAAACG	4260
15	ATTTTGTGTC CGAATTTTTT CAGTAAAAAT CCTATAACAA AGTTTGTAGA TGCATCAGCA	4320
	ATAAAATGTA TTGAAAATGC TAGAGACGTT ATTGCTACAG CAATGGATGT AACTGTTGGC	4380
	AAGAAATTAA TATAGCTTAG GATATACATG CCTCTCGCAA ATTCCATTAA AAATAAGATA	4440
20	ATAAGCaTTA AAATGAAATT TTTATGATTA GCGTAATTAT TTAACGAAGA ATCTTGCATA	4500
	TAAAGGAACC TTTCCATAAA TCTCTTGTTG TTGTGATGAA TGACCGATTA AATCAAGTAA	4560
	GTCTCGACAT ATTGTCTGTG TAGCATACTT AATTTTATCT TGTTCATTG TACTAATCAT	4620
25	GTTAGTTAAT TGCTCATTAC CGTTAGTTAA ACTTGCTACA ATTTTATTG CTTCTTCTGG	4680
	AGTATCAGCG ATTTTACCAA AACCTTTTTT TCAAAAGTAA AGGGCATTIT CAAGCTCTTG	4740
	ACCAGGTGCA GGATTTAGGA AAATCATTGG AATACAACGG GCGAAACCTT CAGTTATTGT	4800
30	GATACCACCA GGTTCGTAA TCATAAGTTG ACTTGATGCC ATCCATTCTAT TCATGTGTTT	4860
	GGTATAACCT AGAATCAATA CATTCTCGTT AGATTTAAAC TTAGCTGTTA AAGAACGCTT	4920
	TAGCTCTTTG CTCTTACCAC AAATCATAAC TACTTGTGCA TTTGCaCTTT tCGCTAATAT	4980
35	ATCAGTAATC ATCGTGTCAA AACCTTTAGA TACACCAAAT GCACCAGCTG aCATTAAAAT	5040
	AGTTTGCTTA TCTGGATCTA AGTTGTTGTC TATTAAACCAC TGCTTTTGAT TAATAGGCGT	5100
	TTCAAATTTG TTATCAATAG GAATACCTGT CaCTTTAACT GTTGAAGGAT CAATACCTAC	5160
40	GTCTATGAAG TCTTGTTTCG TTTCTTTTGT TGCCACATAA TATCTTGTTG AATACGGCGT	5220
	AATCCAGTTT TTATGTAAGC GATAGTCTGT CATCACTGTA GCAACTGGAA TATTAATGTT	5280
45	AAATTGCTCA GTTAGTACCG ACATAACTGG TGTAGGAAAC GTTAATAATA TTAAATCTGG	5340
	CTTTCTTTT ATCAATAAAT TAATTAACIT ATTAAGTCCA TAGTATTTGT AAAACATTT	5400
	GTCTAGTTTA TCTGGGCGGC TGTAATAAAA CCCTTTGTAC ATATTCTAA AATATTTAAA	5460
50	GCTATTGATA TACCATTTTT TACAAATAGA AGTCAAAATT GGATGAGCTT CCATAAATAA	5520
	ATCGTGCTCA ATGACGCTTA AATGGTCTAG ATTCATATCA TTAAGTTGAT TAACGATACT	5580

55

	TTGAGTAACC ATTAATAGCC ACCCTCCGTT AGTTTGAAAA TTTTATTTAA GTGTAACCTA	5700
	TTTTACGGCA TTATAAAAGA AATAAAGACG CAAAGTCGTT ACATTTATAG CAATTTTAAT	5760
5	CTATAGATGA ATTGATACAA AATAAAACGT TATTTTATAA AGCAATTTAT TGTTCTATGT	5820
	TTTATTTGTA TATTTAAAAT TATCCAGTAT ACAATTATAG CATATTTTTG GAAACAATTA	5880
	TGATATTATA CCATGTTACA AGATGGTTTT AATAATTTAA GATGAGCCAT AATTGTAAAA	5940
10	CTAATTCATA ATACCGTATG TTTTATTTTT AATAGTAGAA ATTAGAAAAT GCTGATTAGT	6000
	AGGATATAAC AGTGAAATTA TAAATTTATT AACATCAACA AAACGTGTAT AATAAACATA	6060
	TTGTAGAAAA AGGAGCGGTT CAGTTTGGAT GCAAGTACGT TGTTTAAGAA AGTAAAAGTA	6120
15	AAGCGTGTAT TGGGTTCTTT AGAACACAA ATAGATGATA TCACTACTGA TTCACGTACA	6180
	GCGAGAGAAG GTAGCATTTT TGTCGCTTCA GTTGGATATA CTGTAGACAG TCATAAGTTC	6240
20	TGTCAAAATG TAGCTGATCA AGGGTGTAAG TTGGTAGTGG TCAATAAAGA ACAATCATT	6300
	CCAGCTAACG TAACACAAGT GGTGTGCGG GACACATTAA GAGTAGCTAG TATTCTAGCA	6360
	CACACATTAT ATGATTATCC GAGTCATCAG TTAGTGACAT TTGGTGTAaC GGGTACAAAT	6420
25	GGTAAACTT CTATTGCGAC GATGATTCAT TTAATTCAA GAAAGTTACA AAAAAATAGT	6480
	GCATATTTAG GAACTAATGG TTTCCAAATT AATGAAACAA AGACAAAAGG TGCAAATACG	6540
	ACACCAGAAA CAGTTTCTTT AACTAAGAAA ATTAAAGAAG CAGTTGATGC AGGCGCTGAA	6600
30	TCTATGACAT TAGAAGTATC AAGCCATGGC TTAGTATTAG GACGACTGCG AGGCGTTGAA	6660
	TTTGACGTTG CAATATTTTC AAATTTAACA CAAGACCATT TAGATTTTCA TGGCACAATG	6720
	GAAGCATACG GACACGCGAA GTCTTTATTG TTTAGTCAAT TAGGTGAAGA TTTGTGCGAA	6780
35	GAAAAGTATG TCGTGTTAAA CAATGACGAT TCATTTTCTG AGTATTTAAG AACAGTGACG	6840
	CCTTATGAAG TATTTAGTTA TGAATTGAT GAGGAAGCCC AATTTATGGC TAAAAATATT	6900
	CAAGAATCTT TACAAGGTGT CAGCTTTGAT TTTGTAACGC CTTTGGGAAC TTACCCAGTA	6960
40	AAATCGCCTT ATGTTGGTAA GTTTAATATT TCTAATATTA TGGCGGCAAT GATTGCGGTG	7020
	TGGAGTAAAG GTACATCTTT AGAAACGATT ATTAAAGCTG TTGAAAATTT AGAACCTGTT	7080
	GAAGGGCGAT TAGAAGTTTT AGATCCTTCG TTACCTATTG ATTTAATTAT CGATTATGCA	7140
45	CATACAGCTG ATGGTATGAA CAAATTAATC GATGCAGTAC AGCCTTTTGT AAAGCAAAAG	7200
	TTGATATTTT TAGTTGGTAT GGCAGGCGAA CGTGATTTAA CTAAAACGCC TGAAATGGGG	7260
50	CGAGTTGCCT GTCGTGCAGA TTATGTCATT TTCACACCGG ATAATCCGGC AAATGATGAC	7320
	CCGAAAATGT TAACGGCAGA ATTAGCCAAA GGTGCAACAC ATCAAAACTA TATTGAATTT	7380

55

	GTTTTAGCAT CAAAAGGAAG AGAACCATAT CAAATCATGC CAGGGCATAT TAAGGTGCCA	7500
	CATCGAGATG ATTTAATTGG CCTTGAAGCA GCTTACAAAA AGTTCGGTGG TGGCCCTGTT	7560
5	GATTAATAAA AGATTTATTG ATGAAGGTAA AACTATTGAT GTTTATTTAT TCGAAGCATT	7620
	AAATAACCAG ATAATCATTG CTATACCAGA TTGGTTTTGG TCATATCAGA TGGCAATGAC	7680
	ATTAGATGAA GAAACTTGTT TTGAAGCAAT ACTCATGCAA TTGTTTGTTT TTAAAGAAGA	7740
10	GGAAGAGGCA GAATCGATTG CATCACAAC AACAGATTGG ATAGAAACAT ATAAAAAGGA	7800
	GAAAGACTAA TGAACCTAAA GCAAGAAGTT GAGTCTAGAA AGACTTTTGC GATTATTTCA	7860
	CATCCCGATG CAGGGAAAAC AACGTTAACT GAAAACTAT TGTACTTCAG TGGTGCTATT	7920
15	CGTGAAGCGG GTACAGTTAA AGGGAAGAAG ACTGGTAAAT TTGCGACAAG TGACTGGATG	7980
	AAAGTTGAAC AAGAGCGTGG TATTTCTGTA ACTAGTTCAG TAATGCAATT TGATTACGAT	8040
20	GATTATAAAA TCAATATCTT AGATACACCA GGACATGAAG ACTTTTCAGA AGATACGTAT	8100
	AGAACATTAA TGGCAGTTGA CAGTGCTGTC ATGGTCATAG ACTGTGCAA AGGTATTGAA	8160
	CCACAAACAT TGAAGTTATT TAAAGTTTGT AAAATGCGTG GTATTCCAAT CTTTACATTC	8220
25	ATTAATAAAT TAGACCGAGT AGGTAAAGAA CCATTTGAAT TATTAGATGA AATCGAAGAG	8280
	ACATTAAATA TTGAACATA CCCTATGAAT TGGCCAATTG GTATGGGACA AAGTTTCTTT	8340
	GGCATCATTG ATAGAAAGTC TAAAAAATT GAACCATTTA GAGATGAAGA AAATATATTA	8400
30	CATTTGAATG ATGATTTTGA GTTGAAGAA GATCATGCAA TTACAAATGA TAGTGATTTT	8460
	GAACAAGCGA TTGAAGAATT AATGTTGGTT GAAGAAGCGG GTGAAGCCTT TGATAATGAC	8520
	GCGCTGTTGA GTGGAGACTT AACACCTGTA TTTTCGGTT CAGCTTTAGC TAACTTTGGT	8580
35	GTACAAAATT TCTTAAATGC ATATGTTGAT TTTGCGCCAA TGCCAAATGC GAGACAAACA	8640
	AAAGTAAAGC TTGAAGTAAG CCCGTTTGAT GATTCATTTT CAGGATTIAT CTTTAAATTT	8700
	CAAGCCAACA TGGACCCTAA ACACCGTGAT AGAATTGCCT TTATGCGTGT CGTTAGTGGT	8760
40	GCATTTGAAC GTGGTATGGA TGTTACTTTG CAACGTACTA ATAAAAAGCA AAAGATCACA	8820
	CGTTCAACGT CATTTATGGC AGACGATAAA GAACTGTGA ATCATGCTGT AGCAGGCGAT	8880
45	ATCATTGGAC TATATGATAC TGGTAATTAT CAAATTGGAG ATACTTTAGT TGGTGAAAAA	8940
	CAACCTACA GTTTCCAAGA TTTACCACAA TTTACGCCAG AAATTTTAT GAAAGTTTCT	9000
	GCTAAAAACG TCATGAAACA GAAGCATTTT CATAAAGGTA TTGAACAATT AGTACAAGAA	9060
50	GGTGCGATTG AATACTATAA AACATTACAC ACAAACCAA TTATTTTAGG TGCTGTTGGT	9120
	CAGTTACAAT TTGAAGTTTT CGAACATAGA ATGAAAAACG AATATAATGT TGATGTTGTT	9180

55

	AAGATGAACA CATCAAGATC GATTTTAGTG AAAGATAGAT ATGACGATT AGTATTCTTA	9300
	TTTGAAAATG AATTGCAAC AAGATGGTTT GAAGAGAAAT TCCCTGAAAT TAAATTGTAT	9360
5	AGTTTACTTT AACAGCTCAA TTGTATAATC GAATTTGTTA CATTAAAAAT AATTGTTTCG	9420
	TTGAAGAAAA ATAAATTGTA TATTTTAAAA GAAAAAGGTA TACTATGATG TATCAAATGA	9480
10	ATAACCTATG GCATTTTGTC AGAGGGGAGT AACTTAAGAA TCATGACCGT ATAAATGaTT	9540
	CGACACTTTA TCGTCATTAC GArGATATCT TCCGGTAAAG TGGGCAATTT AAATTGCTTA	9600
	GTGAGACCTT TGCTATTTAT TTAGCATAGG TCTTTTGTGTT TGTACTTAAC TTATTTATTT	9660
15	AAAGGAGTTG TACATGTAA TGGATCCAAG TTTGATCTTA CCTTATTTAT GGGTACTTGT	9720
	CGTTTTAGTA TTTTAGAAG GCTTATTAGC AGCAGATAAC GCGATTGTTA TGGCTGTAAT	9780
	GGTTAAGCAC TTACCACCCG AACAACTGTA AAAAGCTTTG TTTTACGTT TGTTAGGTGC	9840
20	ATTTGTATTT AGATTTTTAG CATTATTCTT AATTAGTATT ATCGCGAACT TTTGGTTTAT	9900
	TCAAGCTGCA GGAGCGGTTT ACTTAATTTA TATGTCAATC AAAAATCTGT GGCAGTTCTT	9960
	TAAACACCCA GAAATTGAAA GTCCTGAAGC TGGAGATGAT CATCATTATG ATGAATCTGG	10020
25	TGAAGAGATT AAAGCAAGTA ACAAATCATT CTGGGGAACT GTGTTGAAAA TAGAATTGTC	10080
	AGATATCGCA TTTGCCATTG ATTCTATGCT TGCTGCTTTA gCTATTGCTG TAACACTTCC	10140
30	TAAAGTTGGT ATTCACTTTG GTGGTATGGA CTTAGGTCAG TTCGTAGTCA TGTTCCTAGG	10200
	TGGAATGATT GGTGTTATTC TAATGCGTTA TGCAGCAACA TGGTTTGTAG AGCTATTAAA	10260
	CAAATATCCA GGACTTGAAG GTGCAGCCTt CGCGATCGTT GGTTGGGTAG GTGTAAATT	10320
35	AGTTGTCATG GTATTAGCGC ACCCAGACAT CGCTGTATTG CCTGAGCACT TCCCACATGG	10380
	CGTATTATGG CAATCTATTT TCTGGACAGT ACTAATTGGA TTAGTAATTA TCGGTTGGTT	10440
	AGGTTCAGTT GTTAAAAATA AAAAATCGCA TAAATAATTG ATGTGAAGCG GACAATCTTA	10500
40	ATTTAGTTTA AGGTTGTCCT TTTTCATTTA ATTGAGTGAT TTATGAAAAA TGGATTTTGA	10560
	AGAATGTGAA TCAAAAGATG CGATATAGTA TTAAGAAAAT GTGCCTTTTA TATTTAGCAT	10620
	TTTTTCAATA GAAATTATAT AGATTTTAAA GCAAATTAGG TGTTAATGTG TCATAATGAT	10680
45	AAGTGATTTT ATTGAATGGA GTGGACATTA GTGGATATTG GTAAAAACA TGTAATTCCT	10740
	AAAAGTCAGT nACCSaCGTA AGCGTCGTGA ATTCTTCCAC AACGAAGACA GAGAAGAAAA	10800
50	TTTAAATCAA CATCAAGATA AACAAAATAT AGATAATACA ACATCAAAAA AAGCAGATAA	10860
	GCAAATACAT AAAGATTCAA TTGATAAGCA CGAACGTTTT AAAAATAGTT TATCATCGCA	10920
	TTTAGAACAG AGAAACCGTG ATGTTAATGA GAATAAAGCT GAAGAAAGTA AAAGTAATCA	10980

55

AAATTCATTA GATTCAGTGG ACCAAGATAC AGAGAAATCA AAATATTATG AGCAAAATTC 11100
 TGAAGCGACT TTATCAACTA AATCAACCGA TAAAGTAGAA TCAACTGAAA TGAGAAAGCT 11160
 5 AAGTTCAGAT AAAACAAAG TTGGTCATGA AGAGCAACAT GTACTTTCTA AACCTTCAGA 11220
 ACATGATAAA GAGACTAGAA TTGATTCTGA GTCTTCAAGA ACTGATTCAG ACAGCTCGAT 11280
 GCAGACAGAG AAAATAAAAA AAGACAGTTC AGATGGAAAT AAAAGTAGTA ATCTGAAATC 11340
 10 TGAAGTAATA TCAGACAAAT CAAATACAGT ACCAAAATTG TCGGAATCTG ATGATGAAGT 11400
 AAATAATCAG AAGCCATTAA CTTTACCGGA AGAACAGAAA TTGAAAAGAC AGCAAAGTCA 11460
 15 AAATGAGCAA ACAAAAACCT ATACATATGG TGATAGCGAA CAAAATGACA AGTCTAATCA 11520
 TGAAAATGAT TTAAGTCATC ATATACCATC GATAAGTGAT GATAAAGATA ACGTCATGAG 11580
 AGAAAATCAT ATTGTTGACG ATAATCCTGA TAATGATATC AATACACCAT CATTATCAAA 11640
 20 AACAGATGAC GATCGAAAAC TTGATGAAAA AATTCATGTT GAAGATAAAC ATAAACAAAA 11700
 TGCAGACTCG TCTGAAACGG TGGGATATCA AAGTCAGTCA ACTGCATCTC ATCGTAGCAC 11760
 TGAAAAAGA AATATTTCTA TTAATGACCA TGATAAATTA AACGGTCAAA AAACAAATAC 11820
 25 AAAGACATCG GCAAATAATA ATCAAAAAA GGCTACATCA AAATTGAACA AAGGGCGCGC 11880
 TACGAATAAT AATTATAGTG ACATTTTGAA AAAGTTTGG ATGATGTATT GGCCTAAATT 11940
 AGTTATTCTA ATGGGTATTA TTATTCTAAT TGTTATTTTG AATGCCATTT TTAATAATGT 12000
 30 GAACAAAAAT GATCGCATGA ATGATAATAA TGATGCAGAT GCTCAAAAAT ATACGACAAC 12060
 GATGAAAAAT GCCAATAACA CAGTTAAATC GGTCGTTACA GTTGAAAATG AAACATCAAA 12120
 35 AGATTCTTCA TTACCTAAG ATAAAGCATC TCaAGACGAA GTGGGATCAG GTGTTGTATA 12180
 TAAAAATCT GGAGATACGT TATATATTGT TACGAATGCA CACGTTGTCG GTGATAAAGA 12240
 AAATCaAAAA ATAACTTTCT CGAATAATAA AAGTGTGTT GGGAAAGTGC TTGGTAAAGA 12300
 40 TAAATGGTCA GATTTAGCTG TTGTTAAAGC AACTTCTTCA GACAGTTCAG TGAAAGAGAT 12360
 AGCTATTGGA GATTCAAATA ATTTAGTGTT AGGAGAGCCA ATATTAGTCG TAGGTAATCC 12420
 ACTTGGTGTA GACTTTAAAG GCACTGTGAC AGAAGGTATT ATTCAGGTC TGAACAGAAA 12480
 45 TGTTCTTATT GATTTGATA AAGATAATAA ATATGATATG TTGATGAAAG CTTTCCAAAT 12540
 TGATGCATCA GTAAATCCAG GTAACCGGG TGGTGCTGTC GTCAATAGAG AAGGAAAATT 12600
 50 AATAGGTGTA GTTGCACTA AAATTAGTAT GCCAAACGTT GAAATATGT CATTGCA 12658

(2) INFORMATION FOR SEQ ID NO: 128:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6048 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

	TGAAATnGAA TAGTACTATT GCAAGTGTAAGAGGTTAAT TTTTGCCnCA CGCGGGACTT	60
10	AAAAAGGCAA CCACTGGTTG TGACATATCC TTATTTACAT TTATAAATAT AAGGAGGAGG	120
	TAGTAGTGAA AGACTTATTG CAAGCACAGC AAAAGCTTAT ACCGGATCTC ATAGATAAAA	180
	TGTATAACCG TTTTCTATT CTTACTACTA TCTCAAAAA TCAGCCTGTC GGACGTCGAA	240
15	GTTTAAGCGA ACATATGGAT ATGACTGAAC GTGTACTGCG TTCTGAAACA GATATGCTTA	300
	AGAAACAAGA TTTGATAAAA GTTAAGCCTA CCGGAATGGA AATTACAGCT GAAGGTGAGC	360
	AACTGATTTG GCAATTGAAA GGTACTTTG ATATCTATGC AGATGATAAT CGTCTGTCAG	420
20	AAGGTATTAA GAATAAATT CAAATTAAGG AAGTTCATGT TGTTCTGGT GATGCTGATA	480
	ATAGTCAATC TGTTAAACA GAATTAGGTA GACAAGCAGG TCAATTACTT GAAGGCATAT	540
	TACAAGAAGA CGCGATAGTT GCTGTAACG GCGGATCCAC GATGGCATGT GTTAGTGAAG	600
25	CAATTCATTT ATTACCATAT AATGTATTCT TCGTACCAGC CAGAGGTGGA CTAGGCGAAA	660
	ATGTTGTCTT TCAGGCAAAC ACAATTGCAG CCAGTATGGc aCAACAAGCT GGCGGTTATT	720
30	ATACGACGAT GTATGTACCT GATAATGTCA GTGAaCAAC ATATAATACA TGTTGTTAG	780
	AGCCATCAGT CATAAACACT TTAGACAAAA TTAAACAAGC AAACGTTATA TTACACGGCA	840
	TTGGTGATGC GCTGAAGATG GCGCATCGAC GTCAATCACC TGAaAGGTC ATTGAACAAC	900
35	TTCAACATCA TCAAGCTGTC GGAGAGGCAT TTGGTTATTA TTTTGATACA CAAGGTCAAA	960
	TTGTCCATAA GGTAAACA ATTGGACTTC AATTAGAAGA CCTTGAATCA AAAGACTTTA	1020
	TTTTTGCACT TGCAGGAGGC AAATCGAAAG GTGAAGCAAT TAAAGCATACTTGACGATTG	1080
40	CACCCAAGAA TACAGTGTTA ATCACTGATG AAGCCGCAGC AAAGATAATA CTTGAATAAG	1140
	AGATAAAAAG TTTAATACTT TTTAAATATC ATTTTAAAGG AGGCCATTAT AATGGCAGTA	1200
45	AAAGTAGCAA TTAATGGTTT TGGTAGAATT GGTGTTTTAG CATTGAGAAG AATTCAAGAA	1260
	GTAGAAGGTC TTGAAGTTGT AGCAGTAAAC GACTTAACAG ATGACGACAT GTTAGCGCAT	1320
	TTATTAAAAT ATGACACTAT GCAAGGTCGT TTCACAGGTG AAGTAGAGGT AGTTGATGGT	1380
50	GGTTTCCGCG TAAATGGTAA AGAAGTTAAA TCATTGAGTG AACCAGATGC AAGCAAATTA	1440
	CCTTGGAAG ACTTAAATAT CGATGTAGTA TTAGAATGTA CTGGTTTCTA CACTGATAAA	1500
	GATAAAGCAC AAGCTCATAT TGAAGCAGGC GCTAAAAAG TATTAATCTC AGCACCAGCT	1560

55

	ACAGTTGTTT CAGGTGCTTC ATGTACTACA AACTCATTAG CACCAGTTGC TAAAGTTTTA	1680
	AACGATGACT TTGGTTTAGT TGAAGGTTTA ATGACTACAA TTCACGCTTA CACAGGTGAT	1740
5	CAAAATACAC AAGACGCACC TCACAGAAAA GGTGACAAAC GTCGTGCTCG TGCAGCGGCA	1800
	GAAAACATCA TCCCTAACTC AACAGGTGCT GCTAAAGCTA TCGGTAAAGT TATTCCTGAA	1860
10	ATCGATGGTA AATTAGATGG TGGTGACAAA CGTGTTCTTG TAGCTACAGG TTCATTAAC	1920
	GAATTAACAG TAGTATTAGA AAAACAAGAC GTAACAGTTG AACAAGTTAA CGAAGCTATG	1980
	AAAAATGCTT CAAACGAATC ATTCCGGTAC ACTGAAGACG AAATCGTTTC TTCAGACGTT	2040
15	GTAGGTATGA CTTACGGTTC ATTATTCGAC GCTACACAAA CTCGTGTAAT GTCAGTTGGC	2100
	GACCGTCAAT TAGTTAAAGT TGCAGCTTGG TATGATAACG AAATGTCATA TACTGCACAA	2160
	TTAGTTCGTA CATTAGCATA CTTAGCTGAA CTTTCTAAAT AATTTTAGTA TAGTTTTTAT	2220
20	TCAAATACGC TAGTGCTCAG AACTATTTAG CATTAAATTA AGCTTATGAG TAAGCGGGGA	2280
	GCACAAACGC TTCTCCGCTT ATTTTATAT AAAATTCCT AATTACAAGG AGGAAACACC	2340
	ATGGCTAAAA AAATTGTTTC TGATTAGAT CTTAAAGGTA AAACAGTCCT AGTACGTGCT	2400
25	GATTTTAACG TACCTTTAAA AGACGGTGAA ATTACTAATG ACAACCGTAT CGTTCAAGCT	2460
	TTACCTACAA TTCAATACAT CATCGAACAA GGTGGTAAAA TCGTACTATT TTCACATTTA	2520
30	GGTAAAGTGA AAGAAGAAAG TGATAAAGCA AAATTAACCT TACGTCCAGT TGCTGAAGAC	2580
	TTATCTAAGA AATTAGATAA AGAAGTTGTT TTCGTACCAG AAACACGCGG CGAAAACTT	2640
	GAAGCTGCTA TTAAAGACCT TAAAGAAGGC GACGTATTAT TAGTTGAAAA TACACGTTAT	2700
35	GAAGATTTAG ACGGTAAAAA AGAATCTAAA AATGATCCAG AATTAGGTAA ATACTGGGCA	2760
	TCTTTAGGTG ATGTGTTTGT AAATGATGCT TTTGGTACTG CGCATCGTGA GCATGCATCT	2820
	AATGTTGGTA TTTCTACACA TTTAGAACT GCAGCTGGAT TCTTAATGGA TAAAGAAATT	2880
40	AAGTTTATTG GCGGCGTAGT TAACGATCCA CATAAACCAG TTGTTGCTAT TTTAGGTGGA	2940
	GCAAAAGTAT CTGACAAAAT TAATGTCATC AAAAAGCTAG TTAACATAGC TGATAAAATT	3000
	ATCATCGGCG GAGGTATGGC TTATACTTTC TTAAGAGCGC AAGGTAAAGA AATTGGTATT	3060
45	TCATTATTAG AAGAAGATAA AATCGACTTC GCAAAAGATT TATTAGAAAA ACATGGTGAT	3120
	AAAATTGTAT TACCAGTAGA CACTAAAGTT GCTAAAGAAT TTTCTAATGA TGCCAAAATC	3180
	ACTGTAGTAC CATCTGATTC AATCCAGCA GACCAAGAAG GTATGGATAT TGGACCAAAC	3240
50	ACTGTAAAAAT TATTTGCAGA TGAATTAGAA GGTGCGCACA CTGTTGTATG GAATGGACCT	3300
	ATGGGTGTAT TCGAGTTCAG TAACTTTGCA CAAGGTACAA TTGGTGTATG TAAAGCAATT	3360

55

	TCTTTAGGTT TTGAAATGA CTTCACTCAT ATTTCAACTG GTGGCGGCGC GTCATTAGAG	3480
	TACCTAGAAG GTAAAGAATT GCCTGGTATC AAAGCAATCA ATAATAAATA ATAAAGTGAT	3540
5	AGTTTAAAGT GATGTGGCAT GTTTGTTTAA CATTGTTACG GGAAAACAGT CACAAGATGA	3600
	CATCGTGTTC CATCACTTTT CAAAAATATT TACAAAACAA GGAGTGTCCT TAATGAGAAC	3660
10	ACCAATTATA GCTGGTAACT GGAAAATGAA CAAAACAGTA CAAGAAGCAA AAGATTCGTC	3720
	AATACATTAC CAACACTACC AGATTCAAAA GAAGTAGAAT CAGTAATTTG TGCACCAGCA	3780
	ATTCAATTAG ATGCATTAACT TACTGCAGTT AAAGAAGGAA AAGCACAAGG TTTAGAAATC	3840
15	GGTGCTCAAA ATACGTATTT CGAAGATAAT GGTGCGTTCA CAGGTGAAAC GTCTCCAGTT	3900
	GCATTAGCAG ATTTAGGCGT TAAATACGTT GTTATCGGTC ATTCTGAACG TCGTGAATTA	3960
	TTCCACGAAA CAGATGAAGA AATTAACAAA AAAGCGCAGC CTATTTTCAA ACATGGAATG	4020
20	ACTCCAATTA TATGTGTTGG TGAAACAGAC GAAGAGCGTG AAAGTGGTAA AGCTAACGAT	4080
	GTTGTAGGTG AGCAAGTTAA GAAAGCTGTT GCAGGTTTAT CTGAAGATCA ACTTAAATCA	4140
	GTTGTAATTG CTTATGAACC AATCTGGGCA ATCGGAACTG GTAAATCATC AACATCTGAA	4200
25	GATGCAAATG AAATGTGTGC ATTTGTACGT CAACTATTG CTGACTTATC AAGCAAAGAA	4260
	GTATCAGAAG CAACTCGTAT TCAATATGGT GGTAGTGTTA AACCTAACAA CATTAAAGAA	4320
30	TACATGGCAC AAAGTATAT TGATGGGGCA TTAGTAGGTG GCGCATCACT TAAAGTTGAA	4380
	GATTTTCGTAC AATTGTTAGA AGGTGCAAAA TAATCATGGC TAAGAAACCA ACTGCGTTAA	4440
	TTATTTTAGA TGGTTTTGCG AACCGCGAAA GCGAACATGG TAATGCGGTA AAATTAGCAA	4500
35	ACAAGCCTAA TTTTGATCGT TATTACAACA AATATCCAAC GACTCAAATC GAAGCGAGTG	4560
	GCTTAGATGT TGGACTACCT GAAGGACAAA TGGGTAACTC AGAAGTTGGT CATATGAATA	4620
	TCGGTGCAGG ACGTATCGTT TATCAAAGTT TAACTCGAAT CAATAAATCA ATTGAAGACG	4680
40	GTGATTTCTT TGAAAATGAT GTTTTAAATA ATGCAATTGC ACACGTGAAT TCACATGATT	4740
	CAGCGTTACA CATCTTTGGT TTATTGTCTG ACGGTGGTGT ACACAGTCAT TACAAACATT	4800
	TATTTGCTTT GTTAGAACTT GCTAAAAAAC AAGGTGTTGA AAAAGTTTAC GTACACGCAT	4860
45	TTTTAGATGG CCGTGACGTA GATCAAAAAT CCGCTTTGAA ATACATCGAA GAGACTGAAG	4920
	CTAAATTCAA TGAATTAGGC ATTGGTCAAT TTGCATCTGT GTCTGGTCGT TATTATGCAA	4980
50	TGGATCGTGA CAAACGTTGG GAACGTGAAG AAAAAGCTTA CAATGCTATT CGTAATTTTG	5040
	ATGCCCAAC TTATGCAACT GCCAAGAAG GTGTAGAAGC AAGCTATAAT GAGGGCTTAA	5100
	CTGACGAATT CGTAGTACCA TTCATCGTTG AGAATCAAAA TGACGGTGTT AATGATGGAG	5160

55

CGAACAGAGC ATTCGAAGGC TTAAAGTTG AACAAAGTTAA AGACTTATTC TATGCAACAT 5280
 TCACTAAGTA TAATGACAAT ATCGATGCGG CTATCGTCTT CGAAAAAGTT GATTTAAATA 5340
 5 ATACAATTGG TGAAATTGCA CAAAATAACA ATTTAACTCA ATTACGTATT GCAGAACTG 5400
 AAAAATACCC TCACGTTACT TACTTTATGA GTGGTGGACG TAACGAGGAA TTAAAGGTG 5460
 AACGCCGTCG TTTAATTGAT TCACCTAAAG TTGCAACGTA TGACTTGAAA CCAGAAATGA 5520
 10 GTGCTTATGA AGTTAAAGAT GCATTATTAG AAGAGTTAAA TAAAGGTGAC TTGGACTTAA 5580
 TTATTTTAAA CTTTGCTAAC CCTGATATGG TTGGACATAG TGGTATGCTT GAGCCGACAA 5640
 TCAAAGCAAT CGAAGCGGTT GATGAATGTT TAGGAGAAGT GGTTGATAAG ATTTTAGACA 5700
 TGGACGGTTA TGCAATTATT ACTGCTGACC ATGGTAACTC TGATCAAGTA TTGACGCaTG 5760
 ATGATCAACC AATGACTACG CAWACAACGA ACCCAGTACC AGTGATTGTA ACAAAGAAG 5820
 20 GCGTTACACT TAGAGAACT GGTGCTTAG GTGACTTAGC ACCTACATTA TTAGATTTAT 5880
 TAAATGTAGA ACAACCTGAA GATATGACAG GTGAaTCTTT AATTAAACAC TAATATTGTA 5940
 AAAGATGTTA AGTAAACGCT TAATGACACT TATTTTTTGA AAATAATAGT AATATCnTTT 6000
 25 TGTTAAATGA AAGAATAAAG CTATAATAAT TATAGAATAA CTATTTAn 6048

(2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5602 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

AAAGAAAGTGC AAGATATCAT CGCATTAAATT AAGTCGTTAC AAAGTGTAAT TGTAGACaTC 60
 40 GCTTCCAATA ATGTTGATAC AATTATGCCT GGTATACTC ATTTACAGCG TGCACAGCCA 120
 ATTTCAATTG CACATCATAT TATGACTTAT TTTTGGATGT TACAACGAGA CCAACAACGA 180
 TTTGAAGATA GTTTAAACG AATCGATATT AATCCTTTAG GTGCAGCAGC CTTAAGTGGT 240
 45 ACCACATACC CTATCGATAG ACACGAGACA ACAGCATTGT TGAACCTTGG CAGTCTCTAT 300
 GAGAATAGCC TAGATGCTGT TAGTGACAGA GACTATATTA TTGAAACATT GCATAATATT 360
 TCTTTAACGA TGGTTCACTT ATCACGCTTT GCAGAGGAAA TTATTTTCTG GTCCACAGAC 420
 50 GAAGCTAAAT TCATTACATT ATCAGATGCA TTTTCAACTG GCTCATCTAT TATGCCACAA 480
 AAGAAAAATC CTGATATGGC AGAATTAATT AGAGGTAAAG TTGGTCGAAC GACTGGTCAT 540

	GAAGATAAAG AAGGTTTATT CGATGCTGTC CATAACAATTA AAGGTTCTTT ACGTATTTTC	660
	GAAGGTATGA TTCAAACGAT GACAATTAAT AAAGAACGAC TCAATCAAAC TGTAAAGAA	720
5	GATTTTTCAA ATGCAACGGA ACTAGCAGAT TATTTAGTAA CTAAAAATAT TCCATTTAGA	780
	ACTGCACATG AAATTGTAGG AAAAATCGTC TTAGAATGTA TACAACAAGG TCATTATTTA	840
10	TTAGATGTTT CTTTAGCAAC ATATCAACAA CATCATTCTA GTATTGATGC CGATATTTAC	900
	GATTATTTGC AGCCTGAAAA TTGTTTAAAA CGACGTCAA GTTACGGTTC AACAGGTCAA	960
	TCATCGGTCA AACACAACCT TGATGTTGCT AAACAATTAC TATCACAATA AATACGTTAA	1020
15	TCTACCTACC CACAATGTCT ATTAAAATTA CATTGTGGGT ATTTTAATGC TCTCTTCGTC	1080
	TTGTTGAACA TCACATTTTT AAGATTCCTA AAATGTTTGA TAATTCTTTT AAATTTATAT	1140
	TACAAAAATG TTATAAATTG TAAAAGAAAT GTGTAAAGCG TTTTCACAAG CAGGTTTTTG	1200
20	TAGTATTTTA AAATTGTTAG ACTACAAATA AAGAGATGAA AGGATAAAGA CTATGACTAA	1260
	CTCTTCGAAA AGCTTCACTA AATTTATGGC TGCTTCTGCT GTTTTTACTA TGGGATTTTT	1320
	ATCAGTACCT ACTGCTGGCG CTGAACAAAC AAATCAAATT GCAAATAAAC CTCAGGCTAT	1380
25	TCAATGGCAT ACAAATTTAA CGAATGAGCG ATTCACTACT ATCGCACATC GTGGCGCAAG	1440
	TGGCTATGCA CCCGAGCATA CGTTTCAAGC ATATGATAAG AGTCATAATG AGTTAAAGC	1500
30	ATCTTATATC GAAATTGATT TACAACGTAC CAAAGATGGC CATTTAGTTG CTATGCATGA	1560
	TGAAACTGTT AACCGTACAA CAAATGGACA CGGTAAAGTT GAGGATTATA CCCTTGATGA	1620
	'ATTAAAACAG TTAGATGCAG GAAGTTGGTT TAATAAAAAA TATCCAAAAT ACGCAAGAGC	1680
35	AAGTTATAAA AATGCTAAAG TACCCACTTT AGATGAAATT TTAGAACGTT ATGGCCCGAA	1740
	TGCAAACTAT TATATTGAAA CAAAGTCACC TGATGTATAC CCAGGAATGG AAGAACAATT	1800
	ATTAGCTTCA TTGAAAAGC ATCACCTTTT AAATAACAAT AAATTAAAAA ATGGACATGT	1860
40	AATGATTCAA TCATTTTCTG ACGAAAGTTT AAAGAAAATT CATCGTCAA ATAAGCATGT	1920
	GCCATTAGTA AAATTAGTTG ATAAAGGTGA ACTACAACAA TTTAACGACC AACGCTTAAA	1980
45	AGAGATACGC TCTTATGCGA TTGGATTAGG TCCTGATTAT ACAGATTTAA CTGAACAAAA	2040
	TACCCATCAT TTAAAAGACT TAGGATTTAT AGTACATCCT TATACAGTGA ATGAAAAGC	2100
	TGATATGTTA CGATTAAATA AATATGGCGT TGATGGTGTC TTTACAAATT TCGCTGATAA	2160
50	ATATAAAGAA GTCATTAGT AGTAATGTTA AACTAGAAAA CATAAATACA AAAATATAGC	2220
	TATTACTATA AAAACAGCA GTAAGATATT TCCAAATTGA AATTATCCTA CTGCTGTCTT	2280
55	TTTGGGAGTG GGACAGAAAT GATATTTTCG CAAAATTTAT TTCGTCGTCC CACCCCACT	2340

	TTGTCTGTAG AAATTGAGGA GCTAATTTCT CTGTGTCGGG GCTCCACCCC AACTTGCACA	2460
	CTATTGTAAG CTGACTTTCC GCCAGCCTCT GTGTTGGGGC CCCGCCAACT TGCACACTAT	2520
5	TGTAAGCTGA CTTTCCACCA GCCTCTGTGT TGGGGCCCCG ACTATTTTTG AAAAGAGCGT	2580
	GTTACACGGG CATTGTTTTA CAGTCAACTA CTGCTAAAAT AAAATTAACG AGCTTAGGGC	2640
10	TTTGTTTTCT GTCCCAAGCT CGTTAAATCA CATATGATAA TTAATTATGC CCAACCACGA	2700
	TATCTAGCTG CTTCTGCTGT ACGTTTAATA CCTATGATAT ATGCTGCAAG TCTCATATCT	2760
	ATTTTTCGGT TTTGAGACAA TCGTAAATC GTATCAAATG CCGCTTCTAA TTTTTCACGT	2820
15	AGCTTTTCAT TAACTTCTTC TTCAGACCAA TAATAACCTT GATTATTTTG TACCCATTCTG	2880
	AAGTAAGAAA CCGTACACC ACCAGCACTT GCTAATACGT CTGGAACATA TAATATACCA	2940
	CGTTCAGTTA AAATACGTGT TGCTTCTGGT GTTGTAGGTC CATTAGCAGC TTCAACAACG	3000
20	ATACTAGCTT TAATATCATG TGCATTGTCT TCTGTAATTT GGTTTGAAAT AGCCGCTGGT	3060
	ACTAAAATGT CACAATCTAA TTCAAACAAT TCTTTATTTG AGATTGTTTC TTCAAATAAA	3120
	TTTGTTACCG TACCAAACT ATCAGCAGCG TCTAATAAAT AATCTATATC TAAGCCATTT	3180
25	GGATCGTGTA ATGCACCGTA AGCATCAGAG ATACCTACAA TTTTGCACC TAAATCATAT	3240
	AAGAATTTAG CTAAGAACT TCCGGCATT ACGAAACCTT GAATAACAAC CTTGGCACCT	3300
30	TCAATTTGCA TATTACGACG TTTTGCAGCT TGTTCAATTG CAATAACTAC ACCTAGTGCA	3360
	GTTGATCTGT CGCGTCCATG AGAACCACCC AATACAATTG GTTTACCTGT GATGAAACCT	3420
	GGTGAATTAA ATTTATCTAA TGCACTATAT TCATCCATCA TCCAAGCCAT AATTTGTGAG	3480
35	TTTGTAATA CATCTGGTGC TGGAATATCT TTGTTGGGAC CTACGAATTG TGAAATTGCT	3540
	CTTACATATC CGCGTGATAA ACGTTCAACT TCATGAATGC TCATTTGACG TGGATCACAA	3600
	ACGATACCAC CCTTACCACC ACCGTATGGT AAGTTTACAA TGCCACATTT CAAAGTCATC	3660
40	CACATTGATA ATGCTTTTAC TTCTTCTTCA TCAACATCTG GGTGGAAACG CACGCCCCCT	3720
	TTTGTTGGTC CAACAGCATC ATTATGTTGC GCACGGTAAC CTGTGAATGT TTTTACTGTG	3780
	CCATCATCCA TTCGTACAGG GATACGCACT TGTAACATTC TTAAAGGTTT TTTAATTAAA	3840
45	TCGTACATTC CTTCGTCAAA TCCCAATTTA TGCAATGCTT CTTTAATAAT TCCTTGAGTA	3900
	GAAGTTACTA AATTATTGTT CTCAGTCATG ATCCTTTTCG CCTCTTCTTT ACCTAATGAT	3960
50	TTGCTTTTCA AACATATTGT AACATAACGT ATTCCTTTT AAAGCCCTTA CAAACTGATT	4020
	GTTACAACCT TTTGACATTA TTGAAATACA TGTCTTATTT TTTCAAGTGC AAGGTCCAAT	4080
55	TCTTCTTTAG TAATAATTAA TGGTGGTGCA AAACGAATGA CAGTATCATG CGTTTCTTTA	4140

ACACCTATAA ACAACCACG TCCACGGACT TCTTTAATTG ATGGATGATC AATTTGCTTT 4260
 AATTGTTCTT TAAAATAATC TCCTAATTCT AAAGAGCGGC CTGGTAAATC CTCATCAACG 4320
 5 ATAACATCTA ATGCAGCAAT TGATGCAGCA CAAGCAAGTG GATTACCACC AAATGTTGAA 4380
 CCATGTGAGC CAGGTGTAAA GACATCTAAT ACTTCTTTAT CTGCTAATAC AACAGAAATT 4440
 GGGAAGACTC CACCACCTAG TGCTTTACCT AAAATATAGA CATCAGGTTT TACATTATCC 4500
 10 CAATCCGTAG CAAATAATTT ACCCGAACGA CCTAATCCTG CTTGGATTTC GTCAGCAATA 4560
 AATAAGACAT TATGTTTCATC ACATAATTCT CTAATTGCTT TCAAATATCC TTCTGGCGGT 4620
 ATATTTATAC CCGCTTCACC TTGAATTGGT TCTACTAAAA CTGCTGCAGT ATTTTCATTA 4680
 15 ATTGCAGCTT TCAATGCATC TACATCTCCA AAATCAACTT TTCTAAATCC ATCTAATAAC 4740
 GGACCATAAC CACGTTGGTA TTCTGCTTCT GAAGATAATG AAAGTGGCGC CATTGTTTGA 4800
 20 CCATGGAAGT TACCATTAAA TGCAATGATT TCTGCTTTAT TTGGCTCAAT TCCTTTAACA 4860
 TCGTATGCCC AGCGTCGTGC TGCTTTCAAA GCTGTTTCTA CTGCTTCAGC ACCTGTATTTC 4920
 ATTGGAAGG CTTTATCTTT ACCTGCCAGT TTACAAATTT TTTCTGACCA TTCACCTAAG 4980
 25 TTATCACTAT GAAAAGCACG TGAAACTAAA GTCACCTTAT CAGCTTGATC TTTTAATGCT 5040
 TGAATAATTT TCGGATGTCT ATGACCTTGG TTAACAGCGG AATATGCAGA TAACATATCC 5100
 ATATATTTAT TGCCTTCAGG ATCTTTAACC CATACCCCTT CAGCTTCTGA AATGACAATT 5160
 30 GGCAATGGTA AATAATTATG TGCTCCGTAA TGATTGTGTA ACTCAATAAT TTTTTCAGAT 5220
 TTAGTCATCA TATCTCCCCT TTTCATCATT TATAACTATT ATACATGAAA CATTATCCAA 5280
 ATAATTACAT TAGTTTTCAA AGCAGATACT TTTCCACCAA AAAAGATGAA ATAATCACTA 5340
 35 AGTTTCATTA AATTTGTCTA TTTTGAAAAC CCTTACATT ATAATGACAT AATTACTTAA 5400
 ATGATTACAA GCAAAGAAT TGATAATTTT AACTTAATC AAAAGTATAT TTTACTAAGA 5460
 40 ATATTTTTAT TTATAAATAT TGAAAACCAC TAACAAATTG CATAACAAT ATCATTAGTG 5520
 GTAACAGTTA AACACTTATT TATCTTTACG GGGTAATGGG TTAAAACCCT TnCATTAAAA 5580
 TTGGATGnCC ATAAAATTAG GG 5602

45 (2) INFORMATION FOR SEQ ID NO: 130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5924 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

	TAACCCCATTTTACCTGGAA AAATCgTTTG CGATGCaATm GCaTTtGaAT ATAAATACAT	60
	TTTACGTATa GAATTATAAA AgGTTTCATT CaaATCTTAG GGTCAAAAAT GTTATAATAT	120
5	TTTTATGTCA AATTTAAAAC AGTAACACTT ATTTACAAGG TTGCAATATT TTGAAGTAAT	180
	AAAGGAAGTG TCGCGTATTT TAACTTTTTTC AGAGCAAAAT GCACTCGCGA AAATAGATGA	240
10	TTTAATGAAT ACTTATTGCA ATCAATGTCC AATCAAACT CGTCTGCGTA AATTAGAGGG	300
	GAAAACGAAG GCGCATCATT TTTGTATCAA TGAGTGTTCa ATAGGGAAAG AAATAAAACA	360
	ATTAGGAAAT GAACTTCAAT AGGAGGAAGT CAAATGAAAA TTATATCTAT ATCAGAAACA	420
15	COGAACCACA ACACAATGAA GATTACACTT AGTGAAAGCA GAGAAGGTAT GACATCAGAT	480
	ACGTATACTA AAGTTGATGA TTCACAGCCA GCATTTATTA ATGACATCTT AAAGGTTGAA	540
	GGCGTTAAAT CAATTTTCCA TGTTATGGAC TTTATTTTCAg TAGATAAAGA AAATGACGCA	600
20	AATTGGGAAA CAGTATTGCC AAAAGTAGAG GCTGTATTG AATAAATTTT TCATCAACTA	660
	GTATTCGGGG GGAATAAAGT ATATGGAAAT TTTACGTATA GAGCCAACAC CAAGTCCAAA	720
	TACAATGAAA GTTGTTTTGT CATATACAAG AGAAGACAAG TTATCTAATA CTTATAAAAA	780
25	AGTAGAAGAA ACACAACCAA GATTTATAAA TCAGTTGTTA TCTATAGATG GTATCACTTC	840
	CATTTTTTCAT GTCATGAACT TCTTAGCTGT TGATAAGGCA CCAAAAGCTG ATTGGGAAGT	900
	CATATTACCT GATATTAAAG CTGCTTTTTTC TGATGCGAAT AAGGTTTTAG AATCTGTAA	960
30	TGAACCTCAA ATTGACAATC ATTTTGGTGA AATTAAAGCT GAATTATTAA CTTTTAAGGG	1020
	TATACCGTAT CAAATTAAGC TAACTTCTGC TGACCAAGAA TTAAGAGAAC AATTACCACA	1080
35	AACATATGTT GACCATATGA CTCAAGCGCA AACAGCACAT GACAATATTG TTTTATGCG	1140
	TAAATGGCTA GATTTAGGAA ATCGCTATGG AAATATTCAA GAAGTAATGG ATGGTGTCTT	1200
	AGAAGAAAGTG CTAGCTACCT ATCCAGAATC ACAGTTACCC GTATTGGTAA AACATGCTTT	1260
40	AGAAGAAAAT CACGCAACTA ATAATTATCA TTTCTATCGA CATGTCTCTT TGGATGAATA	1320
	TCATGCAACT GATAATTGGA AGACTCGATT ACGAATGTTA AACCATTTTC CAAAGCCGAC	1380
	TTTTGAAGAT ATACCGCTGC TTGATTTAGC TTTATCTGAT GAAAAAGTAC CGGTTAGACG	1440
45	TCAAGCGATT GTATTATTAG GTATGATTGA AAGTAAAGAA ATTTTACCGT ATTTATATAA	1500
	GGGGCTTCGT GATAAAAGTC CTGCTGTAAG AAGAACAGCA GGGGATTGCA TAAGCGATTT	1560
	AGGGTATCCA GAGGCACTAC CAGAAATGGT GCTACTATTA GATGATCCAC AGAAAATCGT	1620
50	TAGGTGGCGT GCTGCTATGT TTATCTTTGA TGAAGGTAAT GCAGAGCAGC TTCCCGCACT	1680
	AAAAGCCCAT ATTAATGACA ATGCGTTTGA AGTTAAATTA CAAATTGAAA TGGCCATATC	1740
55		

	AATTTAATTG GAGGAATTAA ATATGAATGC ATATGATGCT TATATGAAAG AAATTGCGCA	1860
	ACAAATGCGT GGCGAATTAA CTCAAAATGG TTTTACAAGT TTAGAAACGA GCGAACAGct	1920
5	ATCGGAGTAT ATGAACCAAG TAAATGCTGA TGACACTACT TTTGTAGTTA TTAAGTCTAC	1980
	ATGCGGCTGT GCAGCTGGAT TAGCAAGACC AGCTGCAGTA GCAGTTGCAA CACAAAATGA	2040
10	ACATAGACCT ACAAATACAG TTACAGTTTT TGCTGGGCAA GATAAAGAAG CAACTGCTAC	2100
	AATGCGAGAA TTCATTCAGC AAGCACCATC TAGTCCTTCG TATGCTTTAT TCAAAGGTCA	2160
	AGATTTAGTT TATTTTATGC CTAGAGAATT TATCGAAGGT AGAGATATTA ATGACATTGC	2220
15	AATGGACTTA AAGGATGCCT TTGACGAAAA TTGTAAATAG TACACATAAA TAAATATAAA	2280
	GGTTAACACA TTTTATAATA TTAATAATGG TGTCTGTCAT TGAAAATAGA GAATATAGTT	2340
	GTATTCTATT TGTAAATAA AGTCCGTTTT TACCaaCTAT ATTTTCTAGA AATTTAACTG	2400
20	TTTTAATAGG ACATCAAACA TAATATTCaA ATCaTGTGTT AACCTCTTTT TTAATAATTTT	2460
	TTAGCATTAA AGTTATAGAT TTGGGTAAAC AATTACCAAT TGGAAACATA TATCACGTTA	2520
	CGATGGGGTA GGTACTTAAT CAGCATTTTA TAAATAAAGT AACGGAATTC ATGATATTAA	2580
25	TATCATATTC CTAAATGAG TGATAACAAA ATGCTACATA AAGTTAAGTT ATATCAAAT	2640
	AAATATACAT ACTATAAATA ATGAAAATGA GGTGTTATCG CATATGTTGA ATTCATTGTA	2700
	TGCAGCATAT CACAGTCTTT GTGAAGAAGT TTTAGAAATA GGAAATACAC GAAATGATCG	2760
30	CACAAATACA GGTACGATTT CGAAATTTGG TCATCAACTT CGCTTTGACT TATCTAAAGG	2820
	ATTTCCACTA TTAACGACAA AGAAAGTTTC TTTTAAATTA GTAGCAACCG AATTATTATG	2880
35	GTTCAATAAA GGAGATACAA ACATCCAATA CTTATTAAAA TATAATAATA ATATATGGAA	2940
	CGAATGGGCT TTTGAAAATT ATATCAAATC AGACGAGTAT AAAGGTCCAG ATATGACAGA	3000
	TTTCGGGCAT CGTGCAATGA GTGATCCTGA ATTTAACGAA CAATATAAAG AACAAATGAA	3060
40	ACAATTTAAG CAACGTATTC TTGAAGATGA TACATTTGCG AAGCAATTCG GGGATTTAGG	3120
	AAATGTTTAT GGTAAACAAT GGCGAGATTG GGTGATAAAA GATGGTAATC ATTTTGATCA	3180
	ACTTAAACAA GTAATTGAAC AAATTAAGCA TAATCCAGAT TCAAGGCGAC ACATCGTATC	3240
45	TGCATGGAAT CCAACAGAAA TTGATACAAT GGCACCTCCG CCTTGTCATA CCATGTTCCA	3300
	GTTTTATGTC CAAGATGGTA AGTTAAGTTG CCAGTTATAC CAACGTAGCG CAGATATCTT	3360
	TTTAGGTGTG CCATTTAATA TCcGCagctA CGCTTTATTG ACACACCTTA TTGCCAAAGA	3420
50	ATGTGGACTT GAAGTGGGTG AATTTGTGCA TACATTTGGA GATGCACATA TTTATTCAAA	3480
	TCATATTGAT GCGATTCAAA CACAATTAGC ACGTGAAAGC TTCAATCCTC CAACATTAAA	3540

55

	TGAATCACAT CCAGCAATAA AAGCTCCAAT AGCAGTGTAG TCATTGCATA GTTAGCTAAC	3660
	CATATAGACA TCAAAATGAC ATCATAGTAT TTTCAAGTGC AAAAAAGTAC TTTTTTGTGT	3720
5	TAAACGTTTT CATAAATTAT GCAAAATCAT TATTTCTATC ACACTTTATG ATAAAAATTG	3780
	TGTTAAATTA AAGATAACTT AGTAATAAAA AATGAAATGA TAGAAGAAGG AGGATAATTA	3840
	TGACTTTATC CATTCTAGTt GCACATGACT TGCAACGAGT AATTGGTTTTt GAAAAATCAAT	3900
10	TACCTTGGcA CCTACCAAAT GATTTGAAGC ATGTTAAAAA ATTATCAACA GGTCACTT	3960
	TAGTAATGGG TCgTAAGACA TTTGAATCGA TTGGTAAACC ACTACCGAAT CGTCGAAATG	4020
15	TTGTACTTAC TTCAGATACA AGTTTCAACG TAGAnGGCGT TGATGTAATT CACTCTATTG	4080
	AAGATATTTA CCAACTACCG GGCCATGTTT TCATATTTGG AGGGCAAACA TTATTTGAAG	4140
	AAATGATTGA TAAAGTGGAC GACATGTATA TTA CTGTTAT TGAAGGTAA TTCCGTGGTG	4200
20	ATACGTTCTT TCCACCTTAT mCATTkGAgx CTGGGAAGTT GCCTCTTCAG TTGAAGGTAA	4260
	ACTAGATGAG AAAAAATACAA TTCCACATAC CTTTCTACAT TTAATTCGTA AAAAAATAGG	4320
	GGGAAAACGA CCATGACAAA ACAGATTATA GTAACAGACT CAACATCCGA TTTATCTAAA	4380
25	GAATACTTAG AAGCAAACAA CATTCAATGTA ATTCCTTTAA GTTTAACTAT TGAAGGAGCT	4440
	TCATACGTTG ACCAAGTAGA TATTACATCA GAAGAATTTA TTAATCATAT TGAATATGAT	4500
	GAAGATGTAA AGACAAGTCA GCCAGCCATA GGTGAATTTA TATCTGCTTA TGAAGAACTA	4560
30	GGAAAAGATG GCTCTGAAAT CATAAGTATT CATCTTTCTT CAGGATTAAG TGGTACATAT	4620
	AACACTGCTT ACCAAGCAAG TCAAATGGTA GATGCTAATG TAACTGTTAT TGATTCAAAA	4680
35	TCTATTTCTT TTGGTTTAGG GTATCAAATA CAACACCTAG TAGAGCTTGT AAAAgAaGGT	4740
	GtCTCAACTT CTGAAATAGT TAAAAAGTTA AATCATTTAA GAGAAAACAT TAAATTATTT	4800
	GTAGTTATAG GGCAATTGAA TCAATTAATT AAAGGTGGCA GAATTAGTAA AACAAAAGGT	4860
40	TTGATTGGTA ATCTTATGAA AATTAAACCA ATTGGTACAC TAGATGATGG TCGCTTAGAG	4920
	CTTGTGcmCA ATGCGAGAAC TCaAAATTck AGTATCCAAT ACTTGAAAAA GGAAATTGCT	4980
	GAATTTATAG GAGATCATGA AATCAAATCC ATTGGTGTCTG CACATGCTAA CGTCATTGAA	5040
45	TATGTTGATA AATTGAAGAA AGTTTTTAAT GAAGCTTTTC ATGTGAATAA TTACGATATA	5100
	AATGTAAC TA CACCAGTTAT TTCTGCACAT ACTGGTCAAG GTGCGATTGG CCTCGTAGTC	5160
	CTTAAGAAGT AAATTTAATC TTTTCAGTGT TAATTACTTC CATTTCAATC CTTTATAGAC	5220
50	TAAATTTATA ATTAGATAGA TAGAGGAGGT AATTCATATG ACAAAGAAT ATGCAACATT	5280
	AGCAGGAGGA TGTTTCTGGT GCATGGTTAA ACCATTTACA TCATATCCAG GCATCAAGTC	5340
55		

GAATCAAACC GGCCATGTCG AAGCAGTACA AATTACGTTT GATCCAGAGG TTA CTTCTT 5460
 TGAAAATATA TTAGACATAT ATTTCAAAC ATTTGACCCA ACTGATGATC AAGGGCAATT 5520
 5 TTTCGATAGA GCGGAAAGCT ATCAACCAAGT CATTTTCTAT CATGATGAAC ATCAGAAAAA 5580
 GGCTGCTGAG TTTAAAAAGC AACAAATTAAG TGAACAAGGT ATTTTCAAGA AACCAGTGAT 5640
 10 TACACCTATT AAACCATATA AAAATTTCTA TCCAGCTGAA GACTACCATC AAGATTATTA 5700
 CAAAAAGAAC CCGGTACATT ATTACCAATA TCAACGTGGT TCAGGTAGAA AAGCGTTTAT 5760
 AGAATCACAT TGGGGGAATC AAAATGCTTA AAAAAGATAA AAGTGAATA ACAGATATAG 5820
 15 AATATATTGT TACACAAGAA AACGGCACTG AACCACCATT TATGAATGAA TATTGGAATC 5880
 ATTTTGCTAA AGGATTTATG TAGATAAAT TCnGGTAAAC CTTG 5924

(2) INFORMATION FOR SEQ ID NO: 131:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9280 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

GGCCGTTnAA AATCTCCAAA ATAnAAAAAC CCATCTTGTT CCAATGTTTT AAAATCGCCa 60
 30 TCCaACACTT GaTCaATAGC TTGCAACAAC GTTGAACGTG TTTTaCCAAA AGCATCaAAC 120
 GCTCCCACTA AAATCAGTGC TTCAAGTAAC TTTCTCGTTT TGA CTTCTCTT CGGTATACGT 180
 35 CTAGCAAAAT CAAAGAAATC TTTAAATTG CCGTTCTGAT AACGTTTCATC AACATCACT 240
 TTCACACTTT GATAACCAAC ACCTTTAATT GTACCAATTG ATAAATAAAT GCCTTCTTGG 300
 GAAGGTTTAT AAAACCAATG ACTTTCTGTTA ATGTTCTGGTG GCAATATAGT GATACCTTGT 360
 40 TTTTTTGCTT CTTCTATCAT TTGAGCAGTT TTCTTCTCAC TTCCAATAAC ATTACTTAAA 420
 ATATTTGCGT AAAAATAATT TGGATAATGG ACTTTTAAAA AGCTCATAAT GTATGCAATT 480
 TTAGAATAGC TGACAGCATG TGCTCTAGGA AAACCATAAT CAGCAAATTT CAGAATCAAA 540
 45 TCAAATATTT GCTTACTAAT GTCTTCGTGA TAACCATTTT GCTTTGCACC TTCTATAAAA 600
 TGTTGACGCT CACTTTCAAG AACAGCTCTA TTTTTTTTAC TCATTGCTCT TCTTAAATA 660
 TCCGCTTCAC CATAACTGAA GTTTGCAAAT GTGCTCGCTA TTTGCATAAT TTGCTCTTGA 720
 50 TAAATAATAA CACCGTAAGT ATTTTTTAAT ATAGTTCTA AATGCGGATG TAAATATTGA 780
 ACTTTGCTTG GATCATGTCT TCTTGTAATG TAAGTTGGAA TTTCTTCCAT TGGACCTGGT 840

55

	ACACTTCTTA CACCGTCAGA CTCTAATTGG AATATGCCAG TCGTATCTCC TTGCGACAAC	960
	AATTCAAACA CTTTTTGATC ATCAAACGGA ATCTTTTCGA TATCAATATT AATACCTAAA	1020
5	TCTTTTTTGA CTGTGTGTTAA GATTTGATGA ATAATCGATA AGTTTCTCAA CCCTAGAAAA	1080
	TCTATTTTTA ATAACCCAAT ACGTTCGGCT TCAGTCATTG TCCATTGCGT TAATAATCCT	1140
	STATCCCCCT TCGTTAAAGG GGCATATTCA TATAATGGAT GGTCAATTAAT AATAATTCCT	1200
10	GCCGCATGTG TAGATGTATG TCTTGGTAAA CCTTCTAACT TTTTACAAAT ACTGAACCAG	1260
	CGTTCATGTC GATGGTTTCG ATGTACAAAC TCTTTAAAAT CGTCAATTTG ATATGCTTCA	1320
	TCAAGTGTA TTTCTAATTT ATGTGGGATT AACTTTGAAA TTTCAATTTAA TGTAACTTCA	1380
15	TCAAACCCCA TAATTCTTCC AACATCTCTA GCAACTGCTC TTGCAAGCAG ATGACCGAAA	1440
	GTCACAATTC CAGATACATG TAGCTCGCCA TATTTTTCTT GGACGTAATG AATGACCCTT	1500
20	TCTCGGCGTG TATCTTCAAA GTCAATATCA ATATCAGGCA TTGTTACACG TTCTGGGTTT	1560
	AAAAAACGTT CAAATAATAG ATTGAATTTA ATAGGATCAA TCGTTGTAAT TCCCAATAAA	1620
	TAAGTGACCA GTGAGCCAGC TGAAGAACCA CGACCAGGAC CTACCATCAC ATCATTGCTT	1680
25	TTGCGATAAT GGATTAAATC ACTTACTATT AAGAAATAAT CTTCAAAACC CATATTAGTA	1740
	ATAACTTTAT ACTCATATTT CAATCGCTCT AAATAGACGT CATAATTAAG TTCTAATTTT	1800
	TTCAATTGTG TAACTAAGAC ACGCCACAAA TATTTTTTAG CTGATTCATC ATTAGGTGTC	1860
30	TCATATTGAG GAAGTAGAGA TTGATGATAT TTTAATTCTG CATCACACTT TTGAGCTATA	1920
	ACATCAACCT GCGTTAAATA TTCTTGGTTA ATATCTAATT GATTAATTTT CTTTTCAGTT	1980
	AAAAAATGTG CACCAAATC TTCTTGATCA TGAATTAAGT CTAATTTTGT ATTGTCTCTA	2040
35	ATAGCTGCTA ATGCAGAAAT CGTATCGGCA TCTTGACGTG TTTGGTAACA AACATTTTGA	2100
	ATCCAAACAT GTTTTCTACC TTGAATCGAA ATACTAAGGT GGTCCATATA TGTGTCATTA	2160
	TGGGTTTCAA ACACTTGTAC AATATCACGA TGTGATCAC CGACTTTTTT AAAAATGATA	2220
40	ATCATATTGT TAGAAAATCG TTTTAATAAT TCAAACGACA CATGTTCTAA TGCATTCAAT	2280
	TTTATTTCCG ATGATAGTTG ATACAAATCT TTTAATCCAT CATTATTTTT AGCTAGAACA	2340
45	ACTGTTTCGA CTGTATTTAA TCCATTGTGC ACATATATTG TCATACCAA AATCGGTTTA	2400
	ATGTTATTG CTATACATGC ATCATAAAT TTAGGAAAAC CATACAATAC ATTGGTGTCA	2460
	GTTATGGCAA GTGCATCAAC ATTTTCAGAC ACAGCAAGTC TTACgGCATC TTCTATTTTT	2520
50	AAGCTTGAAT TTAACAAATC ATAAGCCGTA TGAATATTTA AATATGCCAC CATGATTGAA	2580
	TGGCCCCCTT CTATTAGTTA AGTTTGTGTC GTAAAGCTGT AGCAAGTTGC TCAAATTCAT	2640

55

	CAATATCATT AATAATCAAT TGCCCTTTAG AACGTAATCG ACATCTGATT TCATTACCTT	2760
	CATCGACTGC AAATACCCAT ATTTTCAAGC CTTTGATGTC AGCAATTGTA TTAACAAACT	2820
5	GAGATGCTTC ATTTGGCTGA ATACCGAATT GCTCCAATAC ATCTTCAGTT ATTTTAACTT	2880
	GGCAGAATCC ATCATCCATA AGTTCGAAAT GTTGTAAAC ATAACCTTGA AACGGCAACA	2940
10	TTTTTGGGTC CTTCTCCATC ATTTTATTTA AAAGCGCATT ATGATCAATA TCATGCCCAA	3000
	TTAACTTTCC AGCAATTTCC ATAGTATGTT CTGAGGTATT GTTAAAAAGG AATCGCCCAG	3060
	TATCACCGAC GATACCAAGA TATAAACGC TCGCGATATC TTTATTAACA ATTGCTTCAT	3120
15	CATTAAATG TGAGATTAAA TCGTAAATGA TTTCACTTGT AGATGACGCG TTCGTATTAA	3180
	CTAAATTAAT ATCACCATAC TGATCAACTG CAGGATGATG ATCTATTTTA ATAAGTTTAC	3240
	GACCTGTACT ATAACGTTCA TCGTCAATTC GTGGAGCATT GGCAGTATCA CATACAATTA	3300
20	CAAGCGCATC TTGATATGTT TTATCATCAA TGTATCTAA CTCTCCAATA AAACCTAATG	3360
	ATGATTCCGC TTCACCCACT GCAAATACTT GCTTTTGGCG AAATTTCTGC TGAATATAGT	3420
	ATTTTAAACC AAGTTGTGAA CCATATGCAT CAGGATCTGG TCTAACATGT CTGTGTATAA	3480
25	TAATTGTATC GTTGTCTTCG ATACATTTCA TAATTTTATT CAAAGTACTA ATCATTTTCA	3540
	TACTCCCTTT TTTAGAAAAG TTGCTTAATT TAAGCATTAG TCTATATCAA AATATCTAAA	3600
	TTATAAAAAT TGTACTACC ATATTAACT ATTTGCCCGT TTTAATTATT TAGATATATA	3660
30	TATTTTCATA CTATTTAGTT CAGGGGCCCC AACACAGAGA AATTGGACCC CTAATTTCTA	3720
	CAAACAATGC AAGTTGGGGT GGGGCCCCAA CGTTTGTGCG AAATCTATCT TATGCCTATT	3780
35	TTCTCTGCTA AGTTCCTATA CTTCTGCAAA CATTTGGCAT ATCACGAGAG CGCTCGCTAC	3840
	TTTGTGTTTT TGAATATGCA TGTTCACCTC TATTTTGGCG AAGTTTCTTC CGACGTCTAG	3900
	TATGCCAAAG CGCACTGTTA TATGTGATTC AATAGGTACT GTTTTAATAT ACACGATATT	3960
40	TAAGTTCTCT ATCATGACAT TACCTTTTTT AAATTTACGC ATTTTCATATT GTATTGTTTC	4020
	TTCTATAATA CTTACAAATG CCGCTTTACT TACTGTTCCG TAATGATTGA TTAAAAGTGG	4080
	TGAAACTTCT ACTGTAATTC CATCTTGATT CATTTGTTATA TATTTGGCGA TTTGATCGTT	4140
45	AATTGTTTCA CCCATCTGAG GCTGTCTTCC TAAAAGTTGC ATAGACTTTA AAACATCTTG	4200
	TCTATTAATC ACACCCACTG TCTTTTTATT ACTCGAAACG ACAGGAATCA ATTCAATACC	4260
	TTCCCAAATC ATCATATGCG CACAACCTGC TACTGTACTC ATAGCATTTA CATAAATAGG	4320
50	ATTTGCGGTC ATCACTTTAT CTATTTGTC GTCGTCTTT GTATTAAATCA TCTCTCGACT	4380
	TGTTACAATA CCTACTAATT TATACGACTC ATTGACTACC GGAAATCTTG TATGGCCAGT	4440

55

ATCTAATGGC GTCATTATAT CTTGAACTAT TAAGATATCT TTTCGTATTT TCTGATTAAA 4560
 AAGTGCTTTG TTGATAATAT TTGCAACTAG GAATGTATCA TAACTTGATG ATAGAACAGG 4620
 5 TAAATCATGT TCATTTCGCA AATTAATAAC TTTATTAGAT GGCTTAAATC CACCAGTAAT 4680
 TAATATAGCC GTACCTCTTT TTAAGCTTC AATCTGCACA TCTTCACGAT TTCCGACAAT 4740
 CAATAATGTC TTTGGACCAA TATACTTTAA AATATCTTTG AGTTCCATTG CTCCAATTGC 4800
 10 AAATTTAGAT ACCATCTTAG TGATACCTTT GTTGCCACCT AACACTTGGC CATCAATAAT 4860
 ATTGACAATT TCATTAAAAG TTAAATGTTT AATTTTCAAT CGATTACGTT TTTCGATTCTG 4920
 AACCGTACCA ACACGATCTA TCGTTGCGAC CATGCCCAT TATCAGCAT CTTTmATTGc 4980
 15 ACGATATGCT GTCCCYtCaG ATACGTTTAA AAATTTAGCG ATTTTACGCA CCGAAATTTT 5040
 AGAGCCTATA GATAACGATT CAATATAATC TAAAATTTGT TCATGTTTTG TCATTCTTTA 5100
 CCTCTTCTTT TCGAACAGTA TTAACACAT TATAACTTTA TTTTGGATAA AAAGCATTGA 5160
 20 AGTGAATGA AATAATGATC GTTtCACCTA TTTTATTTTT TGAAAATATA CAACAAACAC 5220
 AAAGATCACA AAATCTTTAA TTTTAAATGG AAAAATCCAT TATTATTTAT TAGAATGTAA 5280
 25 GTGAGGAGGG ATGTACTAAT GTATAAAAT ATATTACTTG GTGTAGACAC TCAGTTAAAA 5340
 AATGAAAAAG CACTAAAAGA AGTGTCTAAA TTAGCTGGCG AAGGTACAGT CGTAACAGTT 5400
 TTAAACGCAA TCAGCGAACA AGaTGCTCAA GCATCAATTA AAGCAGGTGT TCATTTAAAC 5460
 30 AAACCTACTG AAGAACGAAG CAAGCGATTG GAAAAACAC GCAAAGCTTT AGAAGATTAT 5520
 GGTATTGATT ATGACCAAAT AATTGTTCGT GGTAATGCAA AAGAAGAACT ATTAAACAT 5580
 GCTAATAGCG GTAAATATGA AATTGTTGTT TTAAGTAACC GTAAAGCAGA AGACAAAAAG 5640
 35 AAATTTGTAC TTGGAAGTGT CAGCCACAAA GTAGCAAAAC GTGCGACTAT CCCTGTATTA 5700
 ATCGTTAAAT AAAATTTTTA TCCAGAATCA CAAATAATCT TTCAATCATG ATGCAGTCTC 5760
 AAACGACTGA GTAAATACAA GAAACGATTA TGACTGTGGT TCTGGATTTT TTATATCGTA 5820
 40 GTAAATTTAT AATCAATGTC TAATTGTATA AAACATAAAT TACGAGAGTA GGTCAGAAAT 5880
 GATAAAGAAC CACTGATGTC CCCCGTCCAC GTCGTAACGT AATCAGTAGA ATATAAAAAC 5940
 ACCCACTAAA AATATGCAGA CGATAACTTC CACATAGATT AGCGAGGTGT TTTTtagTGT 6000
 45 AAAATCTATA TTCTATTTAA AACTGAACAG ATTCACCTGG TTTTAAAATT TGCACGTCCC 6060
 CTACATTAAC AGCATCTTTA AATTGTTGTG GATCTTGTC GATTAATGGG AATGTATCAT 6120
 50 AATGAATCGG TACAGAAATT TTTGGTTTAA TAAATTCATT AATAGCATAA CTTGCATCAT 6180
 CAATACCCAT CGTAAAATTA TCTCCAATTG GTACAAAACA TACATCAACT GGATGACGTT 6240

55

TTCAACTTCA AACACGATAC CCATTGGCAT ACCTAAATAA ACTGGgAATA CCATTTTCAT 6360
 GTGTAAAAC TGAACATGA AATGCTTGAA CAAATTTAAC GCTTCCGAAA TCAAAGTTTG 6420
 5 CTTTACCACC AaTATTCATA CCATGAACAT TTTCAACACC GTGATATGAA GAAAGATAGT 6480
 CAGCCATTTC TGCACTTCCA ATTACTGTTG CTCCTGTTTT CTTTGCTAGT TCCACAACAT 6540
 CACCAAAATG ATCAAAATGA CCGTGCGTTA AAACGATATA GTCTACCTGC ACTGTTTCAA 6600
 10 TATTCAAATC AACTTAGGG TTATTTGAAA TAAACGGATC TACGATAACC TTTTGTGTGT 6660
 TCCCTTCTAA ATAAATCGTT GATTGACCAT GAAATGATAA CTTCAATTGA GCATCCTCCT 6720
 ATCAATTACT ATATAAATTT AGTACCCTTT TGCCACTTAA TTATAACAAA TTCTCAAATT 6780
 15 TTA AAAATG AAAATCTAGT TAATGTATTA GCTCGATTTT GAAATCTAAT AATAATTGGC 6840
 ATAAATGGA AGTAATATTA TGTGAGGAG TGTTTATAAA ATGACAAAAA TATCAAAAAT 6900
 20 AATAGACGAA TTGAACAATC AACAAGCTGA TGCAGCATGG ATTACAACAC CGTTGAATGT 6960
 ATATTATTTT ACTGGATACC GTAGCGAACC CCATGAAAGA TTATTGTCAT TATTGATTAA 7020
 GAAAGATGGT AAACAAGTAC TATTTTGTCC AAAATGGAA GTCGAAGAAG TCAAAGCATC 7080
 25 ACCTTTCACA GGTGAAATCG TTGGATATTT AGACACTGAA AACCTTTTTT CACTTTATCC 7140
 TCAAACAATC AATAAATTAC TAATTGAAAG CGAGCACTTA ACASTAGCAC GCCAAAAACA 7200
 ATTAATCTCT GGTTCATG TCAATTCATT CGGAGATGTT GATTTAACAA TCAAACAATT 7260
 30 GAGAAATATT AAATCCGAAG ATGAAATTAG CAAAATACGT AAAGCTGCTG AGTTAGCAGA 7320
 TAAGTGTATC GAAATAGGTG TTTCTTATTT AAAAGAAGGT GTGACTGAAT GTGAAGTAGT 7380
 CAACCATATT GAGCAAATA TCAAACAATA TGGCGTCAAT GAAATGAGTT TTGATACGAT 7440
 35 GGTTTTATTT GGAGATCATG CCGCATCACC TCATGGCACA CCAGGAGATC GCAGATTAAA 7500
 AAGCAATGAA TATGTACTAT TTGATTTAGG TGTAATTTAT GAGCATTATT GTAGCGATAT 7560
 GACACGTACT ATTAATTTG GTGAACCTAG CAAAGAAGCA CAAGAAATTT ATAATATTGT 7620
 40 ATTAGAAGCA GAAACATCTG CAATCCAAGC AATTAAACCT GGAATACCAT TAAAAGATAT 7680
 CGATCATATC GCTAGAAATA TTATTTGAGA AAAAGGTTAT GGTGAATATT TCCCTCATCG 7740
 CTTAGGTCAT GGCCTAGGAT TACAAGAACA TGAATATCAA GATGTTTCAA GTACTAATTC 7800
 45 TAATTTGTTA GAAGCTGGCA TGGTTATTAC AATCGAACCA GGTATTTATG TACCTGGTGT 7860
 TGCAGGTGTA AGAATTGAAG ATGACATACT TGTCACTAAT GAAGGATATG AAGTATTAAC 7920
 50 ACATTACGAA AAATAAGGAG TGGGATAAAA ATGAAAAGCT TGTACAAGC GCATTCTCAT 7980
 TCAGTCAAAC ACTGCCAATA TAACATTGTA GCGCCTAAGA CATAAATTTT TATCCAAGTC 8040

55

	TGTAATGAAT CAAATCAATA TCATTTCATGT TCGATGATTT CTTGCGATTG TTTCTAGCTT	8160
	TAATTTATCA TTATTTAATT TTAATAACCA AGGAGATGAT AACGTCATTG TTTAGTACGC	8220
5	TGTAATCCAT TCCCTTTTCA TCAAATTCAA ATTATAATTG TAATGCTTCT TCTACAGATT	8280
	TATATTCAT TTCAAATGCC TCTGCAACGC CTTTATTGGT TACGTGACCT TTGTAAGTAT	8340
	TTAAACCTAA TGATAATGGT TGATTTGATT TAAATGCTTC TCTATACCCT TTATTAGCTA	8400
10	GCATGAGCGC ATAAGGTAGC GTAGCATTAT TTAAAGCTAA CGTCGAAGTA CGCGGTACTG	8460
	CACCTGGCAT ATTTGCAACT GCATAATGAA CCACACCATG CTTAATATAT GTAGGATCAT	8520
	CATGTGTCGT AATTTTATCA GTTGTTCCTAA AAATACCGCC TTGATCAATA GCAATGTCAA	8580
15	TAATAACTGA CCCATTTTTC ATTTGTTTAA TCATGTCTTC TGTTACAAGT CTTGGCGCTT	8640
	TAGCACCTGG AATTAAACT GCACCTATTA CTAAATCACT TTGTTTAACTA TACAACTCAA	8700
	TATTCACCGG ATTTGACATA ATTGTATGTA CACGTCCACC GAATAAATCA TCTAATTGTT	8760
20	GTAAACGCTT TGGATTAACA TCTAAATCG TAACATCTGC ACCTAGTCCT AGTGCAATTT	8820
	TAGCTGCATT TGTTCCTGCT TGACCACCAC CGATAATAGT TACTTTACCC TTAGGTACTC	8880
25	CTGGGACACC ACCTAGTAGA ATTCCCATAC CACCATTAGG TTTTGTAGG AACTCTGCGC	8940
	CAACTTGAGC TGACATTCTT CCTGCTACCT CACTCATTTG TGATAACAAT GGTAAAGATC	9000
	GGTCTGGTAA CTGCACAGTC TCATATGCAA TACTAATTAC TTTTCTATCT ATCAAAGCTT	9060
30	GTGTTAATTT TTCTTCATTT GCTAAATGAa gatAaGTGAA TAATACAAGC CCTTCTTTAA	9120
	AATATGGATA TTCAGATTCA AGTGGTTCTT TAACTTTAAT AACCATATCC ACATCCCAAA	9180
	CTTTTGCTTG TTCAGCAACA ATCTCAGCAC CTGCTTCTTT GTAATCTACA TCTTCAAAGA	9240
35	ATGATCCTGA ACCCGcATTG GTTTCCACTA AAACAGTATG	9280

(2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:

- | | |
|----|-----------------------------|
| 40 | (A) LENGTH: 4669 base pairs |
| | (B) TYPE: nucleic acid |
| | (C) STRANDEDNESS: double |
| | (D) TOPOLOGY: linear |

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

	CTGATTAATC TCTTGTGTC GTGTATTAC TAATTGAATC GTTGGTGTCT GAACACGTCC	60
50	CAGGGATAGC TGTGCATCAT ACTTTGTTGT TAGTGACGCG GTTGCAATTAA TCCCAACAAT	120
	CCAATCTGCC TCACTTCTCG CTAACGCTGC ATAATACAAA TCGTTATATT GACGACCGTC	180

55

	ACGGATTGGC TTTTGTGTTAC CAACTTTATC CAAAATCAAT CTTGCAACTA GTTCACCTTC	300
	TCGTCCaGCA TCTGTTGCAA TAATAATATC TTTCACTTTA TTATCTAAAA TTAACGCTTT	360
5	TACTGTTTTA AATTGTTTGC TTGTTTTACC AATAACAACA GTTTTCATAT ATTTAGGTAT	420
	AATTGGAAGG TCTTCTAATC GCCATTCCCTT TAAATTTTTA TCGTATTGTT CAGGTGTGCG	480
	ATTTGTCACT AGATGACCTA ACGCCACAGT GACAATATAT TGGTTATTTT CAAAGTAAAC	540
10	ATTACGCTTC TGATTTATTT GTAAAGCATC AGCAATATCT CTTGCGACTG ATGGTTTTTC	600
	AGCTAATATT AAAGATTTCA TAAATTATCC TTTCTCATAC GTTCTTTTAT TTCGAACGTG	660
	CTTCATCTAT TCCACTAATC TTTGATTTAA ATTCAATGAT TGCAAATGAT GTGTTAAATG	720
15	TATTGTAACA TGTTAATATC ACTATTAACT TTCATTTTCTG TTGAAATACT ATATAATAAA	780
	AGTAACAAAA AGTACGGAGG TAATGACATG AGCATAGTTC AGTTATATGA TATTACACAA	840
20	ATAAAATCGT TCATTGAACA TTCGAATTAT GAATCAGCAT CATACTTATA TAACTTCCT	900
	CAACAGTACA ATGAAATAGA TGTATTAATA ACCGATGCGA TTGAATCACC TGGTGTATTT	960
	TCGATTAAAG AAAACGATTC AATCAAAGCA ATCATATTGT CTTTTGCATA CGATAAAAAAT	1020
25	AAATTCAAAG TCATAGGCCC TTTTCGTGGCT GACAATTATG TATTATCTGT CGATACGTTT	1080
	GAAACGCTAT TTAAAGCAAT GACTTCGAAC CAACCTGACG ATGCCGTCTT TAACTTTTCT	1140
	TTTGAAGAAG GCATTCAACA ATACAAACCA TTAATGAAAG TTATTCAAGC AAGTTATAAC	1200
30	TTCACTGACT ATTACATAGA AGCCCGTACA AGATTAGAAG AAGATATGCA CCAACCAAAT	1260
	ATCATTCCCTT ATCACAAAGG GTTTTATCGT GCTTTTACGA AATTACACAC AACTACATTT	1320
	AAATATCAGG CACAGTCACC ACAAGATATC ATTGATAGTT TAGACGACCA TCATCATTTG	1380
35	TTTTTATTTG TTAGCGAAGG TTTACTTAAA GGTATTATTAT ACCTTGAAAT TGATTACAA	1440
	CAGTCAATCG CCGAGATTAA ATACTTCAGT TCTCATGTAG ATTACCGTTT GAAAGGTATC	1500
	GCTTTTCGAGT TGCTTGCCTA TGCATTGCAA TATGCTTTTG ATAATTTTGA TATTAGAAAA	1560
40	GTTTATTTTA AAATTCGTAA TAAAAATAAT AAACATCATG AACGATTTAA TGGTCTAGGT	1620
	TTCCATATCA ACTATGAGTA CATTAAATTC AAATTCGAAT CACGTAACGT AAAAGATCAA	1680
45	ACAATCCCTG AATAAAACAC CAAGCAAATA CCCTACAGTA CATCATTAGC ATGTATTGTG	1740
	GGTTTTTCTA CTTTTTGTA ATATTGAAAA TTATAAGTAG TTGTTTTTTA CTATTAGGGC	1800
	AGAATGCTTT ACAATAACAT GCAAGTGTCA ATTAAGGGGA GCACTTGCAT AAATAGTATA	1860
50	GGAGAGTGAG TAGTCTTGCA ATTTCTTGAT TTCTTAATCG CACTTTTACC TGCTTTATTC	1920
	TGGGGAAGTG TCGTTCTTAT TAATGTGTTT GTCGGCGGTG GACCTTACAA CCAAATTCGT	1980

55

	TTCAATAATC	CTACTGTAAT	TATTGTCGGT	CTTATTTCTG	GTGCATTATG	GGCGTTTGGA	2100
	CAAGCGAATC	AGCTTAAATC	TATTAGTTTA	ATCGGTGTAT	CAAATACTAT	GCCAGTTTCT	2160
5	ACAGGTATGC	AATTAGTTGG	TACAACATTA	TTCAGCGTTA	TCTTTTTAGG	TGAATGGTCT	2220
	TCAATGACTC	AAATTATCTT	TGGTTTAAATC	GCCATGATAT	TATTAGTTAC	TGGTGTAGCA	2280
	CTTACTTCAC	TTAAAGCTAA	AAATGAACGT	CAATCAGATA	ATCCTGAATT	TAAAAAAGCA	2340
10	ATGGGTATTT	TAATTGTATC	TACAGTTGGA	TATGTAGGTT	TCGTTGTACT	TGGTGACATC	2400
	TTTGGTGTG	GTGGAAGTGA	TGCATTGTTT	TTCCAATCTG	TCGGTATGGC	AATTGGTGGC	2460
	TTTATCCTAT	CCATGAATCA	TAAAACATCA	CTTAAATCAA	CAGCACTTAA	TCTATTGcCA	2520
15	GGTGTGATTT	GGGGAATTGG	TAACCTGTTT	ATGTTCTATT	CTCAACCAAA	AGTTGGTGTA	2580
	GCTACAAGTT	TCTCATTATC	ACAGTTACTT	GTTATCGTTT	CAACCTTAGG	CGGTATTTTC	2640
20	ATTTTAGGAG	AAAGAAAAGA	TCGTCGTCAG	ATGACGGGTA	TTTGGGCAGG	TATTATTATT	2700
	ATCGTGATAG	CTGCTATAAT	TCTAGGTAAT	TTGAAATAGA	AAGTTAAATA	CTCATGTAAC	2760
	GTAAAAATGT	AATCACTTCT	GAAAATAACC	ATTCACTTAT	AGAATGATTA	AAATTAATTT	2820
25	TCGGGAATTT	TACGTTGAAT	GTTCCTCTAT	ATGTCCTAGG	AAATACGTGG	CTCTAAAAAC	2880
	AAAACGCAAT	AACACATCAT	GACATTAATC	ATGCGTTTTA	AGACTTTAAA	ATTAGCGATA	2940
	CTTTTAAAT	CTTGATGATA	TTCATATATC	AAGTATGCGC	CATACATATG	AAGTGGATAG	3000
30	CTGCATAACG	CACTGCATTA	TCAAATTGAA	TGTATGAGTT	GAACAACAT	ATCATAAATA	3060
	AAAGCCCCCT	TTTCACAATA	TACATTTACA	TATTGTGGTA	AAGGGGGCTC	TCATTTTCTA	3120
	CGAATACTAA	AATGGATTTT	ATTTTCAAAT	GTGTAACTA	GACAAACACT	GCCTGATACA	3180
35	CGTACAAAAT	AATGATACTA	ATAATGATTG	TCAAATTGGT	CGTCATACCT	ATAAATGGCA	3240
	GTGTTGCGATA	TTTAAACTGA	ATACCATAAG	AAATAATTGC	AACACCtACC	GGGAACATCC	3300
40	AAGTGACCAA	CAATGTCGTC	TTAATCATAT	CATCTGATAC	TGGTAACAAAC	ACATATACTA	3360
	ACAATCCCGC	AACTAATGCT	AATCCATAAT	GCAAACATAA	ATATTTAATA	GTAGCAGGTA	3420
	TATACTTTCT	TTCCAGAGTA	AAATTCAACA	TGACACCTAG	CAAAATCATT	GATAACGGCA	3480
45	TATTTGCATG	GGAAAGTATG	CTAAAGAAAT	CGATTGCCAC	ATGTGGTAAA	TGGATGTGAC	3540
	TTATATTCAA	TATAAACATT	ACAATGTATG	TAACGAGTGG	CACTGATTGT	AATAATTTCT	3600
	TACCTAAATA	TTTAAATCG	AATTGATCAC	TACCTTCACT	AAAGTAGCTA	CCTACAAAGT	3660
50	AAGTAATTCC	AAACATCACA	AAGGCACCAC	CTATATCAGC	CATAACAAAA	TAAATAAGTC	3720
	CCGTTTTAGG	CCATATCACT	TCAATTAGTG	GATATGCAAA	CAATCCAATA	TTCATAGCAC	3780

55

CAATCATT TT CGCCACAATA CCATATATAA TCATTAAAAAT TGGTAAAATG GAGAATGACA 3900
 ATTTTAATTC TGCAC TGT T AAATTCACAA TAACTAAAGA TGGGAGTGTG ACATTAAGAA 3960
 5 CTAATGTAGC AATGACTTGA CTATCTGTTG CTTTATAAA ATTAATGCGC TTCAAAAAGT 4020
 AACCAAGCGC AATTAATAAA ATAATCATAG TAAATTGTTT TGTCAC TGT ATCCCTTCTT 4080
 TCAATAATCT TCATAATTTA TAACTTTAAC ATACTCCACA GATATTTTAG AAGTCTACTG 4140
 10 TTTTCATGCTA TAATCTACAT TAAATGCACT TAATTATATT TCAAAGGAGT GTTATAGTAT 4200
 GTCTTTAGAA AACCAACTAG CCGAACTTAA ATATGATTAT GTTCGTCTTC AAGGTGACAT 4260
 AGAAAAACGG GAATCTTTGA ATTTAGATAC TTCCGCACTT GTTCGTCAAC TTAAAGATAT 4320
 15 TGAAATGAA ATTAGAAACG TTCGTGCTCA AATGCAAGAT TAATAATCTA TCATTCAAGC 4380
 AATAATGCT TTTTGTTACA TAAATTTGAC TAGCATTGCT CTGAATACGT TATATTGATG 4440
 AATTGCTTCA TTTTTCGCTC AATTACATCT AGAATCACAA GATGTTGTCG TGTATGATT 4500
 20 TAGTGTTC TTAACAACAT ACACGCATAT CTATCCCAAC ACTGCTATTT ATGTTTTCTA 4560
 CGCTGnTGTA CTACATGAAC CCTTTGAAAC GGAGAGGAAG TTATCATATG CAATTTTAnC 4620
 25 TGATTTTACT AGCAATACTT TAACnAATTG nTAGTTTAAAT AGAATTTTA 4669

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2785 base pairs
 30 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

TTTGCACCCA TCTGaTACAA TGCACCATGC GGTTTAACAT GATTAAATTTT AACTTGATGA 60
 ATGCGACAAA ACCCTTGTA TGCACCTAAT TGATAAATCA TCAAATTATA AATCTCGTCG 120
 40 TTAGAGATAT CTATATTTTCG TCTGCCAAAG CCTTTCAAAAT CAGGTAAACC AGGATGTGCA 180
 CCTACTGCAA CATTATGTGC TTGGCAAGT TTTACCGTTT CATTCAATTAC ATTTTCATCA 240
 CCAGCGTGAA AACCACAAGC AACATTGCGA CTTGTAATTA ACGGAATAAT TTGATGATCA 300
 45 CCACCAAAGG AATAATTTCC AAATGCTTCG CCTAAATCAC AATTCAAATC AACTCGCATT 360
 ATAATTCAC CCCTTTAACA ATTTGATGTT TTTCTAAAAA TTAAATATCA ACATCTTTTG 420
 50 CATCTCCATC ACGATATAGT GGATAATTTA AAAC TGCATA TAAAAAATCG GCAGTTGTAG 480
 AAAATCCATC TATCACCATT TCATCTAAGG TGACTTTCAA CTTATCAATT GCTGAAGCTC 540

55

	AACCGTGATA TAGTAAAGAA TCGACTCGCA CATTAAAGCC TTGAGGTAAA TGTAACGCTG	660
	TCACTTTACC TGGTGTGGT TGAAATTTCT TTTCaGGATT TTCGGCATT ATTCTCGCTT	720
5	CTATCACATG ACCATTAAAT TGAATATCGC TTTGTGAAAA AGGTAAATGA TTATGTTCCA	780
	ATAAATACAG TTGTGCTGCA ACCAAATCAC GTTCTGCTCG CATCTCTGTA ACAGTATGTT	840
	CAACTTGAT TCGAGCATT ATTTCAATAA AGTAATGTGC GGTATCAGTT ACTAAAAATT	900
10	CAATCGTACC TGCACCTCTA TAATTTGCTG CACGTGCAAC TTTAACAGCA TCGTTACATA	960
	TTTGTGTGCG TCTTTCTTCA GTTAATGCTG CACAAGGAGA TTCTTCGATT AATTTTGTAT	1020
	TTTTACGTTG TACAGAACAA TCACGTTCCC CTAAATGTAC ATAATTATCC TGCCCATCTC	1080
15	CCaTAACTTG AACTTCAACA TGTTTTGcAA CAGGTATAAA AGCCTCAACA TAAACACGAT	1140
	CATCATCAAA GTATTTTTTT CCTTCACTTT TAGCTTCTTT AAATGCCTTT TCTAAATCTT	1200
20	CAGCTTTCTT TACAATACGT ATACCTTTAC CACCACCGCC ACTGGCAGCT TTGATAACAA	1260
	CTGGATAACC GATGTCTTTG GCAAGATTCT CAATTCAGA CACATGATTC ACAGCACCAT	1320
	TTGATCCTGG AATCACAGGA ACACCTGCAT GATGAACTGT TTGTCTTGCT GTTATTTTAT	1380
25	CCCCCATCAT TTCCATCGTT TTTTGTAGTAG GCCCTATAAA CGCTATGCCT TGTTCCCTCA	1440
	CGGTTTGAGC AAATTTTGTT GATTCTGATA AAAAGCCATA TCCTGGGTGA ATTGCATTAG	1500
	CACCAGTGAT TTGTGCAGCA GATATGATGC GGTCAATATT TAAATAACTA TCTAAAGCAT	1560
30	TATcwtCCCC AATACATATA GCTTGATCTG CTAAATGTAC ATGCAAGCTT TGCTCGTCCC	1620
	CTTTTGATA AACTGCTACA GTTCAATCC CATATTCTCT GCAAGCTCTT ATAATCCTTA	1680
	CAGCAATTC ACCTCTGTTG GCAATTAAAC AACGAAGCAT TTACTTACCC CCTTTACTTA	1740
35	ATACGTACCA AAACCTGGTC GTATTCAACA TTTGTGCCAT GATCAGCTAC TATTTAGTA	1800
	ATTCTCCAG CAACATCTGT TGTACCTCG TTTAATACTT TCATCGCTTC AACATATCCT	1860
	ATAATATCTC CCTTGTTAAC TTTGTCACCG ACATTACAA TTGGTTCAGT TAATTCTTTA	1920
40	CTATCTTGTA AAAAGAATGT ACCTATCATT GGTGATTTAA TGTATGATA ATCATTTGTC	1980
	GAAACATCGG AGTTATCATT CGCTTTTGAA GCTGTCAAAT CATTATTGTT CATACTTTGA	2040
45	TTTGATTGAT TACTGTGTGC AGCCAAATGA TTCGAGTCAG TGAAGTCAAT TTCTATTTC	2100
	TCTTCAAAAT TTTTATATTT AAATTTCTTA ACATCATTTT CCTTCACTAA TTTGATTATT	2160
	TGTTGATTT nTTCAATATT CATTTTACAA ATCCCCTTTT AAAATTGTTG CTAATTTTTT	2220
50	CGAAGTATGT CGCAAGCTAG ATGTATCAAA AATTGGAGTC TTTTGATGAC TCTTAAGAAT	2280
	TTCATTAAC AGAGACATTT GTCCCGATT CTTATCTACA GCTTCTTGA ATGATATCCA	2340

TACAGTTGCA ATTTTGGTAT AACCACCTAT CGTTTGTITA TCATTAAGCA GAATAATAGG 2460
 TTGACCATCA TTTGGTACCT GAACACTACC AAGAGCAACC GGTTCAGAAA TGATATCTGC 2520
 5 TTGATTAaCT GGTGCAACGC TGTCACCTTC CAAACGATAG CCCATACGGT CTGATTGTTC 2580
 AGTAATTAAA TATGGATGAT TTACAATTTT CGCTCTAGCC TCTTCAGAAA ATGCCTCGAA 2640
 TTGAGGTCCT TGAAGAATGT GTATAATATT ATTTTCTGGC AATAAATCGT CCTGTAAATG 2700
 10 AATCGTCTTT CCAATGTTTT CTTTAAAGTC ATTATTTATT TTCACTGTTA TTACATCATC 2760
 AGCTAATAAC TTTCTACCTT TGAAT 2785

(2) INFORMATION FOR SEQ ID NO: 134:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1010 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

25 AATGGAACCG GTTGAAACAG CAATTATTAC TATTTCTATG GGTGAAGGTA TTTCAGAGAT 60
 ATTTAAATCA ATGGGTGCCA CACATATCAT TAGTGGTGGA CAAACGATGA ATCCTTCTAC 120
 AGAAGATATC GTTAAAGTCA TTGAACAATC AAAATGTAAA CGTGCAATTA TTTTACCGAA 180
 30 TAATAAAAAT ATCTTAATGG CAAGTGAACA AGCAGCGAGT ATTGTTGATG CAGAAGCTGT 240
 TGTTATTCCA ACGAAATCTA TTCCTCAAGG TATAAGCGCA CTATTCCAAT ATGATGTGGA 300
 CGCAACACTT GAAGAAAATA AAGCGCAAAT GGCTGATTCA GTAAATAACG TTAAATCTGG 360
 35 TTCAITTAACG TACGCTGTTC GTGATACGAA AATTGATGGC GTTGAGATTA AAAAAGACGC 420
 GTTTATGGGC TTGATTGAAG ATAAGATTGT AAGCAGCCAA AGTGATCAAT TAACAACGGT 480
 TACTGAGTTG TTAAATGAGA TGTTAGCAGA AGATAGTGAA ATATTGACTG TGATTATTGG 540
 40 TCAAGATGCA GAGCAAGCAG TTACAGATAA CATGATAAAC TGGATCGAAG AGCAATATCC 600
 AGATGTAGAA GTGGAAGTTC ATGAAGGTGG ACAACCAATT TATCAATATT TCTTTTCAGT 660
 45 AGAATAAAAA TTTAAAATAA AAAACTACCA ATGATAAATC ATCAGTTGGT AGTTTTTTAT 720
 TTTGCTATTT TAGTGATATT GCGGGTTAAA AGTATCGTTC TCGAGTTGCT AACAAATGTCA 780
 TGTTCAACTT AGTCATGATA AAATAAATAA CATACTAAAT GATACGTAAA ATCAAATAAA 840
 50 ACATAGGTGA TTTATTTTGG CTAAAGTAAA CTTAATAGAA AGTCCATATT CTCTTTTACA 900
 ATTAAAGGT ATAGGTCCTA AGAAAATAGA AGTATTGCAA CAACTAAATA TTCATACAGT 960

55

(2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1540 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

	TGTAGTTGAA CATGAACAAC AAAAGAAAGA AAAGACAAAA AAGCAATACA AGCCATTTTG	60
	GATTGTCATG AGTTTTATAA TACTTATAGT TGTACTATTA CTCCCGGCAC CTTCAAGTCT	120
	GCCGATAATG GCTAAGGCAG TACTAGCTAT TTWAGCTTTT GCAGTTATTA TGTGGGTAAC	180
	GGAAGCTGTA TCATATCCGG TGTCAGCAAC TTTAATTATT GGCTTAATGA TATTACTTTT	240
	AGGATTTAGC CCTGTTCAAA ATTTAGGGGA GAAGCTAGGT AATCCGAAAA GTGGCAGTGC	300
	TATTTTAGCT GGAAGTGACC TTCTAGGAAC TAATCATGCA TTATCATTAG CGTTTAGTGG	360
	ATTTGCAACT TCAGCTGTAG CTCTCGTTGC AGCTGCATTA TTTTGGCTG CTGCTATGCA	420
	AGAAACGAAT TTGCATAAAA GACTAGCTCT TTTAGTGTTA TCAATTGTTG GTAATAAAAC	480
	TAGAAATATA GTTATTGGAG CAATTATCGT TTCAATTGTA CTGCAATTTT TCGTTCCTTC	540
	TGCAACAGCT AGAGCAGGGG CAGTTGTACC AATCTTGCTG GGTATGATTG CGGCATTTAA	600
	AGTTTCCAAA GATAGCAAGT TAGCGTCTTT ATTAATAATT ACTTCAGTAC AAGCTGTGTC	660
	AATTTGGAAT ATTGGTATCA AAACGGCGGC AGCACAAAAT ATCGTAGCGA TTAATTTTAT	720
	AAACCATCAA TTAGGATTG ATGTTTCATG GGGCGAGTGG TTCTTATATG CAGCGCCTTG	780
	GTCCATAGTT ATGTCCGTAG CTTTATATTT CATCATGATT AAAGTGATGC CTCCAGAAAT	840
	TAATACAATA GAAGGTGGTA AAGATTTAAT AAAAGAAGAA TTGCATAAAC TTGGCCCCGT	900
	TAGCCACGTT GAATGGCGTT TAATTGTTAT ATCGATGTTA TTATTACTGT TTTGGTCAAC	960
	TGAAAAAGTA TTACATCCGA TTGACTCTGC ATCCATTACT ATTATTGCTT TAGGTGTTAT	1020
	GTTAATGCCG AAAATTGGTG TCATGACATG GAAACATGTT GAAAATAAAA TACCATGGGG	1080
	AACAATTATC GTGTTTGGTG TAGGTATTTT ACTAGGTAAC GTTCTTTTGA AAACAGGTGC	1140
	AGCTCAATGG TTAAGTGATC AAACTTTGGG TGTTTTAGGT TTAAACATT TACCTATTAT	1200
	CGCGACAATT GCACTTATCA CGCTTTTAA TATATTGATT CATTTGGGCT TTGCGAGTGC	1260
	AACAAGTTTA TCATCAGCGT TAATACCTGT TTTTATTTTG CTAACCTCTA CGTTACACTT	1320
	AGGAGACCAG TCTATAGGAT TTGTTTAAAT TCAACAATTT GTTATTAGTT TTGGTTTCTT	1380

AGATTCTTG AAGGCAGGTA TACCATTGAC AATTGTAGGG aATAtctAgT GaTAGTTTTT 1500

AGCATGACTT ATTGGAAATG GGTAAGGTTG CnTTAATTAA 1540

5 (2) INFORMATION FOR SEQ ID NO: 136:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11823 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

15 ACTTCTCACA ATAAGAAATA TGAAATTGTT ATGTGTTAGT TGAGATTCAG TGATGAATTA 60
 CTTTATCAT TTAATGTT GTTATCATTG TCATGCGTTA CCAAATCGCT TACGTATACA 120
 20 CGATTCCCAA TCTTAACATA GACGATTGT ATATCAGAAT TTTCTGATTA CTAACAGTTT 180
 ACCTAAGTTT AAATATCTGT TCAATGATTT TCAGTTATTT TAAAAGAAA AATCGTAATG 240
 CTGCCATGAT AACAATCCCA CTAATAATTG TAATAGTTAA ATACGCGTGA TTATAGATAA 300
 25 AATAACCGTC GGAATGAGCG CGATAATGTA AGGGATGTTT AATGTATACC CCTCACCATG 360
 AGGCGTCTGT TGAATAATGC TGTCAATGAC AAGTGCCGTA AATAGTGTGA TTGGGATAAA 420
 TGATAGCCAT CGAACCACGA CATCAGGCAA TTGCACTTTT GAAATCATGA TAAAAGGTAT 480
 30 AATTCGAATT AATAGCGTTA CGATACCACA CAATAAAATA AGTATTAACA TGTTCATATG 540
 AGTTATCATT GTTCCATCAT CACTCCTAAC GCTGCTGAAA TTGTGGCTGC AATTAATATT 600
 GCTAGATATG AAGGCATAAA CATACTTAGC GATAACATCA TTACTATGAC GGCAATAATG 660
 35 AGTACTATGT AAATCTTAA TCGCGATTTA GTAATTGATT CAAATTGCGC AATGGCCAAA 720
 AAGATAAACA TAGCCGTGAT AGCAAAATCT AACCTAGCG TTGCGGATT TGAGATATAT 780
 TCGCCAAATA AAGCCCCAGC TACACATGAA ATTGCCCAA ATAAATATGC TGTGATGTTA 840
 40 AGACCATGCA TCCAACGATC ATTGATAGCT TCTCCTTTTA AATAAGGTGT AATGGCGACG 900
 CCAAACGTTT CGTCAGTTAC TAATGAACCT AATCCAACAC GGTTCACAAA CCCATATGTC 960
 45 TTGAAGTTTG GTGCAAGCGA CATACTTAAA AGGAACATTC TTGAATTTAC GATAAATACA 1020
 GTTAGTACAA TCGCTGATAT AGGTGTACCT GCTATAAACA ACGCGCACAT AATAAATTGC 1080
 GCagcaCCGG CATATATAAC AAGACATAAC AAGACAATTT CTAAAATACT AAAGTTTTGA 1140
 50 GACGAAGCCA CAATACCAA TGAAATACCA ACACCGGCAT AACCCAATAA TGTTGGGATA 1200
 CACTCTTGCA CGCCTTGTCT AAAACTTAAA TGTGTTGTCA TCTCAATTAC CTCCTTTGCC 1260

55

	TAAGCAATAA CATTAGACAT CAGTTTGTCT GAGGTTAGAC ATTCCGGAGT CTTTAGTCAG	1380
	CTTCATATTA ACTTTTTATT TTTGAGAATT TTCAATTTTT TATTAAAGAC TACCTCCATA	1440
5	TTTTCTATGG aTTTGTAGTT GTTTTTAAGT ATCAATTTTA TAAATTTTAA TATCTGATGA	1500
	TGAGTCTGGG aTATTGaTTC ATGTACCACT CCCTTaTaAT CATCCCCTCC CCCTaCCCTA	1560
	CTCCATCGAT ATAACTCATA CTACATATCA ACGAAATCAG TATTTTATCG CTTCCCTTCC	1620
10	TATATTAGTG ATGCTCAAAC TTGTTACGTT TTAGATTGTT TTAGTTCATC ATAATTATCC	1680
	CGTATTGTTG CTATAATGAA ATGCGTTCAC CCCATTAAAC CACAAACTTA ATTTATTGTT	1740
	GTTATGTGCA TTGGCTCACT ATTATATTTT TACAGCACAA AAAAAGTGGC GACAGTTCGT	1800
15	CACCACTTTT TAAAATATTA TTTAAAGTAT CTTGCCCTTG CTTTAAGTAT ACGTAGATAT	1860
	ATACTTTTTTA AAGCTTGTAG CTAAAGCCTT TATTTAACTG GTTTTGAAAT TTGTGTTTTA	1920
	CCACCCATAA ATGGTACTAA TGCTTCTGGA ATTGTTACTG TTCCATCTTC ATTTTGGTAA	1980
20	TTTTCAACAA TAGCAGCAA TGTACGTCCA ACTGCTAAAC CACTACCATT TAATGTATGT	2040
	GCTAATTCTG GTTTAGCTGC TTTGTCACGC TTGAAGCGGA TGTTAGCAGC ACGCGCTTGG	2100
	AAATCCGTAC AGTTTGAGCA TGAACAAAT TCTTTATAAT CATTGTAGCT TGGTAACCAA	2160
25	ACTTCTAAAT CATATGTTTT GCTTGCACTA AATCCAATAT CACCTGTACA TAAAATAACA	2220
	CGACGGTATG GTAAACCTAA CTCTTCTAGA ATTGCTTCTG CGTTTGTTGT CATTTCTTCT	2280
30	AAAGCATTCC ATGAATCTTC AGGTTGTTCA AAACGTACCA TTTCCACTTT ATCGAATTGA	2340
	TGTAAACGAA TTAATCCTCT TGTATCTCTA CCTGCTGATC CTGCTTCACT ACGGAAACAT	2400
	GCAGATTGAC CAGTGAATTT TTCAGGAAGT ACACCTGGTT GAATAATTTT ATTACGGTAG	2460
35	AAATTCGTTA ATGGTACTTC AGCAGTTGGA ATTGTATATA ATCCTTCTTT TTCTACTTTA	2520
	AATAAATCTT CTTCAAATTT AGGTAATTGA CCTGTACCAT ACATTGTATC TGC GTTCACA	2580
	AGCTGTGGTA CCATCATTTT TGTATAACCA TGTGTGTTG TATGTTTTGT AATCATATAG	2640
40	TTCATTAAAG CACGCTCTAA TTGCGCACCT TCATTTGTTA AATATACAAA ACGCGCACCT	2700
	GAAACTTTTG CTGCACGATC AAAATCAGCC ATTTTCAATT CTTCTACAAT ATCCCAATGT	2760
45	GCTTTGGGTT CAAATGAAAA CTCaCGTGGT GTACCCCACT TTTTAACTTC AACGTTATCT	2820
	TCATCAGATT CACCTTGAGG TACATCATCA CTTATTAAAT TTGGAATACG ACAAAGGATA	2880
	CCTGTCATTT TATTATCAAT TTCATTTAAT TGACTATCTT TTTCTTTAAT ATCGTCACCT	2940
50	AATGTGCGCA TTTCAGCAAT CACATCATCA GCATTTTCTT TATTACGTTT TTTTAAATGCG	3000
	ATTCTTTCGC TTACTTTATT ACGACGTGCT TTCATTTCTT CTGTTGCACT AATTAATTTA	3060

55

	TCAATTTTGC TCTTAAGTGT GTCAGGCTCA TTTCTGAATA ATCTAATGTC TAACATTAAAC	3180
	CTTCATCCTT TCCCAAATAA TTATCATTTA TTATGGAATG ACGTACGTCT TTATTTTTTA	3240
5	GAAAATAAAA AAAGACCACA TCCCTACAAG GGACGTGGTC TACGCGTTGC CACCCTATTT	3300
	AACAATTTAA GTTATAAAGA TACACTAAAC CTAAATTGCA CTTCACTAAA ATAACGGTTA	3360
10	TCACCGATTG TTCTTTTAAA TTAAGTAGGT AGATTCAATAT ATATGTTGAT TCTTGTTTAC	3420
	ACTAACCACA AGCTCTCTGA TATCGAACAC TATATATTAC TTGTCCTACG AACAAATGTCT	3480
	TATTAAGTTA TTTTAAATAT AGCAAACTAT ATTTGCTTTT TCAAGTAACG ATTTCAAACA	3540
15	TCACTCATGT CGATTTAGTG ACATGCAGTC GTTTGATAAA TTGATTGCTT TAAATACTGT	3600
	GCAACCGCTT CAATATCTTT ATGAAATTGA CGATCATGTG TAATGGATGG CACGATACTT	3660
	CGAAACTCAT CATACTTGCG ACGTGTPTTT GGTGATAATC CTTCAACACC TTTTAACTCT	3720
20	GCTGCTTGTA ATGCAATAAC ACATTGCGATT GCCAGCACAC GTCTTGCATT TTCAATAATT	3780
	TGATAACCAT GTCTAGCAGC TGTAGTCCC ATAGATACGT GATCTTCTTG GTTCGCAGAT	3840
	GAAGTGATAG AATCAACACT CGCTGGATGC GCTAAAGTTT TATTTTCAGA AACGAGACTT	3900
25	GCAGCAGCAT ATTGCATAAT CATCGCGCCA CTTTGCAATC CTGGCTCTGG ACTAAGAAAT	3960
	GCTGGTAAAT CACCATTAA TTGAGGATTT ACTAGTCGCT CTAGACGACG TTCCGATACG	4020
	TTTGCTAATT CACTTACACC TAATTTAAGA TGATCTAATG CAAAAGCAAT AGGTTGTCCA	4080
30	TGGAAGTTAC CACCTGAAAT AACAAACGTT TCATTGCTT CCTCAAATAT AAGTGGATTA	4140
	TCATTAGCCG CATTCAATTC AAATTCTAAT TGCTGTTTAA CATAATTGAA TACTTGAAAA	4200
	CTCGCGCCAT GGATTTGTGG TATACAACGC AACGTATATG CATCTTGATC ACGTATTTCT	4260
35	GATTGTGCGC TCGTTAATGT TGATCCTTCT AACCAATCAC GCATACGCGC TGCCACATTA	4320
	ATCTGTTCTT GAAAATTACG AACTGCGTGC ACATCATGTC GATATGCATC TATAATGCCA	4380
	TTAAGAGACT GATGCGTTAA TGCAGCAATC CATTGAGATT GGTAACCTAA ATCTTCTGCT	4440
40	TCTATATAAC TAATGACACC TTGAGCTGTC ATAGCTTGCG TACCATTAAAT CAATGCTAAA	4500
	CCTTCTTTAG CCTGAAGGTT CAAAGGTTGT CTATTTAATT CTCTTAATAC ATCGTCACTA	4560
45	TCCTTTTCTT CCCCTCTGTA CAATACTTTC CCTTCACCAA TTAATGCTAA TGCTAAATGT	4620
	GATAATGGCG CTAAATCTCC TGATGCACCG AGAGAGCCTT GCTGTGGGAT TATCGGTATA	4680
	ATACGTTTAT TTATAAAAAA TTGTAATTGT CTCCTAATT CTAAAGTGGC ACCTGAATGA	4740
50	CCTTTTAATA ATGTATTCAA TCGTAAATC ATCATGACTA ATGCTACTTC TTTTGAAAT	4800
	GGCTCACCTA GTCCACAGGC ATGTGAGCGT ATCAGATTCA CTTGTAATTC ATTATATTGC	4860

55

	TCCTCATT	TTT	CAATAATACG	TTCAACTACC	GCTCTACTTT	TTTTGACACG	TTCTAACGCA	4980
	TCATCAATAA	TTT	CAATCTT	TGATTGTTGT	TGTAAAAATG	ATTTAATATC	CTCAATTGTT	5040
5	AGTGT	TTT	CAC	CATCTAAATA	TAAAGTCATA	TATGTTACCC	CCTTGTTTAT	5100
	CCATCCTTCT	TGAAGTATAC	GTTTTCA	TTT	TATTGAAAC	AATGGTTTTA	CGTACATT	5160
	TAACCTATTA	TCAGAGCACT	ATTG	TAGTGC	GTTAAAGGAT	ATTAAGATTG	TTGTAAGCAT	5220
10	ATTTAATAAT	TTATCTATTG	ACGAATTGCA	TATACAGGTA	TAGTATTTTC	TATTGTATTT		5280
	AACGACAAAT	AATAATGAAT	TCAGAAATTT	ATAATACATT	TTGTTAAAAG	TTACTATATA		5340
	TTTTTAAAAT	TGAATAAATT	CGGAAAAGGC	TTTTACATGG	GAGGTTATAT	CACTATGGAA		5400
15	ACGT	TAAATT	CTATTAACAT	TCCTAAGCGT	AAAGAAGATT	CACATAAAGG	TGATTATGGC	5460
	AAAATTTTAT	TAATTGGTGG	ATCTGCTAAC	TTAGGTGGTG	CCATTATGTT	AGCGGCTCGT		5520
20	GCATGTGTAT	TTAGCGGTAG	TGGTTTAAATC	ACTGTAGCTA	CACATCCAAC	AAATCATTCA		5580
	GCATTACATT	CTCGTTGCC	AGAAGCGATG	GTTATTGATA	TTAATGATAC	GAAAATGTTG		5640
	ACGAAAATGA	TTGAAATGAC	TGACAGTATA	CTAATTGGTC	CAGGTCTTGG	CGTTGATTTT		5700
25	AAAGGAAATA	ATGCCATTAC	ATTCTACTA	CAAAATATAC	AACCGCATCA	AAATTTAATC		5760
	GTAGACGGCG	ATGCGATTAC	AATCTTTAGT	AAACTGAAAC	CGCAATTACC	TACATGTCGT		5820
	GTGATCTTTA	CACCACACCT	CAAAGAATGG	GAACGATTAA	GTGGTATTCC	TATTGAGGAA		5880
30	CAGACATATG	AGCGTAATCG	TGAAGCAGTT	GATCGTTTAG	GTGCAACTGT	TGTACTTAAA		5940
	AAACATGGTA	CTGAAATTTT	CTTTAAAGAT	GAAGACTTTA	AATTGACAAT	CGGTAGCCCA		6000
	GCAATGGCGA	CTGGTGGTAT	GGGCGATACA	CTTGCTGGTA	TGATTACAAG	CTTTGTCGGT		6060
35	CAATTTGATA	ACTTAAAAGA	AGCGGTTATG	AGTGCCACAT	ATACACATAG	TTTTATTGGC		6120
	GAAACCTTG	CAAAAGATAT	GTATGTGGTG	CCACCATCAA	GACTTATCAA	TGAAATACCT		6180
	TACGCAATGA	AACAATTAGA	AAGTTAGTCA	TTACTAATCA	TTGAATATAG	TAAAGCATT		6240
40	CTTTCTAGCA	TAAAAATAAG	ACTCCCCTAC	ATATAGGGAA	GTCTTATTTT	TTATTATTCT		6300
	TCATCTGATG	ATTGTTGTAT	ATCTTCTTCA	ACACGATCCA	TGAAATCTTG	TCTTACTTCA		6360
	ATACGTCCAT	CTTCATCATT	TTCTTCTGAA	TCAATCACTT	CAGTATGAAT	TGCATTTTCT		6420
45	GGTGT	TTT	CAT	CATTTaCAAC	CGCTTCACGT	TGTTGTTT	CAG	6480
	GAAGTAGATT	GCTCATCTTC	ATTCGTTTCA	TCTTCTGCAT	CTTCTTTTAC	TTTAGCAACC		6540
	GTTGAAACAA	ATTGATCATC	ACCTAAGCGA	ATTAAGCGAA	CACCTTGTGC	TGCACGACCA		6600
50	TTTTGAGAAA	TATCTGCAAC	ATCTAGTCGA	ATAATGACAC	CTGCATTAGT	AACAATCATT		6660

55

	GTAGCTGTTT TAATACCTTT ACCACCACGA TTTGATAAGC GATAGTCATT AACTGGCGTA	6780
	CGTTTACCAT AACCATTTTC AGTAACTACT AATACTTCAT CAACACTGTT TGCATGAGCT	6840
5	ACATCAAGCC CTACAACTTC GTCACCTTCA CGAAGTGTA TACCTTTCAC ACCCGTTGCT	6900
	GTACGGCCTA AAGGACGTAA TGTGATTCA GGAATCGAA TTAATGATGC ATGTGATGTA	6960
	CCAATCAAGA TATCTTCTTG ACCACTTGTT AAGCGAACTG CAATTAACTC ATCATCTTCT	7020
10	CTGAACGAAA TCGCAATCTT ACCATTTCTA TTTATTCTTG AGAAGTTACT TAATGCTGAA	7080
	CGTTTAACGA CACCACGTTT AGTTGCAAAC ACTAAGAAGT TGTCTTCACT TTCAAGGTCT	7140
	TTAACAGCAA TCATTGTACT AATGACTTCA TCATTTTCAA GTTCAATAGC ATTCACTACA	7200
15	GGAATACCTT TAGACTGTCT TGATAACTCA GGCACCTCGT AACCTTTAAG TTTGTATACA	7260
	CGACCTTTGT TAGTAAAGAA CAATACATGG TCATGTGTAC TTAAAGTTAC CAATTGACTG	7320
	ACAAAATCTT CTTCCAATGT ATTCATACCT TGAACACCAC GACCACCACG GTTTTGAGCA	7380
20	CGATATGTAG ATACCGGCAA ACGTTTAAAG TAGTTATTAT GGCTTAGTGT AATTACTATT	7440
	TGTTCTTCTG GAATTAAGTC TTCGTCCTCT AAGTCTTCAA ATCCACCTAA TTGAATTTCT	7500
25	GTACGACGAT CATCACCGAA ACGATCTCTA ATTTCACTCA ATTCATCTCT AACTAACTGT	7560
	AATAACACTT CTTCATCAGC TAAGATTGCT TCTAATTCAC TAATATAATT TAATAACTCA	7620
	TTATATTCAG CTTCAATTTT GTCTCTCTCT AAACCTGTGA GACGTCTTAA ACGCATGTCT	7680
30	AAAATAGCTT GAGCTTGTTT TTCAGAAAGT TTGAAGCGTT GTTGCAAGCT TTCCATTGCA	7740
	ACTTTATCTG TATCTGACTC ACGAATCGTT GAAATAATTT CATCGATATG GTCAAGTGCG	7800
	ATACGTAATC CTTCTAAAT GTGGGCACGA TCCTTAGCTT TACGTAAGTT GTATTGCGTA	7860
35	CGTCTTCTAA CAACTGTCTT TTGATGCTCT AAATAATGTA CCAACGCTTC TTTTAAATTA	7920
	ATAAGCTTCG GTCTACCATT TACAAGTGCA ATCATATTCA CACCAAATGA TGTGTAAGA	7980
	GGTGTGTTGT TGTATAAGTT ATTTAAAATG ACACTAGCAT TTGCATCCTT ACGCACATCA	8040
40	ATAACGACAC GCACACCAGT ACGTAACTT GTTTCATCAC GTAAATCAGT GATACCGTCA	8100
	ATTTTCTTGT CACGAACGAG CTCTGCAATT TTTTCAATCA TACGAGCCTT ATTCACTTGG	8160
	AAAGGAATTT CAGTGACAAC AATACGTTGA CGTCCGCCTC CACGTTCTTC AATAACTGCA	8220
45	CGAGAACGCA TTTGAATTGA ACCACGACCT GTTTCATATG CACGTCTAAT ACCACTCTTA	8280
	CCTAAATATA GTCCAGCAGT TGGGAAATCA GGACCTTCAA TATCCTCCAT TAACTCAGCA	8340
	ATTGAAATAT CAGGGTCTT ACTTAAGCTA AGTACACCAT TGATTAAATC TGTTAAGTTA	8400
50	TGTGGTGGAA TATTCGTTGC CATACCTACC GCGATACCTG ATGCACCATT GGCTAATAAG	8460

55

	AAATCTATTG TATCTTTATT AATATCACGT AACAGTTCAA GTGTGATTTT AGTCATACGC	8580
	GCTTCAGTAT AACGCATTGC TGCTGCGCCA TCTCCATCCA TTGAACCAAA GTTACCTTGG	8640
5	CCATCAACAA GCGGATAACG ATAAGTAAA TCTTGAGCCA TACGTACCAT TGCTTCATAA	8700
	ATAGATGAGT CACCATGAGG GTGATATTTA CCCATTACGT CACCAACGAT ACGTGCTGAT	8760
	TTTTTATATG ATTTATCCGG TGTACATACCT TGTTCATTTA ATCCATATAG TATACGACGA	8820
10	TGTACTGGTT TTAAACCGTC ACGAACATCT GGCAATGCAC GAGCAACGAT AACACTCATC	8880
	GCATAATCTA AAAATGATTC ACGCATTTCA CTGGTAATAT TTCGTTTCATT TATTCTTGAT	8940
	TGAGGTAATT CAGCCATCAA GAGTTCCTCC TTCAAAAGTT CAGTTCACAG CGCTTAGAAG	9000
15	TCTAAGTTTG CATAAAGTGC ATTATCTTCT ATAAATTGTC TACGGTTTTT TACAACGTCA	9060
	CCCATTAACTA TTTCAAATGT TTGGTCCGCT TCAATCGCAT CTTCAAGTTT TACTTGTAAG	9120
20	AGAGCGCGGT GCTCAGGGTT CATTTGTTGTT TCCCATTAAT GATCTGCATT CATTCTCTCA	9180
	AGACCTTTGT ATCGTGCAAT AGACCATTTT GGTGTTGGAT TCAATTCAGA TTTAAGTTTA	9240
	TCAAGTTCCC TATCATTTGA TACATAATAC TTTTGTTCAC CTTGTGTCAG TTTATACAAC	9300
25	GGTGGCTGTG CAATATACAC ATAGCCTGCT TCAATTAACG GTCTCATAAA TCGATAGAAG	9360
	AATGTTAATA ACAATGTTCT AATATGCGCT CCATCCACAT CGGCATCAGT CATAATGACG	9420
	ATTTTGTGAT ATCTTGCTTT CGTAGATCA AAGTCGCCAC CGATTCTGT ACCAAATGCT	9480
30	GTGATCATTT GACGAATTTT ATTGTTATTC AAAATTCTAT CTAATCGTGC TTTTCAACA	9540
	TTTAATATCT TACCTCGTAA TGGTAAAATC GCCTGCGTTC TAGAGTCACG ACCAGATTTT	9600
	GTAGACCCCC CGGCAGAGTC CCCTTCGACT AAGAAAATCT CACATTCCTC AGGACTTTTA	9660
35	CTAGAGCAAT CGGCTAATTT ACCTGGAAGG CTTGCTACAT CTAACGCTGA TTTACGACGT	9720
	GTTACTTCAC GCGCTTTTTT CGCAGCAACA CGTGCACGTG CCGCCATAAT ACCTTTTTCA	9780
	ACCACTGTAC GTGCGACTTG TGGATTTTCA TATAAAAATC GTTCAAAGTG CTCTGAGAAT	9840
40	AATTTATCTA CAACCTGACG CACTTCAGAA TTACCTAATT TTGTCTTCGT TTGACCTTCG	9900
	AATTGAGGAT CACCATGTTT GATAGATATA ATTGCTGTCA TACCTTCACG TGTATCTTCA	9960
	CCAGAAAGTC TATCTTTTTT TTCTTTCATA ATCTTGCTAC TTAAACCATA ACTATTTAAG	10020
45	ACACGCGTTA ATGCACGTTT GAATCCGTCT TCATGCGTAC CACCTTCATA CGTATGAATG	10080
	TTATTTGCGT AAGTTAAAAG ATTTGTGGCA TATCCTGAGT TATATTGAAT CGCAATTTCT	10140
	ACTTCAATAT CATCTTTAGA TTGATGAATA TAAATTGGCT CATCATGAAT AGGTTCTTTA	10200
50	TTTTCGTTCA ATAAGTCAAC GTACGATTTA ATACCGCCCT CATAGTGATA GGAGTCTTCT	10260

55

5 GCAAGCTCTC TAATACGCTG CTGTAATGTT TCATAGTTGT ATACAGTTGT CTCTGTGAAG 10380
 ATTTCTCCAT CTGCTTTAAA ACGAATGACA GTACCTGTCT TATCAGTnGT GCCAACTTCT 10440
 10 TTTAAGTCAA ATTGAGGTAC ACCTTTTPTA TATGCTTGAT GATATATAGT CTCATTTCTG 10500
 TGTACATATA CTTCTAAGTC TTGTGACAAT GCGTTTACAA CTGATGAACC AACACCATGT 10560
 AAACCACCAG ATACTTTGTA TCCGCCACCG CCAAATTTAC CACCAGCATG TAAAACAGTT 10620
 15 AAAATAACTT CGACAGCTGG ACGTCCCATT TTTTCTTGAA TATCAACTGG GATACCACGT 10680
 CCGTTATCCG TTACTTTAAT CCAGTTATCT TTTTCAATAA CAACTTCAAT TTGATTTGCA 10740
 TAACCaGCTA ATGCTTCATC GATACTATTA TCGACAATTT CCCACACTAA ATGGTGCAAA 10800
 CCTCTCTCTG AAGTCGATCC TATATACATA CCTGGTCTTT TACGTACTGC TTCTAAACCT 10860
 TCTAATACTT GTATTTGCCC AGCACCATAA TTATCCGTGT TGTTTACATC TGACAATGCA 10920
 20 GTCACCATCG CTTTCTGTTA CTTTATAATT TCACCTTGAT TAATACGATA CAATTTAGCG 10980
 TTATTCATGA TTTTATGATC AATACCATCT ACAGATGTCT TAGTGACAAA TGTTTGTACT 11040
 TTATGCTGAA TCGTACTTAA TAAATGCGTT TGACGCGAAT CATCTAATTC ACTGAGTACA 11100
 25 TCGTCTAATA ATAAGATGGG ATATTCCCCA ACTTCGATAT TCATTAACTC AATTTCACT 11160
 AATTTAATGG ACAAAGCCGT TGTACGTTGC TGTCTTGAG AACCATATGT TTGAGCATCC 11220
 ATGCCATTCA CATCAAACT TATATCATCT CGATGTGGTC CGAATAAGCT AATGCCTCGT 11280
 30 TCTTTTTCTC TTTGCATATT ATCGTAAGA ATAGACATAA TTTCTTCAAG TCGTGCCGCT 11340
 TCATTTTGAG CATAATCAA TTTAAGACTA GGTAAATAAT TCAGCGACAA CGCTTCTTTA 11400
 TCATTTGTGA TACCAGCATG AATCGGTTTA GCTAACGACT CTAGCTCTTG AATAAAATGT 11460
 35 GCACGTTTAT CAGTTACTTT CATTCATAT TCAGCAAACCT GCTGATTTAA TACTTCCAAC 11520
 ATTGTTAAGT CCTTTTTTTG GCCTAATTGT AACTGCTTTA AGTAATTATT CTTTTGCTTT 11580
 AAAATACGTT GGTATTGAGC TAAATCATT AAGTAAACAG CAGAAATTG GCCCAACTCC 11640
 40 ATATCTATAA AGCGTCGTCT TATTtGrGGr GAGCCTTTTA CAATATTCAA ATCTTCTGGC 11700
 GCAAATAGAA CCACATTGAG GTGTCCAATA TATTGAGTTA GACGACTTTG CTCTAAGTn 11760
 ATTCACTTTG GACTTGTTTA CCTTTnTTAG TTATAACAT TGTTAATGGG CATCGTGCCG 11820
 45 TGT 11823

(2) INFORMATION FOR SEQ ID NO: 137:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 692 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

5 ATAATTATTA ACATGGTGTG TTTAGAAGTT ATCCACGGCT GTTATTTTGG TGTATAACTT 60
 AAAAATTTAA GAAAGATGGA GTAAATTTAT GTCGGAAAAA GAAATTTGGG AAAAAGTGCT 120
 TGAAATTGCT CAAGAAAAAT TATCAGCTGT AAGTTACTCA ACTTTCCTAA AAGATACTGA 180
 10 GCTTTACACG ATTAAAGATG GTGAAGCTAT CGTATTATCG AGTATTCCTT TTAATGCAAA 240
 TTGGTTAAAT CAACAATATG CTGAAATTAT CCAAGCAATC TTATTTGATG TTGTAGGCTA 300
 TGAAGTTAAA CCTCACTTTA TTACTACTGA AGAATTAGCA AATTATAGTA ATAATGAAAC 360
 15 TGCTACTCCA AAAGAAACAA CAAAACCTTC TACTGAAACA ACTGAGGATA ATCATGTGCT 420
 TGGTAGAGAG CAATTCAATG CCCATAACAC ATTTGACACT TTTGTAATCG GACCCGGTAA 480
 CCGCTTTCCA CATGCAGCGA GTTTAGCTGT GGCCGAAGCA CCAGCCAAAG CGTACAATCC 540
 20 mTTATTIATC TATGGAGGTG TTGGtTTAGG aaaaacccat TTAATGCATG CCATTGGTCA 600
 TCATGTTTTA GATAATAATC CAGATGCCAA AGTGATTTAC ACATCAAGTG AAAAATTCAC 660
 AAATGAATTT ATTAAATCAA TTCGTGATAA NA 692

25 (2) INFORMATION FOR SEQ ID NO: 138:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7900 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

35 ATACTGTAGC GCAAATTTCA CAATGGCATG TTATAGAAGA TTTAGTTACG AATGAATTAG 60
 GTATTAGTAT TTTACCAACA TCAATTTAG AGCAACTAAA TGGAGATGTG AAGCTGtACG 120
 CATTGAAGAT GTCATGTAC ATTGGGAATT AGGTGTTGTT TGGAAGAAGG ATAAACAATT 180
 40 AAGTCATGCC ACAACGAAAT GGATAGAATT TTTGAAAGAC CGTTTAGGCT AACATATTAA 240
 TAAAGCACTC ATTATTTAAG GCGCATCATT ACGTGGGTCA TTGAAATAAT GAGTGTTTTT 300
 45 TTGTGAAAAT GAAGTGAAAT TTAGAGAGCG TTTCCATAGA AAATAGTAAT ACAAACATA 360
 AAAAAAGAGT ATTTTATAT TGTGTACGCC ATCTTTATAA TAGTTATTGT AACAAATTTAG 420
 ACATATTTAG AAAGGGATGG CGCCATGCAC AAAGTCCAAT TAATAATCAA ACTACTACTA 480
 50 CAACTAGGAA TCATCATTGT GATTACTTAT ATTGGCACAG AAATTCAAAA GATTTTTCAT 540

55

	ATTGTACCGC TAAC TTGGGT AGAAGACGGT GCAAAC TTTT TATTAAAGAC GATGGTCTTT	660
	TTCTTCATAC CGTCAGTTGT AGGtATTATG GaTgtgCTTC CGAAATTACG CTAAATTATA	720
5	TACTCTTTTT CGCAGTCATT ATCATAGGAA CATGTATCGT TGCATTATCT TCAGGTTATA	780
	TTGCTGAAAA AATGTCyGtT AAACwTAAAC ATCGTAAAGG TGTAGACGct TATGAATGAT	840
	TACGTGCAAG CCTTATTAAT GATTTTGTG ACTGTCGTTT TATATTATTT CGCTAAAAGG	900
10	TTACAACAAA AATATCCGAA CCCATTTTTG AATCCAGCAT TAATTGCATC TTTAGGAATT	960
	ATTTTTGTCT TACTTATCTT TGGAAATTAGT TATAACGGGT ATATGAAAGG TGGCAGTTGG	1020
	ATCAACCATA TTTTAAACGC AACGGTCGTA TGTTTAGCGT ACCCACTTTA TAAAAATAGA	1080
15	GAGAAAATTA AAGACAATGT CTCTATCATT TTTGCAAGTG TATTAActGG CGTCATGCTG	1140
	AATTTTCATGT TAGTGTCTT AACACTTAA GCATTTGGCT ATTCTAAAGA CGTCATTGTA	1200
	ACGTTATTGC CCCGATCTAT AACAGCCGCA GTAGGTATCG AAGTGTCA CA TGAActAGGT	1260
20	GGTACAGATA CGATGACCGT ACTTTTTTATT ATCACAACGG GTTTAATCGG TAGTATTTTA	1320
	GGTTCGATGT TATTAAGATT TGAAGATT GAATCTTCTA TCGCCAAAGG ATTAACGTAT	1380
	GGGAATGCGT CACATGCATT TGGCACAGCT AAAGCACTAG AAATGGATAT TGAATCCGGT	1440
25	GCATTTAGTT CAATTGGGAT GATTTTAACT GCAGTTATTA GTTCAGTGTT AATACCTGTT	1500
	CTAATTTIAT TATTCTATTA ATTTAGATAT TTAAAATGAT AGACAGAAAG GGAGGCTATT	1560
30	AGTAATAATG GCAAAAATAA AAGCAAATGA AGCATTAGTT AAAGCATTAC AAGCaTGGGA	1620
	TATAGATCAC TTGTATGGTA TTCCAGGAGA CTCAATCGAC GCATAGTCGA TagTTTACGT	1680
	ACAGTGAGAG ATCAATTTAA ATTTTATCAT GTACGTCATG AAGAAGTAGC AAGCTTAGCG	1740
35	GCTGCTGGTT ACACAAAATT AACTGGTAAA ATCGGTGTGG CATTAAGTAT CGGTGGCCCT	1800
	GGTTTAATTC ATTTATTAAA TGGTATGTAT GATGCCAAAA TGGATAATGT ACCGCAATTA	1860
	ATATTATCTG GACAAACGAA TAGTACAGCA CTTGGAACGA AAGCATTCCA AGAAACAAAT	1920
40	TTACAAAAAT TATGTGAAGA TGTAGCCGTT TATAATCACC AAATTGAAAA AGGTGACAAT	1980
	GTGTTTGAAA TCGTTAACGA AGCAATTCGT ACGGCATATG AACAAAAAGG TGTAGCTGTT	2040
	GTTATTTGTC CTAACGACTT ATTAActGAA AAAATTAAAG ATACAACGAA TAAACCAGTA	2100
45	GATACATCAA GACCAACAGT AGTATCACCA AAATATAAAG ACATCAAAAA AGCGGTTAAA	2160
	CTAATTAATA AAAGTAAAA GCCTGTCATG TTAATTGGTG TAGGTGCGAA ACATGCGAAA	2220
	GATGAGCTAC GTGAATTTAT TGAAATGGCT AAAATTCCTG TCATTCAATC ATTACCAGCT	2280
50	AAAACAATCT TGCCGGATGA TCATCCATAT AGTATCGGtA ACTTAGGTAA AATCGGTACC	2340

55

EP 0 786 519 A2

	CCATATGTGG ATTACTTACC TAAGAAAAAT ATTAAAGCCA TTCAAATTGA CACAAATCCT	2460
	AAAAATATCG GACATCGTTT CAATATTAAT GTAGGAATTG TTGGAGATAG TAAAATTGCG	2520
5	TTGCATCAGT TAACTGAAAA TATTAAACAT GTTGCTGAAA GACCATTCTT AAACAAAACG	2580
	TTAGAACGTA AAGCGGTTTG GGATAAATGG ATGGAACAAG ATAAAAATAA TAATAGTAAA	2640
	CCATTACGTC CAGAACGATT AATGGCATCA ATCAATAAAT TTATTAAAGA TGATGCAGTG	2700
10	ATTTACGAG ATGTAGGTAC AGCAACAGTT TGGTCAACTC GATACTTAAA CCTTGGTGTA	2760
	AATAACAAGT TCATCATTTT AAGTTGGTTA GGTACAATGG GTTGCGGTCT TCCAGGTGCA	2820
	ATTGCATCAA AAATTGCATA TCCAAATAGA CAAGCCATCG CAATTGCTGG TGACGGTGCA	2880
15	TTCCAAATGG TAATGCAAGA CTTGCTACA GCAGTACAAT ATGATTTACC TTAACTGTA	2940
	TTTGTAAGTA ATAACAAACA GTTAGCATTT ATTAAATATG AACAAACAAGC AGCTGGTGAA	3000
	TTAGAATATG CAGTTGATTT TTCTGATATG GATCATGCAA AATTTGCTGA GGCAGCAGGT	3060
20	GGTAAAGGTT ATACAATTAA GAGTGCTAGC GAAGTAGATG CTATAGTCGA AGAGGCATTA	3120
	GCACAAGATG TACCAACGAT TGTAGATGTA TATGTTGATC CTAATGCTGC GCCATTACCA	3180
	GGTAAATTG TAAATGAAGA AGCGCTTGGT TATGGTAAGT GGGCATTAG ATCAATTACT	3240
25	GAAGATAAAC ATTTAGATTT AGATCAAATT CCACCAATTT CAGTGGCAGC AAAACGTTTC	3300
	TTATAACTGA TTAAAGGTT ATCACAATTG AATTGAACTA TAAAAACGGT AATTTCTATT	3360
	TCAACAAAT GGAATTGCC GTTTTGTTA TTTATCACAA ATGATCGTAC TGAATTGATG	3420
30	ATAAAATTGT GAAAAAGTTG TTGAAAACGC TTTTACAAAT ATGTATAATA GCTATGAATT	3480
	AGATATCACT TCGTGTTAC TGGTAATGCA GGCATGAGCA AACAAACGCA CTATGAGAAT	3540
35	AGTCTTGTTT GTTCATGCCT GCTTTTMTTG TACATGGAAG CGGAAATTGA GATAGGGGAT	3600
	GTTTETATGT TTAAGAAATT GTTTGGACAA TTGCAACGTA TCGGTAAAGC ATTAATGTTA	3660
	CCTGTTGCGA TTTTACCAGC AGCTGGTATT TTATTAGCGT TTGGTAACGC AATGCACAAC	3720
40	GAACAATTAG TAGAAATTGC ACCATGGTTA AAAAACGATA TCATTGTAAT GATTTGTCG	3780
	GTCATGGAAG CAGCAGGACA AGTTGTATTT GATAACTTGC CATTATTATT TGCAGTTGGT	3840
	ACAGCACTTG GATTAGCAGG AGGAGACGGT GTTGACGAT TAGCAGCGCT AGTAGGTTAC	3900
45	TTAATTATGA ATGCAACAAT GGGGAAAGTG TTGCACATTA CAATTGATGA CATTTTCTCA	3960
	TATGCCAAAG GGGCAAAGA ATTAAGTCAA GCAGCGAAAG AACCAGCACA TGCTTTAGTA	4020
	TTAGGTATTC CAACGTTACA AACGGGTGTG TTTGGTGGTA TTATCATGGG TGCTTTAGCC	4080
50	GCATGGTGTT ACAACAAATT TTATAATATT AACTACCAC CATTTTITAGG ATTCTTTGCA	4140

55

	AGCTTTGCGT GGCCACCAAT TCAAGATGGA TTAAATAGTT TATCGAATTT CTTATTAAAT	4260
	AAAAATTTAA CATTAAACAAC GTTTATATTC GGTATTATTG AACGCTCATT AATTCCATTT	4320
5	GGTTTACATC ATATTTTCTA TTCACCGTTC TGGTTTGAAT TCGGAAGTTA TACAAATCAC	4380
	GCAGGTGAAT TGGTTCGTGG TGACCAACGT ATTGGGATGG CACAATTGAA AGATGGCGTA	4440
	CCATTTACTG CTGGTGCATT TACTACTGGT AAATATCCAT TTATGATGTT TGGTTTACCA	4500
10	GCGGCGGCAT TTGCTATTTA TAAAAATGCA CGACCAGAAC GTAAAAAAGT CGTGGGTGGT	4560
	TTAATGTTAT CAGCAGGATT AACTGCATTT TTAAGTGGTA TCACTGAGCC ATTAGAATTT	4620
	TCATTCTTAT TTGTAGCACC AGTACTTTAT GGAATTCACG TATTATTAGC TGGTACATCA	4680
15	TTCTTAGTAA TGCATTTATT AGGCGTTAAA ATTGGTATGA CATTCTCAGG TGGTTTCATA	4740
	GATTATATTT TATATGGTTT ATTAACTGG GATCGTTCAC ACGCATTATT AGTTATTCCA	4800
	GTCGGTATTG TATATGCTAT CGTGTATTAC TTCTTATTCG ACTTTGCAAT TCGTAAGTTT	4860
20	AAATTGAAAA CACCAGGTCG TGAAGATGAA GAAACTGAAA TTCGTAAGTC TAGTGTGCGA	4920
	AAATTACCAT TTGATGTCTT AGATGCAATG GGTGGAAAAG AAAACATTAA ACATTTAGAT	4980
	GCGATGATTA CACGTCTACG CGTAGAAGTG GTTGATAAAT CAAAAGTAGA TGTAGCAGGT	5040
25	ATTAAAGCTT TAGGCGCATC AGGTGTATTA GAAGTTGGAA ACAATATGCA AGCTATCTTT	5100
	GGTCCAAAAT CAGATCAAAT TAAACATGAT ATGGCCAAGA TTATGAGTGG TGAAATTACG	5160
30	AAACCAAGTG AAACGACAGT GACTGAAGAA ATGTCAGATG AACCAGTTCA CGTAGAAGCA	5220
	CTTGAACAA CAGACATCTA TGCACCAGGT ATCGGTCAAA TCATTCCATT ATCAGAAGTA	5280
	CCTGATCAAG TATTCGCTGG TAAATGATG GGTGATGGTG TTGGCTTTAT CCCTGAAAAA	5340
35	GGTGAAATTG TAGCACCGTT TGATGGTACA GTGAAAACAA TCTTCCCTAC GAAACATGCG	5400
	ATAGGATTAG AATCTGAAAG TGGCGTCGAA GTACTTATTC ATATTGGTAT CGATACAGTG	5460
	AAACTGAATG GTGAAGGATT CGAAAGTCTG ATTAACGTTG ATGAAAAAGT AACACAAGGT	5520
40	CAACCATTAA TGAAAGTGAA TTTAGCATAC TTGAAAGCAC ACGCACCAAG CATCGTTACA	5580
	CCAATGATTA TTACAAATCT TGAAAATAAA GAACCTGTCA TTGAAGATGT ACAAGATGCT	5640
	GATCCAGGTA AGCTAATTAT GACAGTCAAA TAATGATTAA AAATGAAACA GCATATCAAA	5700
45	TGAATGAACT TTTAGTCATT CGTAGTTCGT ATGCGAAGTA GCGAGTTGAA AGAGAATACG	5760
	TTACAAAAGG CAGTAGCTTA AAATGAAGCT ACTGCCTTTT TAGTGCGCAA TGATGTATAG	5820
	CAGGTGTGTT GATGTAAATA AGTTAAATAT TAGTGTTAGA TATAGAAAAC ATTGCTTATG	5880
50	TTTTTGTAC ATTTTAGAAA AATGCATCTT CGCGACTAGC CAAATTAATA GTCTCATTGA	5940

55

	AATAAATTAA CATGATTTTA AATCTATTTG TAAGATAAGG AGATTGTGCA TTATGACAAC	6060
	AGAAGGTCTA TTAGTTGCAG AGAAAGAAAT CGAAGTGAAT GGTACGACA TTGATGCGAT	6120
5	GGGTGTCGTT AGTAATATCG TTTATATTAG ATGGTTTCGAA GATTGAGAA CAGCGTTTAT	6180
	TAATCAGCAC ATGAATTACT CAACAATGAT CAATCAAGGC ATTTACCTA TACTTATGAA	6240
	AACGGAAGCA GAGTATAAAG TACCTGTCAC AATACATGAC AAACCAGTAG GTCGTATTTA	6300
10	CTTAGTTAAA GCAAGCAAGA TGAAATGGGT GTTTCAGTTT GAAATTGTGT CCGCACATGG	6360
	CGTGCAITGT ATTGGTACAC AGACAGGCGG TTTTACAGA TTGAGTGATA AGAAGATAAC	6420
	CTCTGTGCCA CAAGTGTTTC AAGACATTTT AGCAACAAAA TAATGACTTC ATTTTAAAA	6480
15	ATAAAAAGTA AGAAGGTGTT CGAAATGGTT AAGCAATTAA ATAGTGTCGA AGCATTCCGT	6540
	GAATTTATTC ATCAATATCC GTTAGCAGTT GTACATGTCA TGCGCGATCA GTGTAGCGTG	6600
	TGTCATGCCG TTTTACCACA AATTGAAGAC TTGATGCAAT CATATCCCAA TGTGCCATTA	6660
20	GCTGTGATTA ATCAAAGTCA GGTGGAAGCT ATTGCTGGAG AATTAAATAT TTTCaCTGTA	6720
	CCTGTGGATT TAATTTTTAT GAATGGAAAA GAAATGCATC GTCAAGGGCG TTTTATCGAT	6780
	ATGCAACGTT TTGAACATCA TCTTAAGCAA ATGAATGATA GTGTAAATAA CGATGTCGAT	6840
25	GAGCATTAAAT ATCGCAAATG ATTAGCATTG CTAAGATTAT GTAGACATCA TAACTTATTT	6900
	CCCAGTAAAT ATTGGTAGTA ATTAGAATCA GCATGGTACA GTAGAACTAT AGTAGAAATC	6960
	ATCAAAGAGG AGTGACGACA AATGCGTAAA AAATGGTCTA CACTTGCGTT TGGATTTTAA	7020
30	GTTGCAGCAT ACGCACATAT TAGAATTAAA GAAAAACGCA GTGTGAAAAG TTATATGTTA	7080
	GAACAAGGTA TACGATTATC TAGAGCTAAG CGTCGTTTTA TGTATAAAGA AGAAGCGATG	7140
35	AAAGCATTAG AAAAAATGGC GCCACAGACA GCAGGCGAAT ATGAGGGAAC CAATTATCAG	7200
	TTTAAGATGC CAGTAAAAGT GGATAAGCAC TTCGGTTCAA CCGTTTATAC CGTTAACGAT	7260
	AAACAAGATA AGCATCAACG CGTTGTATTA TATGCACATG GAGGCGCATG GTTCCAAGAC	7320
40	CCACTCAAAA TTCATTTTGA ATTTATTGAT GAACTTGCAG AAACACTCAA TGCTAAAGTC	7380
	ATCATGCCAG TATATCCGAA GATTCCGCAT CAAGATTATC AAGCGACGTA TGTGCTTTTT	7440
	GAAAAGTTGT ACCATGATTT ATTGAATCAA GTAGCAGATT CTAAACAAAT CGTTGTAATG	7500
45	GGTGACTCTG CGGGCGGTCA AATTGCTTTA TCATTTGCTC AATTGTTAAA AGAAAAACAT	7560
	ATTGTGCAAC CAGGACATAT TGTATTAAAT TCACCACTTT TAGATGCAAC GATGCAGCAT	7620
	CCTGAAATTC CTGACTACTT AAAGAAAGAC CCAATGGTAG GTGTGGATGG CaGTGTGTTT	7680
50	TTAGCTGAAC AATGGGCAGG GGACACACCT TTAGATAACT ACAAAGTATC ACCAATTAAT	7740

55

CCAGATGCTT TGAACCTATC GCAATTGTTG AGTGCGAAAG GTATCGAACA TGACTTTATA 7860

CCTGGATATT ACCAATTCCA TATTTATCCA GTATTTCCGA 7900

5 (2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1984 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

15 GTCTAAATAA ACAAATTAT CATTGATTaC TGAACGGCA TTTCGAAGTA ATGCTTCAAT 60
 ATCATTCGAA TATTTCTTCA ATTTATGATT GTGAAATAAT TCTTGCATCA AAAATGGTCT 120
 20 TTGGTCACAT GAATGTGCAT CTGAAGCTAC AAAATGAGCC AAATTACATT CTATAAATTG 180
 TAATGATAAC TTTTGAATGT TTTTACCAA TCCACCAACT AAAGAACTCG ATGTTAATTG 240
 ACTCAGTGCC CCATTGCAA CCAATTCATA TAATATTTCC GGATTTTGG CGATACTTCT 300
 25 ATTTCTTTCA GGATGTGCAA TGATTGGTAT GTAACCTCTC GATTGTATTT CAAAAACAA 360
 TTGTTTTGTA TAATGTGGTA CTTGCCCCGT TGGAAATTC AATTAATAAAT ATTCGAACG 420
 ATTAATACCT TGAATACTAC CATTATCTAA GCCTTTCAGA ATCGAATCTG TAATTCTAAT 480
 30 TTCTTGCCCG GGAAATAATT TAATATCCAA TGCTTGAAC TCTGGATGCG TTCTTAATC 540
 CGCCAATTC ACAAGCACTT GTTGAAATGT ATTATCATAT CTCGGATGCA AATGATGAGG 600
 TGTGCTACA ATACTTGTTA CACCTTCATC CTTAGCTTGC TTTAATAGTG CAATACTCTT 660
 35 TTCAATTGTT TTAGGACCAT CATCTATATC AACTAATATA TGGTTATGAA TATCAATCAT 720
 GATTCATCAG TCCATAATA TGCATAGTAA CTAGCACTTT TATCTTAGG CATTCTATTT 780
 AAGACTACAC CTAATAATTT AGCACCTGTT GCTTCAATAA GTTCTTTTCC TTTTAACT 840
 40 TCATCTCTAT TATTATTTTC CGAATTAAC ACGTAGACAA CATTGCCGGT AAACCTTGAA 900
 AATAATTGCG CATCTGTAAC TGTGTTCACT GGTGGCGTAT CGATAATTAC AAAGTTATAA 960
 TTCATCAATA ATGTGTCATA CAAATTTGCA AATGCCCTTG ATGTAATTAA CTCTGACGGA 1020
 45 TTGGGTGGGA TTGGCCAGA CGTCAAGACG TCTAAATCTT GAATTCAGT TGAGATAATA 1080
 CTGTCTTGAT AAGTTGACCA ATTTAGCAAT AAACCTGATA GGCCTTCATT GTTTGGCAA 1140
 TTAAAAATAT AATGCTGCGT AGGTTTACGC ATATCCCCGT CTACGATTAG TGTTTTATAA 1200
 50 CCTGCTTGCG CATATGCAAC TGCTAAATTT GCTGCAATTG TAGACTTACC TGCGCCTGGT 1260

55

GATCTTATGC CTCGAAATTT CTCGCTAATA GGTGACTTTG GTTGTTTCATG GACAATTAAA 1380
 CTTGATGTAC TTCyTCGTGT ATTCGTCATG GTAATTCCTC GTAAATTAAA ATTTTGTAT 1440
 5 TGAACCTAAA ATAGGTAATC CTAGTTGCGA TTCAACATCT TCTTCTGTCT TAATACGCTT 1500
 ATCTAATAAT TCTTTTAAGA AAATAATCAA TATTGCTAAA ACAATACCAA CAATAATGCT 1560
 GATAACTAAG TTGACAGATA CTATTGGAGA TACTTTTACA GCATTATCAT GTGCTGAGGA 1620
 10 AAGTATCGTA ACATTATCAA CACTCATAAT TTTAGGCATG TCATGAGCAA AAACTTTAGA 1680
 TATTTTATTA ACAATTTTGT CAGATTGAGA TTTATTCCCA GTGGTAACTG ATACAGTAAT 1740
 AATTTGAGAG TTTGTTTGAT TGGTTACTTT TAAAAATGAA TTCAACTCAG CTGTTGAATA 1800
 15 CTGACCATCA AATTCTCTAG ATACTTTATC TAGAATTCTA GGACTTTTGA TAATTTCCGT 1860
 ATATGTATTA ACAGACTGCA AACTACTTTG AACATTTTGG AAAGCTAAAT CACTTGAGGA 1920
 CTTTTTCATG TTCACTAATA TTTGAGTAGA AGCAGTATAT TTGTCAGGCA TAACAAAAAA 1980
 20 GGTT 1984

(2) INFORMATION FOR SEQ ID NO: 140:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6272 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

CAAATCCCTT GGTGATGATA AATGLATTGC TGTGTAGCCA AATAATCTTC GTATATATGA 60
 35 CTGACGTTCA ACAACAGCTT GCAATCGTTT CGTTGGTACA GTTACTTTCT TCTTGTTAAA 120
 GAGACCATAT TCAATTTTAA GTTGCTCATT TTCAAGCATC ACCGAAAAGC CATAAAATCT 180
 TATCATTGTT ATAATCGTTC CAATAATATA TGCCACTATT AATACTAGTA AAATGATGAT 240
 40 TAATACTGAA ATACTTACAA TTTGAACCCA TTGACTAATT TCATGATTTA GCTTCGACCA 300
 TGGGATCAAC TCTCTTACAG CCCCCTAAAT CGGTACTAAA GCTGCTAACG TTACACCAAT 360
 GGCGCCACTG GTCATTGCCA TAAATAGTGA TTTCTTAAAA TTCATCTGAT ATATAGGAAT 420
 45 GCGTTTATTT TTCTGATTAA GCATACTATC AGTGTCTGTC ACTTCATCTA AGCGACCTTC 480
 TGCGATGTCT TCCACATTAC CTTCAATGTC ATGATTACAG TTGTCAATTCT TCTCAGCACT 540
 AGACTTTTGC GCCACTTCTG TCTTCAACTC TGTTTGCAAT TGATCAATAT ATCGTTCAAG 600
 50 ATATTCACCT TGTTTTTTTCG AAATAACACT TAAGACAATA CCATCACTTG GTGTTTTGAT 660

55

	AATACGTTTT ATATTTAATT CTTTACGCTT TTTATTAAAA ATACCTGTTG TTAAAATGAA	780
	ATAATTATCC tCAATCCAAT ATCGCGTGTT CATAATTCCG ACAATTTGAG AAATGTATGA	840
5	TATTAAAAAG AATACAAATA CAATACCTAT CCATAAATAT GATTCCGGAT TCGTATAATC	900
	AAAATCTTTC AATTGAAAGA TAATGAAAAT AAAAAAGACG ACTATGTTTT GTTTGATAGC	960
10	ATTGATTATG CCATTAAAAT ATGAAATCGG ATGTAATTTT TGAGGTTTCA ACATCACTTT	1020
	CAACCCCTCT CAAATTCGAC ATAGTTCTCT CTTCGATTAT TTAAACATCG TCATGAGACA	1080
	TCATCGGTAA ATAAATAGTA TGACCTGCAG TCATAAATCC AACTTTATAC AAATTAAGCA	1140
15	CTTTACTAAT TGGATTAGAT TTAATCGACA AGTATTGTAA ACGTTCAATT CGACTCGTTT	1200
	CTTCTTTATA TATAAAAAAT GATGTACGAT ATTGTACACT TAGTTGATCA ACTTTATAAA	1260
	AGCGACAATG ATATTGCCAT AAAGGCTTAA TAAATAATTT TAATGTACTC AGAGCACCTA	1320
20	AAACCAACAA AATATAAAGT AAGTAATGTG GCCATTCAAA TCTTAACCAT ATAAATATAA	1380
	AAATGACATA CACAGCTACA CTCAATATAA ATTCTAAGCC ATTCGTAATG TAGTAATACA	1440
	ACAAATGCTGA CTTAGGACTC TTAGTCAACT TAGTATAATC TGACATATAC CCCTCTCCCC	1500
25	AAATAAAAAA TTATACGGAT TTATAATCTA TTTCATTTTA TTTTATATG ATGATAATTA	1560
	TAGCATATGG AATATTTTCAT GCTAATTTAT TCTTCCTAAA GGTACATCTA AAAATTTAAT	1620
30	TAAGCAGAAA GTGCTTGAAT TGCTAAAAAG ACACCATGTT ATAATTTTAT CAACATGATG	1680
	CCTTTCATCT ATAATCAATC TTTCATCTTA TCAAGAGCGA TATTTAGTTC AAGCACATTC	1740
	ACATAATCAT TTGTTAACAC ACCACGCTGC TTACGATGTT GAATCAAGTC GGCCACTCTT	1800
35	GAAGTAGATA CATGACGAGC ATCAGCAATA CGAGGTGCTT GCTTCAATGC ATTTTCGACC	1860
	GTAATATGCG GATCTAAGCC CGACCCAGAA CTGTGTCAG CATCTATTGT TACATTTGAA	1920
	TTCCCAAATT TAACATGATG TTTCATGCGT GCTATTAATT CGGTGTTTCC ATTCGATTCA	1980
40	TTACTTCCAC CTGAAGATAC GCCGTTTTTA TATAATTTTT CAGGATTCAT ATTATAATCA	2040
	ACTGCACTCG GTCTCCCGTG AAAATATCGT GTCTCTGTCC AGTGCTGTCC AATCAATTTT	2100
	GATCCAAC TAACGATTGTC ATACGTAATT AAAGTCCCAT TTGCTTGTTG ATAAAAAAT	2160
45	ATTTGACCAA TTAACGTGAT AGCTAACGGG AATAAAAATC CACATAATAC CATAGTTATT	2220
	ATCGTTAAAC AAATACTATT TCTTATCGTA TTCATGGTAC AGGCTCCTTC CTCTTTACAC	2280
50	AAAAAATTGT ACAATCATAT CTATTAATTT AATGCCTAAA AACGGGACGA TTAATCCACC	2340
	TAATCCATAA ATCAACATAT TATTTATAAA GATTCTATCA ATGCTGTAAC CCTTTACTTT	2400
55	TACACCTTTC ATGGCAATTG GAATTAAGGC AACAATGATT AATGCATTGA ATATCAAAGC	2460

	AATTGTTGAC ATCATTAGTG CAGGTAAAAT TGCAAAGTAT TTTGCTACGT CATTAGCCAA	2580
	ACTAAATGTC GTTAATGCAC CTCTCGTCAT TAATAATTGT TTGCCTATTT TTACAACCTC	2640
5	TATTAACCTTT GTAGGATTCG AATCTAAATC AATTAGATTA GCTGCCTCTT TAGCACTAAT	2700
	TGTCCCTGAG TTCATAGCTA ATCCTATATT CGCTTtGTGc tAGCGCAGGT GCATCATTG	2760
10	TACCATCTCC TGTCATCGCA ACAATATGGC CTTTCGCTTG TTCATCTTTG ATGACTTTAA	2820
	TTTTATCTTC GGGTTTACAC TCTGCAACAA ATCTATCAAC CCCGGCTTCT TTTGCAATTG	2880
	TAGCTGCTGT TAAAGCATTa TCACCTGTAC ACATAACTGT TTCAATCCCC ATTTTTCTCA	2940
15	ATTCAGTAAA TCGTTCTACA AGACCATCTT TAATCACATC TTTTAAATAA ATCAGGCCAA	3000
	GCATGACATT GTTTTCAATG ACTATTAAaG GnGTGCCACC TTTACTCGAT ACATCCATAC	3060
	AGAGAGACTC AATATTAAGA GGAATATTGC CTTGTTGTTG TTTGACAAGA TTTATCATAC	3120
20	TATTAGGTGC ACCTTTGAAT ACCGATATTT CATTTGTAAT GATTCCGCTC ATTCTAGTTT	3180
	CAGCTGTAAA AGGCTTATAT GTGCCATCAA TGTCTTTAGG CAGCTCATT ATATACATcT	3240
	GcttCGCTAA TCGTACAATA CTTTTTCCTT CTGGCGTATC ATCGTAGATT GATGACATAT	3300
25	AAGCAGCGAC TATCAATTTT TCAAGCATTG GTTGATTAC TGGTAAAAAT TCACTAGCGA	3360
	TTCGATTGCC ATAAGTGATT GTGCCTGTCT TGTCTAAAAT CATTACATCG ACATCTCCAC	3420
30	ATACTTCTAC AGCACGCCCC CTTTTCGCTA ATACATTGAA TTGAGTAACA CGATCCATGC	3480
	CTGCAATACC AATCGCCGAT AACAAACCAC CGATTGTCGT TGGTATTAAA CATACTGTTA	3540
	ACGCAATGAG CATCGCAATA GGTAAAATTA AATGCAGGTA AGATGCTATT GGATATAACG	3600
35	TTACAATAAC GACTAAAAAT ATAATTGTTA ACGTTGTAA TAATGTAAAA AGTGCAATTT	3660
	CATTGGTGT TTTATTTCTT TCCGCCCTT CAACTAAGGC AATCATTTTA TCTAAAAAAG	3720
	ATGTACnCGC TTCACTCTCA ACACGTATTT CTAACCAATC AGATGTTACA AGTGTTACCGC	3780
40	CAATGACTCC ATCAAAATCG CCACCTGATT CTTTTATCAC AGGTGCAGAC TCACCAGTAA	3840
	TTGCAGATTC ATCAACGGTT GCTAATCCAT TTATTACAAC GCCATCAGCA GGGATTGTTT	3900
	CTCCATTTTC TACCCGAATA TTTTGTCCGG CTTTTAACTC TGTGGCGTTC ACTATCCGAT	3960
45	ACGCACCATT TTCTTCTATC AATCGAGCAG TTAATTTGA TTGTGCTTGT CTTAAACTAT	4020
	CAGCTTGCGC TTTTCCACGA CCTTCAGCAA AGGCTTCTGA AAAATTAGCA AACAATATAG	4080
50	TTATTAATAA TATGATAAAA ATTGTAATCA AATAACCTCG CGATAGATAG CTAGTTCCAA	4140
	ATATGTCAGG AAAACATATT AATATCAACG TTAATCAT TCCAACCTCA ACGACAAACA	4200
	TTATCGGATT TTTTATTAAT TGTTTAAGAT TCAGCTTATA AAAACTCATT TTCAAAGCTT	4260
55		

	TTTATTTTAA AGTTAAAAAT TCACCAATAG GACCAAGTAA TAGTACTGGA ATAAATGTCA	4380
	AACCACTTAG TAAAACGATA AATACGATTA GTGATACGCC AAAATAAGGT TTATCAATCG	4440
5	CTATTGTATA TTTATCTTGA TGGTATGATT TTTTATTCAC TAAACTTGAT GCAATCATT	4500
	ATTGCAAAAT AATTGGTATA TAACGAGAAA GCAACATAAT GATTCTGTGA GAGATATTCC	4560
10	AGAATGTTGT ATCATCTTTC AGTCCTTCAA ACCCTGATCC ATTGTTGCA GCAGCTGATG	4620
	TCATTTTATA CATAACTTGT GAAATACCAT GAAAAGACGG ATTCTGTATA CTTTCACTTG	4680
	CTCCAGGAAT CATAAAAGCA AGTGCTGAAA ATACTAAAAT TAAAATTGGG TGTATGAGAA	4740
15	AGACTAAGAC AATACATTTT ATTTACGGG CGCCAATTGG CATATTTAAA TATTCTGGTG	4800
	TTTTACCAAC CATCAAACG CATATAACA CCGTCAGTAA GACAAATATC AATAAATTCA	4860
	TGAGTCCTAC GCCTTCGCCA CCAAATACAA CATTTAGCAT CATTAAATACC ATTGGTCTTA	4920
20	ATCCACCTAT AGGCGTTAAG CTATCATGCA TGTATTAAAC AGAACCCGTT GTAAATGCCG	4980
	TCGTAATAAC TGTAATAGT GCTGACAAAC CTGCTCCAAA CCGTACCTCT TTACCTTCCA	5040
	TATTCGGTCC ATAAATGCCT AAATTCGCTA GTATTGGATT ACCACGATAC TCACTCCACA	5100
25	TAGTTAATGT AAGAATTGCT ATAAAAATGA AAAACATTGC GACAAATAAT ATCAACGCAT	5160
	GACGATGTAC TCGTTTACCA TGTCTACTTA ACATGCGACC AAATAAGAAC AACATTGACA	5220
30	TAGGAAGTAA CATCATACTG CCCATTTCTA TAAAATTGCT CCAAATATTT GGATTTTCAA	5280
	AAGGTGTTGC AGAATTTTCT GCTAAAAATC CTCCACCATT CGTACCAAGA TGTTTTATTG	5340
	ATTCAAGTGA TGCAATAGGT CCAAATGCAA TATGTTGAAT ATGTCCGCTT AAAGTCCGAA	5400
35	TCATTAAATT AGCATGCAAC GTTTGTGGTA CaCCTTGAGT CATCAATAAA ATACTAATTA	5460
	AACATGATAA TGGTAAAGT ACTCGGACAA TAAACCGAAC AATATCTTGA TAAAAATTAC	5520
	CAATGATATT AGTTAATCCA GTTAAACGTC TCAACATCGC TATACAAACG GCGTAACCTG	5580
40	ATGCACTAGA TGTAACATT AAATATGTCA TTACAATCAT TTGCGTTAAA TATGTCACAT	5640
	CTGATTCACC GTTATAGTGT TGtAAATTAC TATTTGTtAA AAAAGATATT GCTGTATTAA	5700
	ACGCTAAATC TATCGATTGG TTTAAATTAT GATTTGGATT TAAAAAAGC CATTGCTGAA	5760
45	CTATTAGCAA TACAAATGTT ATAAACCCCA TAAATCCATT AAATGCCAGA AAATGTTTGA	5820
	CATATGTTTT AGCTGACATG TGTTCTAAAT CTGTGCCGAT AATTTTAAAA CACATATTTT	5880
50	CAAATCTAGT AAATATTAAA TCTACTCTTG ACGATTGCAC CAATGCTACG CGATATAGAT	5940
	ATCCACTAAA AACATACGTA ATCATAACCA TCATTGTTAG AAACAAAATT ATTTCCATGA	6000
55	TAACCCTCAC TTAATATATT TCTAAAATTT TTCACTACGA ATTAAGGCAT AAAATAAATA	6060

ACACAACAAC ATCGTAACAA CTTGTTTATG AGAGAAATnT TAATTTTCAA ACTTAGTTAT 6180
 TAAGAAAnCA TTAAGATGTG TATGCAGAAA TAAATTTTAT AGCATTTAAT TGTGAAGAAT 6240
 5 ATTATGATAT TGCTATCGAG GTGAAGGTTA TG 6272

(2) INFORMATION FOR SEQ ID NO: 141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1978 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

AAATGATGTT TTACAATAAA TATAAaACG TATCAACATA TATCATCATA TTTTtagTTT 60
 20 CAAGTGCAGC CTTTGCAATA TTCTTGTTAA GTGCGnACAT TAGTGCTCAC TCGGAACAAG 120
 TGTACGAAAT GACTGACCAT CAAATTAAGA ACAATACGAT AAATAAAGCA TACGAACATA 180
 AAGACCCTAC AAACAATAGC GAACAAAGAG ATGGGAAAGT GTTCGCTTTA ATAAATTGAT 240
 25 ACATTGTCAC AACGTTATTT TGCCTATTTT TGCGmAATAG CGTTTTTTTAT TACwTTTTTG 300
 CTGATsTTAA ATTTGTTATA TTTTGTTAAA GTATTATAAT GATTGAATAA ACAAATTGAA 360
 GGTAGGTTTT TTAATTGAGT AATTCTGATT TGAATATCGA AAGAATTAAC GAGTTAGCTA 420
 30 AAAAGAAAAA AGAAGTAGGA TTAACtCAAG AAGAAGCAAA GGAGCAAACA GCCTTAAGaA 480
 AAGCTTATCT TGAGAGTTTT AGAAAAGGGT TTAAACAACA AATTGAAAAT ACTAAAGTAA 540
 TTGATCCAGr AGGTAATGAT GTAACACCTG AAAAAATTAA AGAGATACAA CAAAAAGAG 600
 35 ATAATAAAAA TTAAATCACA AATCTGTAAA GAATTTTCTG ACATTATAAC TTGAAATAAG 660
 TATTTTACTT ATCTTTTTAT TTTAAAATAA GTTATAATGT ATTTGATAAA ATTGAAGAAG 720
 40 GGAAGATACA CAAGATGTTT AATGAAAAAG ATCAATTAGC TGTTGATACG CTACGTGCAC 780
 TAAGTATCGA CACAATCGAA AAAGCGAATT CTGGTCATCC AGGATTACCT ATGGGAGCTG 840
 CCCCAATGGC TTACACTTTG TGGACACGTC ATCTGAATTT TAATCCACAA TCTAAAGATT 900
 45 ACTTCAATAG AGACCGTTTC GTATTATCTG CAGGGCATGG TTCAGCATTa TTGTATAGCT 960
 TGTTACATGT TTCTGGTAGT TTAGAATTAG AAGAATTAAA GCAATTTAGA CAATGGGGTT 1020
 CTAAACACC AGGTCATCCT GAATACAGAC ATACAGATGG TGTAAGAAGTT ACTACCGGAC 1080
 50 CACTTGACA AGGTTTTGCT ATGTCAGTAG GATTAGCTTT ACAGAAGATC ACCTAGCAGG 1140
 gAAATTTAAT AAAGAAGGAT ATAATGTTGT AGATCAITAC ACATATGTAT TAGCTtCTGA 1200

AAGTAAATTA GTTGTTTTAT ACGATTCAAA TGATATTTCA TTAGATGGCG AATTAAACAA 1320
 AGCTTTTTCT GAAACACAA AAGCTCGTTT TGAAGCATAT GGTGGAATT ACTTACTAGT 1380
 5 TAAAGATGGT AATGATTTAG AAGAAATTGA TAAAGCGATT ACTACAGCTA AATCTCAAGA 1440
 AGGACCAACG ATTATTGAAG TTAACAAC AATCGGATTT GGTCACCGA ATAAAGCAGG 1500
 AACTAATGGT GTTCATGGG CACCTTTAGG TGAAGTTGAA AGAAATTAA CATTGAAAA 1560
 10 TTACGGTTTA GATCCTGAAA AACGTTTTAA TGTTTCAGAA GAGGTATACG AAATTTTCCA 1620
 AAATACTATG TTAACCGTG CTAATGAAGA TGAATCTCAA TGAATTCAT TATTAGAAAA 1680
 ATATGCAGAA ACATATCTG AATTAGCAGA AGAATTTAAA TTAGCGATTA GTGGTAAATT 1740
 15 GCCTAAAAAT TATAAGGATG AATTACCACG TTTTGAAC TGTCATAATG GTGCATCTCG 1800
 TGCTGATTCT GGTACTGTTA TTCAAGCAAT CAGTAAACT GTCCCTTCAT TCTTTGGTGG 1860
 ATCAGCAGAC CTTGCTGGTT CAAACAAATC CAATGTAAAT GATGCAACTG ATTATAGTTC 1920
 20 TGAAACACCT GAAGGAAAA ATGTGTGGTT TGGGTACGT GAATTGCTA TGGGTGCT 1978

(2) INFORMATION FOR SEQ ID NO: 142:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7588 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

TAGTAGTATT TATTAAATTA TACGAAGGGA CCcAACACAG AAAATTCATT TTATTGAATT 60
 35 TTACATTTAT GTGCCAAGTT GGGAAAAATG TCTATTTTTT TCaAGTATT TAAAGTAAA 120
 ATTACATGTT AATACGTAGT ATTAATGGCG AGACTCCTGA GGGAGCAGTG CCAGTCGAAG 180
 40 ACCGAGGCTG AGACGGCACC CTAGGAAAGC GAAGCCATT C AATACGAAGT ATTGTATAAA 240
 TAGAGAACAG CAGTAAGATA TTTTCTAATT GAAATTATC TTAGTGCTGT TTTTAGGGA 300
 TTTATGTCCC AACCTTTTTA GAATATTAAA TTTCTACAAT TCGTCATCT TCAACAATAA 360
 45 AGCCCATTTGT ATTGACGCTG TTATTTAAGA AAGTCAGAAT ATAACGCATT ACTTCATCAC 420
 GTTCTGGCTC ATTGTGAACC TCGTGGTAAA AACCTTGCCA AGCTTTAAA TATAATTCAG 480
 GTGTTTGATA TTTTCTTTA AACTCATCAA TTGCCCTAGT ATCAACAATT AAATCCTTCG 540
 50 TTCCATACAT TAATAGCGTT GGCATTGGTT GAATGTCATG AATATGAGCC ATCGTATCTT 600
 TCATCGTCTC ATTAATTGTA TTATACCAAT GATACGTTGC TTTTTTAAAC ATTAAACCAT 660

55

	CATTAAAACG TGTGTCCTTTT GAAATTTTAC CTATATTTGA AACAAAGTTTA TCTTTACGAT	780
	TTTTTCCATT CTTTTGAAGT TCTAGCATAG GAGAAATTAA CATCATCCCC TCGATTGGCA	840
5	ATTCTACTTT TTCAAGTAAA TTTAATAAAA TCAAACCGCC AAGTCCTACC CCTAATACAT	900
	AAGTAGGAAT TTTATATTCA TTAGCTATCT TTAACCACTC TAGCAAACCT TCGTGATACG	960
	TTTGAAAGTT TTCAATTTGT CCTTTATTAG CTCTTGAAGT TTGACCTTGA CCAGGCAAAT	1020
10	CTCCCATAAT CACATGATAG CCATTTCTTC TTAACATCGT AATAACATAT GCATATCTTC	1080
	CCGTATGTTT TAATATATTA TGAGCAATAA CAACGACGCC TTTCGCATCA TTTTCAGCTT	1140
	CCCCTTCCA CATTATTATA CTGCCCCCTT TTCATTAATC TTCAATAACA TAATTATAGC	1200
15	AAATCCAATA TGTAGATTTC TATTTATAGT ATTATTGTTG TCCATATTAT TATATATAAA	1260
	TGAAATCAAC ATCAATAATA GTGTAATTAT ACATAATTAT TTTTGATTGT TTTTGATGAA	1320
20	AACGCTTTCT CGAATATTTT TTTTCATGCTA AACTTATTGT AAACACAAGG GTTTGGAGGA	1380
	GTAGCAATGG CACTATTAAA GAATTTTTTT ATCGGATTAT CTAATAATAG TTTTTTAAAC	1440
	AACGCAGCAA AAAAAGTGGG CCCACGTTTG GGCGCCAATA AAGTCGTTGC CGGAAATACA	1500
25	ATTCCAGAGT TAATTAATAC AATCGAATAC TTAATGACA AGAATATCGC TGTACGGTA	1560
	GACAATTTAG GGAATTTGT CGGTACAGTT GAAGAAAGTA ATCATGCTAA AGAACAAATT	1620
	TTACAATTA TGGACGCGCT TCATCAACAT GGCCTAAAGG CACATATGTC TGTAAATTG	1680
30	AGTCAGTTAG GTGCAGAATT CGACTTAGAA TTAGCTTACC AAAATTTAAG AGAGATTTTA	1740
	CTTAAAGCAA ATACTTACAA CAATATGCAT ATAAATATTG ATACTGAAAA ATATGCTAGC	1800
	CTGCAACAAA TTGTTCAAGT TTTAGATCGC TTAAGGCG AATTTAGAAA TGTGGTACT	1860
35	GTAATTCAG CATATTTATA CGATAGCCAC GAATTAGTTG ATAAGTACCA AGATTTACGA	1920
	TTACGTTTGG TTAAAGGTGC ATATAAAGAA AACGAATCAA TTGCATTTCA ATCTAAGGAA	1980
40	GACGTAGATG CAAATTACAT CAAAATAATT GAACAACGTT TGTAAACGC ACGCAATTTT	2040
	ACTTCAATTG CAACACATGA CCATCGCATC ATTAATCATG TAAACAATT TATGAAAGAA	2100
	AATCACATTG AAAAAGATCG TATGGAATTC CAAATGCTCT ATGGTTTTAG ATCAGAGTTA	2160
45	GCAGAAGAAA TCGCAAATGA AGGCTATAAT TTTACTATTT ATGTACCTTA TGGCGATGAT	2220
	TGTTTTGCGT ATTTTATGAG AAGATTAGCA GAACGCCAC AAAACCTATC TCTTGCTGTA	2280
	AAAGAATTTG TGAAACCTGC TGGCTTAAAA CGTGTGGCA TAATTGCAGC TTTAGGAGCT	2340
50	ACAGTTATGT TAGGTTTAAG TACAATTAAA AAATTATGCC GTAAATAGAG CAAGACATAA	2400
	ACAATAATTT AGGAGTCTGG AACAATAATC AATGTTCTAG GCTCCTAAAT GTTATATTGG	2460
55		

	TAGATTTTAA TAAATTAGCC ATTTCAATTG CACTTACTGC TGCTTCAGCA CCTTTATTGC	2580
	CAGCTTTTCGT ACCTGCTCTT TCCACAGCTT GTTCAATACT TTCAGTCGTT AAAATACCAA	2640
5	ATATGACTGG TACATTAGTT TGATCATTCA CTTTAGAAAC ACCTTTCGCG ACTTCATTAC	2700
	AAACATAATC ATAATGAGAC GTAGCACC GC GAATTACGCA TCCTAATGTA ATTACTGCAT	2760
	CATAATTTCC TGATGAGGCT AATTTTTTAG CTACTAAAGG AATTTCAAAC GCACCTGGCA	2820
10	CAAATGCTAC ATCAATATTG TCTTCATTAA CATCATGTCG AATCAAAGTA TCTTTTGCAC	2880
	CTTCAAGTAA TCTTCCAGTG ATAAAATCAT TAAATCGACT AACTACGATT GCAACTTTCA	2940
	AATCTTTTCC AATTAATTTA CCTTCAAAAT TCATGTTAAA ATCTCCTAT ATTAAATGAC	3000
15	CCATTTTTAT TTTTTTCGTT TCCATATAAT CATGATTATG TACCGTTTCT GGTACGATAA	3060
	CTTCAATTCT TTCTGCAATA TCAATGCCAT ATTGTTTTAA TCCCTCAAAT TTACTTGGAT	3120
20	TATTACTTAA TAAATTGATA TGTCGATGT TAAATATTT TAAATCTGT GCAGCAATAT	3180
	GATAATCTCG CAAATCTTCA TCAAAACCTA ATGCTAAATT TGCAGTTACT GTATCATATC	3240
	CTTGCTCAAT TAATTCATAT GCGCGTAATT TGTTTAAACAA TCCTATGCCA CGACCTTCTT	3300
25	GAGGTAGATA AATAATCATG CCACCATGTT CATGATATA CTTCATAGAC GATTCAAGTT	3360
	GAGCACCACA ATCACAACGT TGACTATGGA AAATATCGCC TGTAAGGCAC GCAGAATGTA	3420
	AGCGTACATT TTCATGTTGT CGAATTGCAC CTTTGTGTCAG TACAACTATC TCTTCATCTG	3480
30	TGTATGTCGC TTTAAAACCA TACATATCAA ATGTTCCGAA ATCTGTAGGC ATTTTCACTT	3540
	TTGCCTTAAA TTCAATTTCT GGTCTAATT TTTTACGATA TTCAATTAAA TCATCAATCG	3600
	TAATCATCTT TAATGATGT TTTTCTTTAA ACTTTTGTA ATCTTGTCTT TCGCCATCG	3660
35	TGCCGTCATC ATTCATAATC TCACAAATGA CACCAGCGGG CTTGGCACCA GTAAGTTTAG	3720
	CTAAATCAAC AGCCGCTTCT GTGTGTCCAT TTCTAGCTAA TACGCCTTTA TCTTGTGCTA	3780
40	CTAATGAAA TAAATGACCA GGACGATTAA AATCTTTAGC TTCACTACTA GGATCAATGA	3840
	GCTTTTGGC AGTCAATGTA CGTTCATAAG CACTAATTCC TGTGTTGTA TCTACATGAT	3900
	CAATACTCAC TGTAATTGTC GTACCAAAGA TGTCGGAGTT ATCATCAACC ATTTGTACCA	3960
45	AATCCAAACG TTGTGCAATA TCTTTAGACA CTGGTGCGCA TATTAATCCC CTTGCTTCTT	4020
	TCGCCATAAA ATTAATGGTA TTATCGTTCA TCCATTCACT AACCGCTACT AAATCACCTT	4080
	CATTTTCACG ATTCTCATCA TCTACTACAA TAATTGGTTC TCCATTTTTT AAAGCCATTA	4140
50	AAGCACTGTC AATATTATCG AATTGCATGC TACCCCTCCT AAAAACCAA TGCTCTTAAT	4200
	TTATCTACAG ATAATTGGTC TTTATCTTTA TTTAAAATAT TTTCAACATA TTTAAACAAA	4260

55

	CTCGTTTCTG GAATAAGATG AATGTCAAAA CTGTTATCAT GCTTATCAAA TACCGTTAGA	4380
	CTAACACCAT CCACAGTAAT AGACCCTTGC TTAACCTAAT GATTATTAAT ATGTTGGCTA	4440
5	CATTGAATCG TAATAATTTT TGCATTGGCT GTTTCATTTA TTTTGAAC TGTTCTAGT	4500
	TCATCTACAT GACCGAGGAC AAAATGTCCA CCAAACCTAC CGTTACCACT CATGGCACGC	4560
	TCTAAATTTA CTTCTGATTG TCGCTTAACA TCTGCTAAAT AGGTTTTATT TTCAGTGCCT	4620
10	TTAATTACTT GAACAGTAAA AGATGTCTGA TTAAATCAA TCACTGTTAA ACATGCACCA	4680
	TTAACTCTGA TGGAAATCACC AATATGCATA TCTGCCGTAA TCTTATGTGC TTCAATTTCA	4740
	ATCGTCCTGA CTGATTGACG AATTTGAACA CTTTAAACGA CACCTATTTT TTCAACGATG	4800
15	CCAGTAAACA TGCATCATCA CTTCTTTTCT AAAGTTAATT TAACATTTTG ATTTAATAAC	4860
	TCGGAATGAA CAATTTCAAA TTGGTTCGCA TCTGGTATCT CAATCACATC ATTTGTTTGA	4920
20	TAAAATTGAT AATTTCCAGA TCCGCCAATT AATTTGGGG CATAATAGAG AATAAATTCA	4980
	TCTATATAAT TAGATTGGAG AAATCTGAA GTAGTGGTTG GACCTGCCTC GACTAGCAAA	5040
	GTTCCAACTC CTCTTTTATA TAAATGTGA AGAATTGTTG TTAAATCGCA AGACTTCAAG	5100
25	TAAATAATTT CAATATGTGT TTGATTGGTT GTTAAATTTG GATTTTCAGT ATATATCCAA	5160
	ATTGGTGTG ATTCACTCTG ATAAATTTGC TGATTAAAT GAATATCCC AGACTTAGAC	5220
	AATATTACTT TTATAGGGTT TTTTCCATCT TGAATACGTG TAGTATATTG TGGATCATCT	5280
30	AATTCAACTG TACGTCTTCC AGTTAACT GCGTCGTGTC GATGTCTTAA CTTATAGACA	5340
	TCTTGTTTAA CCTCTTTGTT AGTAATCCAT TGACTTTGTC CATTATCATT CGCTTGTTTA	5400
	CCATCTAAAC TTGCAGATAC TTTCACTGTA ATTTGTGGCA GTTGCTTTGC TTTTGCTTTA	5460
35	AAAAAGTCTT GGTATAATTG TGATGCCCGT TCATCATCAA CGCATTCAAC CTCAATACCG	5520
	TGAGCCCGTA ACGTCTCATC ACCATGTGTG TCTAACGAAT TGTCTTTTGT TGCGTATACT	5580
40	ACTTTTGCTA TCTTACAATC AATTATTTTG TTAACACAGG GTGGTGTGA ACCAAATGA	5640
	CTACATGGCT CTAACGTAAT ATAAATCGTC GCACCTTCAG CATTTTGTG TGCCATATCA	5700
	AGTGCTTGAA CCTCCGCATG CTTGTCACTT TTTCTCAAGT GTGCACCAAT ACCAACAATC	5760
45	CTACCTTCTT TAACTACAAC AGCGCCAACG GGTGGATTAA CACCTGTTTG ACCTGTGACC	5820
	ATATTGCAA GTTGAATCGC ATAATCCATA AATTGACTCA AATGATCACC TCTATAAACA	5880
	AAAATCCTCA CATCATGAAT TAAGATGCAA GGAGAAAAAT TTATCGTTAA ATAAGCCTAT	5940
50	TTGTACACAT TTTTACAAAT ACGCTACATT ATCTTTGTG ATAATTAACA TTCTTTCTCC	6000
	CATCCAGACT TTAACGTGCG GCTCTAGAAT CTCACCTAGAT CAGCCACTAA TATGAAACAT	6060

55

	TTaTATATGA AATTGTTATA GATTATTTGA GTACGTAGTA TGTCAACTAC ATTTAAAATG	6180
	ATACTATATG TTTTCTGAAA AAACAATTAA TGACGGTTTT AATTTAATAT AATCTGAGTA	6240
5	CTATAGGCAT CTCATTGATA TGATTCTTAC TAACAGACAT TAAAATCAAA CCTTCAATTC	6300
	GTCTCTATAG AGCGTTCTCT TTATTATCTT CTAGTTACAA ATTATTGATT GtCACTGCGC	6360
	TGTTGTTGCT CATTGATTTC TAAAGCATCA TATAATTGAG ATACTGTATG CGCAACTTGT	6420
10	TCTACAATCA TTTTCACACC GTTTCGTAGT TTATTAACAC CGTTTGTCAT TTGACCTATC	6480
	GCAATCATAT TTGTTAATGT TCCAAACCTT GGACTAATAA CTTGATTGGT TTCCGGAATG	6540
	ATTTGTATGC CTCCCATGG GTGTGCTTGT ACAATTTGTC TATTTTCAAG ATTTCTAATT	6600
15	AATTGATCAT CTTGATCCAA TTCATTTAAA TGACTTTTTG CACCTGTCGC GTTAATGACA	6660
	ACATTATATA TGTCTACTGA TTCTTGGTTT TTGTATGAAA AATAATACAA CTTGCCATaC	6720
20	ATGTTACAT CTTCTAAATC TTTTTTCAAA ATTAAAGACT TATTTTCTAT TAATTCAATA	6780
	ATTAGTTCAG CAGTTCTTGG AGGCATTGGA TTTGAATTA ATTGAATCAT CTTTGAGTAT	6840
	TTTTGATTAA ATTGATGTTG GTCTTCAATA CTTAAGCTAT TCCATATCCA ATTTAAATTC	6900
25	TCTTTCAAAT GTTCAATCAT ACTTTGGAAA ATGCCCaTTT CTGTTGGACG CGCTAAATCA	6960
	TACTTCAAAT CTGCAATATG ATTTCTGTGTA CGTCTATGTA CTAATTTTTT AAAATCAATG	7020
	TCATATTCAG CACATTCTTT TAAAAATAAA GAACTAAAG TATCAAGCGG TGCATTGCCG	7080
30	AAATGATGTT TTTTAATGTC ATTTAATTTG TCTTTAGTTA AGTACTTGAA TGTACCGTCT	7140
	ATCATTGTAC CTCTTACACT TGGTAAATGA GCAGAACGAC TCGTCATAGT AATTGGTAAT	7200
	TTTGGATGAT GAGCAGCAAC ATAACGGACA ACATCTAAAC TGGCAAGGCC TGTACCAATA	7260
35	ATCGCAATAT CGTCCAGTTC ATTTACTTCG TCTAACGTAT TATATGTTGG ATAAGGCGTA	7320
	gcGATATATC CTTTTTTACC CTTTAAGTTA TATGGATCAT GGTAGGCAAA TGTACCACAT	7380
40	GTAAAAATA CATAATCGTA CGCTTGCCAT GATTGTCCTG AATTTGTAGT ACATATGTAA	7440
	TAAGTTAAAT TCGTTTCATC GATATTAGAA TTTGTATAAA TCTCTGAAC TTTATTATAA	7500
	TTAGTTGATA TATTTGGATA TTTTTCGTG AACATAGATA AATAAGATTT CATATAATGT	7560
45	CCGAATACAA ATCTCGGTAA ATATGCAG	7588

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10320 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

	nCTAGGTATT TTAAACCTAA TCTAGATAAA CTAGCTTCGT AAGCAGCTGC TACATTTTCA	60
5	CGACCGAAAT CCTCAAATA TAATTTTGAA GTAATAAATA AGTCTTCTCT AGCAATACCA	120
	GTTGACTCCA ATCCGGCACG AATGCCAGCA CCTACTTGTT CTTCAATCCC ATAAACTTTT	180
	GCGGTATCAA TACTACGATA TCCTTGTTCA ATGGCATACT TAACACTTTC CATGCAATTT	240
10	TCATCATTTT CCACACGAAA TGTCCCTAAA CCAATTTGTG GCATCGTGTT TCCATTATAA	300
	AATGTTTTAA CCTCCATAAA TATCGCCTCA CCTTTTTGAT GTATTATACC CTGTTATCAT	360
15	AACAAATCTG AGTTGAATAC ATGAGAAAAA ACACTTAGAG CAATCAACCA CTAAAATTCT	420
	AGTAATATCT CTCAAATATT AATCAAATTG TAAAAGTAAT TCTGTTTAAT TTATGACAAA	480
	CTAAAAAGC CGAAGTAACA ACATATAGTC ATCACTTCAG CCTAACATTT AATTGAATGA	540
20	TTCAATTTTA TCCATCATTT GTTGTAAGTC TTCCACGTTG TATTGAATAC GACCATGGAA	600
	TACAAATTTG TTAAAGAACT CGTCTAATTG TTCAGCACCG ACAAGCACTT TGACAGCACT	660
	ATTTTGATTA TAATTTGAAA TCGTTACATC GCCTTCATTT TTAAGATTAA AGTATAAAAT	720
25	TGAAGTTGGT GTATATTTGG CACCTAATTC TTTTGTAG TCTTCAGCCA ATTGTTTAAT	780
	CGCCTCAATT TGATCTGAAT AATTTACAAA TGATAATGAA CGTTTGTGTCAT CATTTTGATC	840
	CATCACAATA GTTTGCGGTC TAGATTTATC TAAATCCAAT GTATCAAATA CTTGTTCCAT	900
30	TGGTGGTAAA TCTTTAAATT GACCGCCACT AATACCATA TAAACATGAC CTTTTAACAA	960
	TTGAGAATCA ATAATATAAA GACCAGTTCT TGTTAATACT AAATGACTAA TTCGTTCAAT	1020
	ATTATTAAAG CCATCCTTTG GTAAAAAGAT ATTTGCCATA ATGTGCATAT CTTCTGGTCG	1080
35	AATTCGTTTT TCTTTAACTA ATCTTTCACG AATACCAATT AATCTCATGT CCGTTACATA	1140
	TTCACTATGA TTTTTCGAGA ACAATTTTAA TGCGTCAATC TCACGATCTT TTGTACTAAC	1200
40	CATGTGATTA TAATCTTCTT GTTGTGTTGT AATTGTCCTT TTATTTTGAA TACGCTCTTT	1260
	CTCTAAAGCT TCTTCATGAG ACTTTTAAAT GTTTTGTCTT TGTGTTCAT ACTTTTCTTC	1320
	TGTTTGTGCG TTAACTTTTT TCTTACTACC TAAGGCAACT AAAAAAGGA CAAAAAGAT	1380
45	TAATGCAATG AgCTACTGCA ATAATGAGTC CAATGACTAT CGGTGAAGAT AAATCCATCA	1440
	CAACAACGCT CCTTTTTAAT ATATGAATAA CTTTAATTAT AATAGAAAG CTAAAGATTT	1500
	TCGATACATA TTATCATTTA TATACCGAAA ATCTTTTATT TAGCTATATT CAATTCATCT	1560
50	TATTATTTTA CTGCGTCTTT TAATCTTCC ACTTTGTCTA ATTTTCCCA TGGGAATAAG	1620
	ACATCTGTAC GTCCAAAATG ACCATAAGCA GCAGTTTGTT TGTAAATCGG TTGTTTCAAA	1680

55

	AGTTGCCCTT CAGAACTTT ACCTGTTCCA AATGTATCAA TTGCAATTGA CACTGGTTCT	1800
	GCAACACCAA TCGCATATGC CAATTGTACT TCACATTGAT CTGCTAAACC TGCTGCAACA	1860
5	ATATTTTTAG CCACATAACG TGCAGCGTAT GCAGCTGAAC GGTCTACTTT TGTAGGATCC	1920
	TTACCACTGA AGCATCCGCC ACCATGACGT GCATAGCCAC CGTACGTATC AACAAATGATT	1980
	TTACGTCCTG TTAATCCTGC ATCACCTTGA GGTCCACCGA TTACAAAGCG TCCTGTAGGA	2040
10	TTGATGTAGA ATTTAGTTTG TTCATTATC AAGTTTTCTG GAACAGTTGG ATAAATGACA	2100
	TGTGCTTTAA TGTCTTCTTG AATTTGTTCA AGTGTACAT CCTCAGCATG TTGTGTTGAT	2160
	ACGACAATCG TATCAATACG TACTGGGTTA TCATTTTCAT CATATTCAAC AGTGACCTGA	2220
15	ACTTTACCGT CTGGTCGTAA ATAATTTAAC GTACCATCTT TACGCACATC TGATAAACGT	2280
	TTTGCCAATT GATGTGATAA ATAAATTGCT AGAGGCATAT ACGTCTCTGT TTCATTGCTT	2340
20	GCGTAACCAA ACATTAAACC TTGGTCACCT GCACCTGTTG CTTCAATTTT TTCTTCGCTA	2400
	TCTTTATCAC GATACTCTAA TGCTTTATCC ACGCCTTGTG CAATGTCAGG TGATTGTTCA	2460
	TCAATCGCAG TTAAAATTGC CATTGTTTCA TAATCATAAC CATATTTTGC TCTTGTGTAT	2520
25	CCAATTTCTT TAATTGTTTC TCTAACAACT TTCGGAATAT CAACATATGT TGTGTAGAA	2580
	ATTTGCGCCG CGATCAATGC CATACCTGTT GTAACAGTTG TTtCACAAAGC TACACGTGCA	2640
	TTTGATCGT CTTTTAAAT AGCATCTAAT ATTGCATCTG ACACTTGGTC AGCGATTTTA	2700
30	TCTGGGTGTC CTTCTGTAAC AGACTCTGAA GTAAATAATC GTTTGTTATT TAACATAGTT	2760
	TGCTCCTTTA AATTTATATT ACGAAAATTC TCTCTCTGTG AGCTAAATAA AAAAGACCTT	2820
	CTAACTATTA ATATAGAGAG AAGGCCTAAT ACGTCCATTC GCTCTTATCG TTCAGACCTA	2880
35	TTTGTCTGCA AACGGTTTGG CACCTTTCTT TTATAAAAAA GAGGTTGCTG GGTTTCATTG	2940
	GGTCCATGTC CCTCCACCAC TCAGGATAAG AGAATCCGTT AAAAATAATA GTACCTAATT	3000
40	AATGAATTAA TGTCAATTTT TCACAAATAA ATTTACAGTA AAATATTGTA GATTAATTAT	3060
	GTTAATGTGT TATACTAATT AAATGTAAAG GCTTACATTT AAATTATCGC TTTGGAGGGA	3120
	TTTAGGATGT CAGTAGACAC ATACACTGAA ACAACTAAAA TTGACAAATT ACTGAAAAAA	3180
45	CCAACGTCAC ATTTTCAACT TTCGACGACA CAACTTTATA ATAAATCTT AGACAATAAC	3240
	GAAGGGGTAT TAACAGAACT TGGTGCTGTT AATGCAAGTA CTGGAAAATA TACTGGTCGT	3300
	TCGCCTAAAG ACAAATTTTT TGTCTCTGAA CCTTCATATA GAGATAACAT TGATTGGGGA	3360
50	GAAATTAATC AACCTATCGA TGAAGAACT TTCTTGAAGT TATACCATAA AGTACTAGAC	3420
	TATTTAGATA AAAAAGATGA ACTATACGTA TTTAAAGcT ACGCTGGTAG CGATAAAGAT	3480

55

	ATGTTTATTA GACCTGAATC AAAAGAAGAA GCTACAAAGA TTAAACCTAA CTTCACTATC	3600
	GTTTCTGCAC CACATTTTAA AGCAGATCCA GAAGTTGATG GTACTAAATC TGAAACCTTT	3660
5	GTCATTATTT CATTTAAACA CAAAGTCATT TTAATCGGCG GTACTGAATA CGCTGGTGAA	3720
	ATGAAAAAAG GTATCTTCTC TGTAATGAAT TATCTCTTAC CGATGCAAGA TATTATGAGC	3780
	ATGCATTGCT CAGCAAACGT TGGTGAAAAA GCGGATGTTG CATTATTCTT TGGTCTATCT	3840
10	GGCACTGGTA AAACAACCTT ATCGGCTGAC CCACACCGTA AACTAATCGG TGATGATGAA	3900
	CACGGCTGGA ATAAAAACGG GGTCTTTAAT ATCGAAGGTG GCTGCTATGC AAAAGCAATT	3960
	AATCTTTCCA AAGAAAAAGA ACCACAGATT TTTGACGCAA TCAAATATGG TGCAATTTTA	4020
15	GAGAACTG TAGTTGCAGA AGATGGTTCA GTGGACTTTG AAGACAATCG TTATACAGAA	4080
	AACACGCGTG CCGCTTATCC AATTAATCAC ATTGACAATA TTGTAGTACC ATCTAAAGCA	4140
20	GCACATCCAA ATACAATTAT TTTCTTAACT GCGGATGCAT TTGGTGTAT TCCACCGATT	4200
	TCAAAGTTAA ATAAAGACCA AGCAATGTAT CATTTCTTGA GTGGTTTCAC TTCTAAATTA	4260
	GCTGGTACAA GCGTGGTGTG ACAGAACCTG AACCATCATT CTCAACATGT TTCGGAGCAC	4320
25	CGTTCTTCCC GTTACACCCT ACTGTTTACG CTGATCTATT AGGTGAACTT ATCGATTTAC	4380
	ATGATGTTGA TGTTTATCTT GTTAATACTG GATGGACTGG CGGAAAATAT GGTGTAGGAC	4440
	GTAGAATCAG CTTACATTAC ACACGTCAAA TGGTAAACCA AGCGATTCTT GGCAAATTGA	4500
30	AAAATGCAGA ATATACAAAA GATAGTACGT TTGGTTTAAAG CATTCTGTGA GAAATGAAG	4560
	ATGTACCGAA AACAATTTTA AATCCAATTA ATGCTTGGAG CGACAAAGAG AAATATAAAG	4620
	CACAAGCAGA AGATTTAATT CAACGTTTGT AAAAGAACTT CGAAAAATTT GGTGAAAAAG	4680
35	TTGAACATAT TGCTGAAAAA GGTAGCTTCA ACAAATAAAT TTGAATACTA AATCaAAACC	4740
	ACCGGTGTGA ACGGGTGGTT TGTCTGCGG CTATAAGCCT TCCTTACTGG CCAGCCCTAA	4800
40	AAGGGCACTG ACAAGTCAGC CAACTGCACT ACTATTCCAG CAACCCTAAA GGGTTACTCT	4860
	TTTTTCTTTC TTTTTTTATT TTTCTCTCCA GTGAAAGGAT CTAAATATTC TTCCATTGAG	4920
	ATTTGGTCTG CAACGATATC CTCTTGTAAT TGATTACGAA TATAATTTTC AATCACTTTT	4980
45	TTATTTCTAC CTA CTACTGTATC CACATAAAAT CCTTTACACC AAACTTTCT ATTTCCATAT	5040
	CTATACTTTA AGTTAGCATG TCTATCAAAT ATCATTAAAC TACTTTTTC TTTTAAATAG	5100
	CCAACAAATG ATGATACCCC AAGTTTGGGT GGTATACTAA CTAACATATG GATATGATCT	5160
50	TTACATGCCT CTGCTTCAAT TATCTCTACA CCTTTTCTTT CACATAATTG ACGCAATATA	5220
	ATCCCTATAT CTTTTTTTAT TTTTCCATAT ATCACTTGTC TTCTGTATTT AGGTGCAAAG	5280

55

AAATAGCATC TCCTCGTGT GATTATTTTG GTTGGCTGAC CAATATTTAT TCTAGCACGT 5400
 AGAGATGCAT TTTTGTGAC AATGGTAGAA CCTTTTctGa ACCATACGCA TAGCGTATGG 5460
 5 TTTTCTTTTT ACAATTAAAG AGCCAACCGT TGTTATAGTC TAACAATGGT TGGCTCCTCT 5520
 TATTTTATGT GCTAAAAATT TATAGGCAAT TTTATTACAA CAATGTACAT TTAAGGTGAC 5580
 CTTTCATGCCA AAATCGCATC ACTCATTAA TGGAAGCAGC ACGTCTTCAT ATAAAGTACC 5640
 10 GATCCCTAAT TCAACGCATG TAGTACCACA TCTTCAAAGC TTGATAGTTC CCATGCGCAC 5700
 ACCACGTTTC ATACTAGCTA TGCGACTCAA CTTGGTTCAT AAACCTTTTA ATATAAGTCA 5760
 ATGTTTCAAC CATCGCTGGT GGTCTTGGCA CATGTCCTTC TGCCATTTGA TAAAATGTTT 5820
 15 CATGCGTGGC ACCTTTTAAC TCTAGTTGGT CCGCTAAATA ATACGCATGA TGAATACCAA 5880
 CTTGCTGGTC TTTCCCTCCA TGTACAATTA ATATTGGCGG ACTGTTTTCA TTAATGTTTG 5940
 20 GAATCGCTTG GCGTGCCTCA TATGCCGCTC GATCTTTTTT CGGATGACCA ATCATTCTTC 6000
 GTAGCATGCC TCTTAAATCG ACACGTTCTT CATACTTAA ATCAATATCT GAGACACCAC 6060
 CCCAGATTGT ATAAGTTGTT ACTGGTAAGT CTTGAAATGT CAACAATCCT TGTAAACCAC 6120
 25 CTCGCGAAAA ACCAACCATG TGGATAAATG CATGTGGATA TTTATCATGT AGCAACCTTA 6180
 ATAATTGCGT CACATCATTT AAATCGCCAC GGTAAAATTC GTCTTTGCCT TCACTCCCAT 6240
 TGTTACCTCG GTAGTATGGC CCAATCACTA AAGTTTGAAT ATCTGAAAAT TGCATTAATC 6300
 30 TACCTGCGCG CACACGTCCT ACTTGACCTT TGCCACCTCG CAAATAAACT ACAATGCGAT 6360
 TTACTTCATG ATGTGGTGTC ATCATTAAG CTTTACTTGG TAAGTCATCT GACAAATATG 6420
 TAATTTCTTC GAATTGATGC GTAAAATATT CAATTGGCAT TCGTTTACGT TTGATAAAAC 6480
 35 CCAAGTGATT GCACCCTCTC TACGCATTTT AAAATGGTAC TATCTTGCAG TAAGAAACTC 6540
 CGTTGTGCGA GTTCAATATC ATTGATACAG TTAACAACA CTGGCCCTGC TGTCTCTAAA 6600
 40 TAATCGTTCT TGCTTACCAA TGATTCAACT TCGATAAAAT ATACATCTTT TACAAAATCA 6660
 GTTTGATCAT GTGTTTCAAT GGTATATTGT GCTATGTAAT AAATATTTTT AACTTTGGCG 6720
 CCTGTTTCTT CATATAATTC aCGTGTAAGT GCTTCAGCAC TACTTTCCCC GCGTTCCCTT 6780
 45 TTACCACCAG GAAATTCAAT CCCCCGTAAA TTATGTTTGG TAAAAAGCAA TTGATTTTTA 6840
 AACGTTGGAA TAGCTAGCAC ATGATTGCCA TCTGCTATCT CATTATCCTT TTTAAATGTC 6900
 AAATTAACCT GACGATTATC TTTATCCCTA AACTTCACGC GCATCACATC CCTACATTGT 6960
 50 ATGTTAATAT AATAGTTAAT TACTATCGTT GGAGGCATTA ATTATGAAAA AGATATTCTT 7020
 GGCGATGATT CATTTTTATC AACGTTTCAT TTCGCCACTC ACTCCACCAA CTTGTCGTTT 7080

55

	CCTTTATTTA GGTATCCGTC GTATTTTAAA ATGTCATCCG CTTCATAAAG GCGGCTTTGA	7200
	CCCTGTTCCG TTAAAAAAG ACAAGTCAGC AAGCAAGCAT TCACATAAAC ATAACCATTA	7260
5	ATATGGTTGT AATTGAGTTA TATCCACTAA AGGGGGGCGA AATTCGAGTC GCCCCCTTTT	7320
	TAATATGCCT GAATGCGCCA CCACATCTTG TTCAAAATAA TAACCTGCTG GTGTAACATC	7380
	TCCTGGATAA TCACCTTTAC GAGCAAGCAT CGCTGTAAAA TAGCGGCTTA AACCATATTC	7440
10	GTACATGCCG CCAATAACCA CTTTTGCACC ATGACTTTTC AAAGTATCAA TTGCCGTTTG	7500
	CACTTTATCA ATGCCACCTA GACGAAATGG TTTTAATACA ACAACTTTCA CATTGTATAA	7560
	TTCTATCAAA TTAATTATGT CCaACAACGA TGTTCCTTT TCATCAAGGG CTATTGGAGG	7620
15	TATTGTTCa TCCGCTACTT CATCAAGCAT GGAGATATCT TTAATGGCT CTTCGATATA	7680
	AAGAACCTGT TCACGCGCTA ATAAGTGTAA CTGTGTGAAA TCTTGACGAT CCAAGGACTC	7740
20	ATTTGCATCT ATAACCAATT GAAAGTGAAA GTCTAATTCC CGTAACACTC TAATTGATG	7800
	CATGATTGA GCGCTCCATT TTAATTTAAT TCTGGTCGGC TTTGTTGCTT TTAATGACTC	7860
	TAGTTGTTTA TTTGATAAGC CGCTCGcTGT CGCTCCATAT GCTACTGAAA ATGAAGGCAG	7920
25	TACATGAAAC ATTTGATACA ATGCCATGAC AATAGTTGCC CTTGCAGCAG GCGTATTTTC	7980
	CAATGAATCT ACTAATTTTA GTGCTGCTTC ATACGTTTCA AATGATTTAT TTCTATTATC	8040
	TTCGAACCAT TGCTCAATTA CATGTTTCAC TGAGGCAATT GTTTCATGAT CATACCAATC	8100
30	TGTTTGAAAA GCGTTACATT CCCCgAAATA TGCAATTCCT TTGTCATCAA TCAATTGAT	8160
	AAACAAACAA TCACGATGCG TTAAAGTGAC TTTGGTGTG ACAATTTGTG ACTTAAATGG	8220
	CTCACTATAT TTATAAAAT GCAAAGCTGT CAACTTCATC AAATCATCCT CTATACAACT	8280
35	TATTTCTTTG TAATTTACCT GTTGATGTAT AAGGTAAAGT ATCAACCTTT TCAAAGTGT	8340
	TCGGTACTTT ATATTTGCT AAATGTTGTG ATAAATATGC AATCAATTGT GCCTTTGAAA	8400
40	TGTCACTTTC ACTGACAAAA TATAATTTAG GCACCTGGCC CCAAGTATCA TCAGGATGCC	8460
	CTACACATAC TGCCTCACTG ATACCTGGAA ATTGctTCGC TACCGTTTCA ATTTGATATG	8520
	GATAAATATT TTCACCGCCA CTAATAATTA AATCTTTACG TCGGTCATAA ATCATGACAT	8580
45	AACCTTCATG ATCTATTTCA GCAATGTCAC CCGTATTAAA ATAACCATTT TCAAACGTAC	8640
	CCGTAAATC TGTGGATAC AAATATACAT TCATCACATT GGCGCCTTTA ATCATTAAAT	8700
	CTCCATGACC TTCTTTATTA GGATTTTAA TTTTACGTC AACATTGGCA CTTGGCATCC	8760
50	CTACAGTGTC AGGACGTGCA TGCAACATTT CCGGTGTTGC TGTTAAAAAT TGCGAACATG	8820
	TCTCAGTCAT ACCAAATGAA TTATAAATTG GCAGGTTATA TTGTAATGCC GTCTCTATCA	8880
55		

AACCTTGTTG CATAAGCCAA TTAAAGTTT GTGGCACAAG CGAAATGTGC GTGATTCGTT 9000
 CATTTTTAAT CATCGTTAAA ATTTGTTTCG CATTGAATTT ATCAACAATG CGCACAGTAA 9060
 5 AACCTTCAAT AACAGCTCTT AAAAGTACAC TGAGACCCGA AATATGATAA ATCGGCAAGA 9120
 CAGATAGCCA ATTAGTGTCA CGATCAAATC CCAAGCTCTC TTTACATCCG ATTGCACTGG 9180
 CATAATGATT ACGAAACGTT TGTGGCACCG CTTTTTGAGG GCGCGTTGTC CCTGATGTAA 9240
 10 ACATAATCGA TGCAATGTCA TCTAAATTAA ATGATGTATT TAATATGTTG GACGGCGACT 9300
 CTTTCGGCAC CACAGTTTCA TTCGATGTTT CATATTGGAT ACCCATTTGTG TTGTCCAACA 9360
 AACTGTTTCGTTGTAATATCC CTTCAGCGA ATTCAATATC ATCCAGCGAT ACAATTTGAA 9420
 15 ACCCTCGTAA TTCCAGTGGC AAGGTACAAA AAATCAATTG TACATCGATT GACTTCATCT 9480
 GATTCGTCAT CTCATTAGGT GTCAACCTTG TATTAATCAT CGCAATTTCA ATATTTGCCA 9540
 20 ACCAACATGC ATGTATTAAA ATGATCGATT GAATCGAATT ATCTATGTAT AGCCCAACAC 9600
 GAGATTGTTG ATAAGCCTTG AGTCTTTTAG CCAATAGACT CGCTTCACAG TATAAATTTT 9660
 GATAAGTATA AGATTCTTGA CCGTCTGTTA TCGCAATATG ATGTCCATTT TGTGTGCTT 9720
 25 GTTTATATAA CCAAAGTCC ATGCGTTATT CCTCCAAAAT CATTTACATT ATAATTATAA 9780
 CGATTTTATG ACATTCTAGC AGTGGTTATG TTTAAAAATA TAAAAAGTA GACGAATTGA 9840
 TGCATTGATA TGATTGTTAT AATGCTCAAT ACATATCGTT ATATCATTCG TCTACTATTA 9900
 30 TCAGTTATTT TTATTTAATT TTAGTGTTCAT TCTGTCAATTT TGATGTGGTG ATTTACCCAT 9960
 TGTGCCCACA TCATCTGCAA TGTCAATTGG TATACGGTTC ATGTCTTGTA ATGCACTTAA 10020
 ATGGAATACT TCATCATCTA AATTTTCAAT GAGATATACA TAATATGTTA CCTTGTCTTT 10080
 35 TTTATATTTT AACGTTTTCC AAAAGTCCGG CTTGCAATTC AATACATTAT CCGGAATATA 10140
 TTCAATAAAT AAGTAACGTT TGCTGCCTAC TTTGTCTATG AAATATTTTG CAGTGCCTTT 10200
 40 TTCTATACCT CTTATATGTG CATAGTCTGC TGAAAAGTAA ATACTACCTA TTGTTTCATT 10260
 ATGTTGTTGT ATTTCAAATC GTTGGCCTAC TATTTTATTA TTTGTGCTAC nGGGGACTTA 10320

(2) INFORMATION FOR SEQ ID NO: 144:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1477 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

GTGTGGATTG GATTTTAAAA TCACCCTCAT AAATACTGTC ATCAATATGA TAAGTTACAA 120
 TTTCACCTAT TATTAAATCA GCCCCATCTA ATACATCTCC AAGCAATATC ATTTGCGmTA 180
 5 GTTTACATTG GAATCTCATT TTCGCATCTT TAATTCCTGG CGTCTTAATC GTTGTAGATG 240
 TTAAAAGTGA TAATTCTGTA CGACTCAACT CACTGTCACC ATATGCTAAC GGCGCTGCAG 300
 TCTCATTAAT ATCTTGAACA TTATCTTCGT CTGTAATATG CACAACAAAG TCTCCAGTCC 360
 10 GTTCTATATT TAATGCAGTA TCTTTTCTCT TACCTCCTGC ACGTTGAACT GCAATAGCAA 420
 TCATTGGCGG ATGATTATTA ACAATATTAA AAAAGCTAAA TGGTGCTGCA TTTACTGATG 480
 CATCTTGATT TAATGTTGTA ACAAAGCTA TAGGTCGTGG AATAATTGAA CCAATTAATA 540
 ATTTATAGTT TTCTCTAGCA GTTAATGATT GTGCATCAAA CGTATACATA ATACCTACCT 600
 CTTTTCTAAG TATATCTAGG TATTTCTCCG ATTTTGTTA ATTTAAACAT CTATTCTCCT 660
 20 CTGAAAATCA CTGTATTTA TTTAGCAAAT CTTTTGAAAT ATGACACATA TGCATATCTT 720
 CTGGATATTT TTCTAAATGT TGCTGATGTT CTTGAGCACT TTTAATGTAG TTAGACAGCG 780
 GTAAGACTTC CACTGCAATT TGATCTCTGT CTTTACGTCG TTCAATGAAC TGACGCGCTT 840
 25 CAATTAAGTG GTCATCTACA CAACTATATA AACCCGTTCTG ATACTTTTGT CCAATATCAT 900
 TTCCTTGTTG ATTCACACTG TAAGGATCAA TGATTTCAAA TAAATAATTC ATAATGTCTG 960
 TAATTGTTAA CATACGATCA TCGAAATGAA GTTTGACACA TTCAGCATAA CCATCATACG 1020
 30 GACCGTCTAA TTTAGAGCTT CTTCCATTTG CTCTTCTGTC TTCTGTATGT ATAATTCCAG 1080
 GTATTGTTGC AAAAAATGCT TCAACACCCC ATAAACATCC TCCTGCTACA TAAACAACTG 1140
 CCATATTTAC ACCTCATCAT CCTTTTTTAT ATTTTAAACA AGGTTATACC ATTTAATACC 1200
 35 GCCATGACAT GATTCTGATA CACCTTCATT ACGATACCCA TATTTTTCAT AAAATGAAAT 1260
 TAATGATTCT CGACATGTTA ACGTTACACC ATGTCGATGA TGATTCTTAG CAAGAGTTTC 1320
 40 AAAATAGTTT AGTAAGCGAC CTGCAATACC CTGACCTTGA TAATTTGGTG CTACAACAAG 1380
 ACCTAACACA CTAATATAGC CACCTTCACT ATTATTTGTG GAGACATTTT TAAATAAATC 1440
 ATCGCTAATG TAACGCTCTT TTATGACTGG ACCGTTG 1477

(2) INFORMATION FOR SEQ ID NO: 145:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3976 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

	AGGTGATTAT CCTAAAAATG CTCATGAGGT CGCTATTAAT GATAAGTTAG CTGCAGACAA	60
	CATTAGAGTC GGGGATAGAT TACATTTTAA AAATAATTCA ACTAGTTATA GAGTTTCTGG	120
5	TATTTTAAAC GACACAATGT ATGCGCATAG TTCCATTGTG CTATTGAACG ATAACGGATT	180
	TAATGCATTG AATAAGGTTA ATACGGCATT TTATCCAGTG AAAAATTTAA CACAACAACA	240
	ACGTGATGAG CTTAATAAAA TAAATGACGT TCAAGTTGTG AGTGAAAAAG ATTTAACAGG	300
10	TAATATTGCG AGTTATCAAG CAGAGCAAGC ACCGTTAAAT ATGATGATTG TTAGTTTGT	360
	TGCTATTACA GCAATCGTTC TAAGTGCAAT TTTCTATGTT ATGACGATTC AAAAAATATC	420
	ACAAATTGGC ATTTTGAAAG CAATTGGTAT TAAGACAAGA CATTTATTGA GTGCGTTAGT	480
15	TTTACAAATT TTAACACTAA CAATAATTGG GGTAGGTATT GCTGTGATCA TCATAGTAGG	540
	ACTATCATTT ATGATGCCCG TAACGATGCC TTTTACTTA ACAACGCAA ATATTTTATT	600
20	AATGGTGGGG ATATTTATAT TAGTAGCGAT TTTAGGTGCC TCACTATCAT TTATCAAATT	660
	ATTTAAAGTG GATCCTATCG AAGCAATTGG AGGTGCAGAA TAATGGCATT AGTCGTTGAA	720
	GATATCGTCA AAAATTTCCG AGAAGGTTG TCTGAAACAA AAGTTTTAAA AGGTATTAAT	780
25	TTTGAAGTGG AACAAGGGGA ATTTGTCATT TTAAATGGTG CCTCTGGTTC TGGGAAAACA	840
	ACATTGCTAA CGATATTAGG CGGATTGTTA AGTCAAACGA GTGGTACAGT GCTTTACAAT	900
	GATGCGCCAT TGTGATAA ACAGCATCGT CCTAGTGATT TACGATTGGA AGATATTGGT	960
30	TTTATTTTTC AATCTTCACA TTTAGTTCCT TATTTAAAAG TGATAGAGCA ATTGACACTC	1020
	GTAGGTCAAG AAGCGGGAAT GACCAAACAA CAAAGTCAA CAAGAGCAAT ACAACTTTTG	1080
	AAAAATATTG GTTTAGAAGA TCGCTTGAAT GTATATCCGC ATCAGTTATC TGGCGGTGAA	1140
35	AAGCAACGTG TTGCGATTAT GAGAGCATTT ATGAATAATC CGAAAATCAT TTTAGCAGAT	1200
	GAGGCCACAG CAAGTTTAGA TGCCGATAGA GCAACAAAAG TTGTTGAGAT GATACGTCAA	1260
40	CAAATTAAAG AACAACAAAT GATTGGTATT ATGATTACAC ACGATCGAAG ATTATTGAA	1320
	TATGCAGATC GAGTGATTGA ATTAGAAGAT GGCAAATAA CTGATTAGTG GCTTGTAAG	1380
	ACGCTAAATG TTAATGATTT AAGACATAGT AGTATAAAG TTAGATAACA GAATACGATT	1440
45	TGGGTTTACA AAAACAGGC TGGGACATTA AGTTCCTAGG CAATGTAAAA AAGCTGATTT	1500
	CTATTAAATTA TTTGATAGAA ATCAGCTTTT TTGATATGTA TTTTATAATG TACAGCTCGT	1560
	TGCATTATA TAGCTGAAG TCACGTTTAA AACCATATCT ATCATTATGG TATGCATATC	1620
50	TTTTAAAACC TATTCCTTTG TTATTAGGAC ATATAAATTC ATCATTAAAGT TCGTCATATT	1680
	TCCAATTTTG AGTGTTAAAA ATGTCACITT TAACTTTCT AGTTTTATCT TTAATAAACA	1740

55

	CACTATCATA ACATGCATCA GCTACAATAT ACTCCGGTAA ATAACCGAAG nTATTTTgAA	1860
	TCATTGTTAA AAATGGAATT AAAGTTCTAG TATCTGTTGG GTTTTGAAAT AGGTCATAGG	1920
5	ATAAAACAAA TTGAGAATTT GTCGCTATTT GTAAATTGTA TCCTGGCTTA AGTTGGCCAA	1980
	AGTGTCTTAT TTTTTTAAAG TATTTAAAAG TAAAATTACA TGTTAATACG TAGTATTAAT	2040
	GGCGAGACTC CTGAGGGAGC AGTGCCAGTC GAAGaCAGGG GCCCCAACAC AGAArcTGAC	2100
10	ATATAGTCAG CTTACAACAA TGTGCCGGTT GGGGTGGCTG AGACGGCACC CTAGGAAGGG	2160
	ACCCGTCATC AAAAATTCTA TTTATAGAAT TTTACAGTAA TGTGCCAGAT GGGCATAGCG	2220
	AAGcCATTCA ATACGAAGTA TTGTATAAAT AGAGAACAGC AGTAAGATAT TTTCTAATTG	2280
15	AAAATTATTT TACTGCTGTT TTTTTTAGGG ATTAATGTCC CAGACTCTTT AGTTTATTIA	2340
	TTTTCAATAT AACAAATTGTC TAATCAAGGA TTAACGAATA TTAAAGATA GTTTGACGCA	2400
20	ATATTAGAAA CAACCTATAA TAATAGTTTG TTTGTGGATT AACTATTATA AATAAAGCG	2460
	GCGTAAAGAC ATATAAACCA ACTACTTGAA CAATATAACG TTAATAACAA TCTATACTGA	2520
	TACATTACGC CTAGATAATC TTTGATGAGC ACATGTAAGA AAAAGTGATA TGGTGTATGA	2580
25	CTTCCGACAC CATCGATAGA TAAACCTAAT TTTTGGGCTA GTCGTAAGGC GCGCAATACA	2640
	TGAAACTGAC TTGTTACACA AACAAITTTA ACTGCTTCAT GATACAAATT GTTGATGATT	2700
	TGTTTAGAAT ATAAAAAGTT TGTGTATGTA TTTATAGAGT GAGATTCCAT TAGTATATCT	2760
30	GTTTTATCAA CACCATGTGC AATCAAATAA CGTTGCATAG CTAAAGCTTC AGAAATTGGT	2820
	TCGTCTGGTC CTTGTCCGCC AGATACAATG ATCTTTGTTG CTGATGCTTG TTGTTGATAG	2880
	ATATCAAGTG CACGATCTAA ACGCGCTGCA AGCATTGGTG TGACAAATTC GGTAAAAATA	2940
35	CCAGCACCTA ACACAATTAT GATATCAACT TCTTTGTTGT ATGATCTATG TCTATATGAT	3000
	ACTGFCCAAA CGAGATAACA AATAAAGGTT AGTAACAGGG AAAGACATAA TATAGCTAAC	3060
40	CACATAGACA AACCTTTCAC AATAGGTGAC TGAATCGTAC TTATAAATAG AAGTGCTGAT	3120
	GTGTAGAGTA CAAATTTATA TGAAAAAGAT AATAATTTTT TAATAAATAA GCGACTAGAA	3180
	GTATGAGAAA ATAAATATCT ATGTTTGAAT AGCATGATAA TACTGATTAT TATAAATGTT	3240
45	ACAAACATAG ACCAAGGGAA AGTATAGGTC ATGATGCTAT AGATGAGTGA CAAAAATATC	3300
	GATATGACAA CTAAGATGTA GCATGTTAAA TTAAAGTCA GAGTATAGTT GAAAATTAAC	3360
	GGACAAATAA CGATAAGTAT AAATATTAAT AATAAATTCA ATAACATACT GACACCTCGC	3420
50	TTATAATAAA TATTAAATAT AATGTAGAT GATTTAATTT ATTAAAGCAA GGAGAAAGCA	3480
	GCAACATGTA AATCTTAATT TGTATATTA TATATGGGTC AATATTTTTG TGTTTTTTAG	3540
55		

TATGGTAAAA CATTTACAAG ACCATATTCA ATTTTITAGAG CAGTTTATAA ATAACGTTAA 3660
 CGCATTAACT GCAAAAATGT TGAAAGATTT ACAAAATGAA TATGAAATTT CATTAGAGCA 3720
 5 GTCTAACGTA TTAGGTATGT TAAATAAGA ACCTTTGACA ATTAGTGAAA TCACGCAAAG 3780
 ACAAGGTGTA AATAAGGCCG CAGTAAGCCG ACGAATTAAA AAGTTAATCG ATGCTTAATT 3840
 AGTTAAGTTA GATAAACCAA ATTTAAATAT TGATCAACGT TTGAAATTCA TAACCTTAAC 3900
 10 TGACAAAGGT AgAGCATATT TGAAAGAACG TAATGCGATT ATGACAGATA TTGCGCAAGA 3960
 TATTACTAAT GATTTA 3976

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3346 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

25 GCTACCTAGG CATTTAAGAG ATCAAAAAAT GTATGAATAT GAACGTTATT TTTATGAGCA 60
 AGAACTTAAT GCGCTTGATG aAGGGGAAAT TTTAAAGAAG TTAAAAGACC CACAAGATGT 120
 TGCAGCTGAA ACAAAGCTA GAAGTGTTAT TGATTATGCT GAATCTAAAC CAACATTTGA 180
 30 AAATATTTCA AGAGCTGTTG CTGCTTCATT AAGTTTAGGC ATTCTATCTA TTTTGTTCAT 240
 CCTTATACCA GTATCTATAG TTGGATTATT TGTATTAGCA TTATTTTAA TATCACTTTT 300
 GCTGCTGTTT TGTCCAATTA TTTTATTAGC ATCAGCAATA TCCAGAGGAA TTGTGGACTC 360
 35 AATTAGTAAT GTATTTTTTG CCATATCATA TTCAGGATTA GGATTAGTAT TTATCATTGT 420
 CATATTTAAG ATTTTAGAAT ACATTTATCG TTTAATCTTA AAATATTTAC TTTGGTATAT 480
 40 TAAACTGTC AAAGGAAGCG TTAGAAAATG AAGAAATTCT TTTTATTGG GCTTTTAGTG 540
 TTTGTTGTCT TTTTACAGC AGCAACCATT ATTTGGTTCA GCTATGATAA AAACAAATAT 600
 GGTACTAAAC AATATGATAA AACATTCAAA gACGATGCTT TTGACAATGT ATCTATAAAT 660
 45 TTGGATAGTA CAGAACTTCG TATAAACCGG GGAATCAAT TTAGAGTTAA ATATGATGGT 720
 GACAATGATA TATTAATTAA TATAGTAGAT AAGACGTTGA AGATTAGTGA TAAAAGGTCT 780
 AAGACAAGAG GATATGCAAT TGATATGAAT CCTTTTCATG AGAATAAGAA AACGTTAACG 840
 50 ATTGAAATGC CTGATAAAAT GATTAAACGT TTAAATCTAT CATCTGGAGC AGGAAGTGTT 900
 AGAATCAGTG ATGTTGATT AGAGAACACA AGTATTCAAA GCATTACCGG TGAAGTAGTT 960

	AGTAAAAGTA ACATTAAAAA TAGCAATATT AAAGTTGTTA TTGGTACGCT ACAAATCGAC	1080
	AAGAGTCAAA TTAAACAATC CATATTTTAA AACGATCATG GTGACATTGA ATTTAAAAAC	1140
5	ATGCCATCAA AAGTAGATGC AAAAGCTTCT ACTAAACAAG GAGATATTCG TTTTAAGTAT	1200
	GATAGTAAAC CTGAAGACAC TATACTAAAG CTAAATCCGG GAACGGGTGA TAGCGTAGTT	1260
	AAAAATAAAA CATTIACATA TGGTAAAGTT GGGAAAAGCG ACAATGTTTT AGAATTTTAT	1320
10	ACGATTGATG GTAATATCAA AGTTGAATAA ATAAAGGATG TAAGCACCGA TATTAGGAAG	1380
	CATAATTTCT CTAATATCGG TGTATTATTT TTGTTGGCAA AAGTTAAGTC GGTATCTATA	1440
15	TTGCCAGTAA AGTGAGTGAT ATTAAGGTCT TGACCATCTA ACCATGATTT GAAATCTATT	1500
	ATTTCTGGTG GCGCATTTTC TCCCAATGTA AAATATGCAG TTAATGTTTC AGGTTGATAC	1560
	ATTGATGTAT GGATGGTGCC AGACCAGCTT TTGAATAGTT TACTGTAAAT TTCATACTGA	1620
20	GGATTATTGA ATAACTTAAA TGCTGTAGTC ATATCTAAAT TATCATTAGT TTGTGAAATG	1680
	GTACGCGCCA GTCTTTCTTT AGATTCTTTT GTATAATTAC GATTTTCATG TGTTAATAIT	1740
	TCAAAATGAT TTGTACATAT ATTATCATAA CGAACATCTA TTGATCTCGG TGCTACTTCA	1800
25	ACAATTGCAT GGTTCATGA TTTGTCCATC AGTATGTAGC TAAATGAGCT TCTGTGTGGT	1860
	ATTTCTTTCA ATAAATTGGAT TGCTTCTGTT ACATTTCCGGC AATTTTCAAG AATTAGACGA	1920
	CCAATCATAT AACATACAAA ACCATTTGCT GGTTCCTTCC GGTGCATAAA GTTATAGCCC	1980
30	ATAGTTAATC CTGACTCATT CATACCATCC ATTCTTCCAG TTACCCTTGA TACAGGACCA	2040
	ATTTGAGCTA AACCGCTATC TGTAGGTTGA TAAAGTAAGT AGCGACCATC ATAAGTTGCA	2100
35	GGGTGGTAAT CATAATTTCT AACCATGAAG TCTTTGCCTT GAAAGACCGT GCAaCCACTT	2160
	TCTTTTAAAT CGGTAAAACG ATAATGTCCA AAGTTTAAAA TAATTTGGCG TGTGGCATT	2220
	TTGAGTATAC TTTGTAGTCC CATTAATTCT TCCCATATTT GAGGTGCGTA TGTTTGGAAT	2280
40	ATTTGATAAG TTTCATTTAC ATCTATATCG AAACGTGGGA CaCnTTTTTT CCATTCTTTT	2340
	TCTCGATTTT TTAGAAGAGG TGTTTGTTGA AGCCATTTAC CAGTTTTAAC ACCTAACTCG	2400
	AAATGTGAAC CTCTAAAAGT CATGATATCT GATGTCACTT GTTGCATATC ATCGGCCCTT	2460
45	TTCTTTTTAG TTGTAATATA TTGTAAATAA ATAGTAATCG TATGTATATT GAATGTCATG	2520
	TTAAATAAAG TTATATTTTA CTAAATGAAA TATAAAATTG TTTGAGGTGA TTTCTCGGTG	2580
	TATAAGACTT ATCAATCAGT TAAAACATAT TTTTATAGAT GGTGGGGATA TTGAGTTAAA	2640
50	AACTTAAAAT CATCTTATCA TAAATATCAA TCTTAAGTTA GCATTCACGA TAATAGTCAT	2700
	TGTTAACATT AGCATATAAG GTCATGTCAC GTTGAAACAG AGGTTCCCTCG GCATTTTGA	2760
55		

TTATTTAATG ATTATTCTAT ATATGATAGT ATAATGAAAT GTAGATAGGT ATTTAATTTA 2880
 ACAGAGGTGA AATTGAGATG TGGAAATTTA TTAATGtGT GkTTAAATTC GTATTTAGCT 2940
 5 TAGTTGCTAT TACAACATTA GTTGCTGGTG TTGGTGTAGT AGCATTTGCT TATATCTTTA 3000
 AAAAAGATTT TGAAGATATT GAAAGAAAAA CTAAAGAAAT TATTTCTGAT ATTGAAAGTA 3060
 AAAATAACTA ATAACATTTA GAGGCTGGGA CATAAATCCC TAAAAAACAG CAGTAAGATA 3120
 10 ATTTTCAATT AGAAAATATC TTACTGCTGT TCTCTATTTn ATcAmTACTt CGTATTGAAT 3180
 GGCTTCGCTT TCCTAGGGTG CCGTCTCAGC CTTGGTCTTC GACTGGCACT GCTCCCTCAG 3240
 GAGTCTCGCC ATTAATACTA CGTATTAACA TGTAATTTTA CTTTGGAAT ACTTTTAAAA 3300
 15 AATAAGACAC TTTGGCCCAA CTTGGCACAT AAATGTAAAA TTCAAT 3346

(2) INFORMATION FOR SEQ ID NO: 147:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2375 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

GTTGAAGAAA GAAATATAAC AGTCAATTAT AATTATAACC TTGTTGAAAT CGACGGTGAC 60
 30 AAAAAAGTGG CTACATTCGA ACATATCAAA GCATACGATA GAAAAACAAT AAGTTATGAT 120
 ATGTTACATG TAACACCACC TATGGGTCCC TTAGATGTAG TAAAAGAAAG TACACTTTCA 180
 GATAGTGAGG GTTGGGTAGA TGTTAACCCA ACCACATTAC AGCATAAAAG CTACTCTAAT 240
 35 GTATTTGCAC TTGGTGATGC TTCAAATGTA CCTACTTCAA AACAGGCGC ACTATTcGTA 300
 AGCAAGCACC TATCGTCGCT AATAATTTAT TGCAAGTGAT GAATAATCAA ATGTTAACGC 360
 40 ATCATTATGA TGGTTATACT TCATGCCCTA TTGTIACTGG ATATAATAGG TTAATACTTG 420
 CAGAGTTTGA TTATAATAAA AATACTAAAG AAACAATGCC GTTTAATCAG GCCAAAGAAC 480
 GTaGAAGTAT GTATATATTT AAGAAAGATT TATTACCTAA AATGTATTGG TACGGCATGC 540
 45 TAAAAGGATT AATATAATAA AGTACAGAAA ACAATAAATT TTTAATGAAA AATCTTTTAC 600
 TATAAAAGAT TAAGTATTTA AATGACGTGT CAGTGTGTGT TTTATATGTC GTGAATTTTT 660
 AGCTCTAAAT AGTATAAGAT TGAAAAAGTT GTTACTGTTT TAAATGATCA CGATGAAGTC 720
 50 ATTCAATAAG AATGATTATG AAAATAGAAA CAGCAGTAAG ATATTTTCTA ATTGAAAATC 780
 ATCTCACTGC TGTTTTTTAA AGGTTTATAC CTCATCCTCT AAATTATTTA AAAATAATTA 840

55

AGATATTCAA ACCACGTGTA CTCAAAATGA TAGCTTGGTA TGTACCTCCA ATAGTAATTT 960
 CAATAACTTT GTCTGTTGAA CACTAAGAGC AATTTTAATT TCATAATGTG TTGTAAACAT 1020
 5 TTTTTTTGAT TGGAGTTTTT TTCTGAGTTA AACGATATCC TGATGTATTT TTAATTTTGC 1080
 ACCATTTCCA AAAGGATAAG TGACATAAGT AAAAAGGCAT CATCGGGAGT TATCCTATCA 1140
 GGAAAACCAA GATAATACCT AAGTAGAAAG TGTTCATATC GTGTTAAATT GGGAAATATC 1200
 10 ATCCATAAAC TTTATTACTC ATACTATAAT TCAATTTTAA CGTCTTCGTC CATTTGGGCT 1260
 TCAAATTCAT CGAGTAGTGC TCGTGCTTCT GCAATTGATT GTGTGTTTCA CAATTGATGT 1320
 CGAAGTTCGC TAGCGCCTCT TATGCCACGC ACATAGATTT TAAAGAATCT ACGCAATCTC 1380
 15 TTGAATTGTC GTATTTATC TTTTCATAT TTGTTAAACA ATGATATATG CAATCTCAAY 1440
 AATCTAATA GTTCyTTGCT TGTGTGTTCT CGTGGTTCTT TTTCAAAAGT GAATGGATTG 1500
 20 TGGAAAATGC CTCTACCAAT CATGATGCCA TCAATACCAT ATTTTCTGC AAGTTCAAGT 1560
 CCTGTTTTTC TATCGGGAAT ATCATCGTTA ATTGTTAACA ATGTGTTTGG TGCAATTTCTG 1620
 TCACGTAAAT TTTTAATAGC TTCGATTAAT TCCCAATGTG CATCTACTTT ACTCATGCGT 1680
 25 TTGATAAAAA CTAAATAAT ATTAATTCGG TCATCAGTGG CGTTAAATCT TTTATCATTT 1740
 TTAGTTATAG TTGATAAATT TATATTTATA AGCATATATG GATATTTTCA CAAAAATTTT 1800
 TATTTATATA AATCCGAACT GCATACATAT TTGTTTAAAT AAGAGGTATT ATTTTTCGGG 1860
 30 AAATTGCTGT CTGAGTTAAA AGGATTAGTT TTATAAAATG AGTTGAAC TAAGCCAAAA 1920
 CGATTAAAT ACTGATAATC CATTTTGTGA TTATGTTAGG GACTTTTTTA CTAAATTTTA 1980
 ACCCTATTGG aGcMAATATA ATACTCCCTA TTATAAGGAA TAAGGCGTCA TATAAaGGGA 2040
 35 TATAACCTTG AATAAGTTTG ATGACAAAAG CACCAATTGA AGATATAAAA GCAATTACTA 2100
 TACTATTAGC GACTACAGTA TTCATTGGTA ATTTGAATAA AACCAATAAT ATAGGAATAA 2160
 40 TAATGAAGGC ACCACCTGCA CCTACTATAC CTGAAATAAT ACCAATGAAA AGGCCAATGA 2220
 TAACTAATAA ATATTTATTA AATGAAGACT TTTCCGAACT AGGTTtCACT TTAATAAACA 2280
 TTAATGTTAA TGCAAGTAAA GCAATAATGA TATATACCGT ATTTACAAAT GTAGCATCAA 2340
 45 ATAAATTGTC TAGAAATGCA CCTAACATAC TCCCT 2375

(2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 6115 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

	GAGGTTTCTA GACAAGCTTT TAATAACTTA CCAAATCAT TAAGrTGGTT gTgtTGGACT	60
5	GCCcATTATC mAAGtATTAT GaGTTGTTTA ATATTAGtGC TAArACATAC GAAGAGTGGT	120
	TTAAACAATT TAGTAGTAAG AAAGCACAAT TCAGTATTAA TCTCACGGAT AAATGGATAA	180
	TTCAAATCGC ATATGGTAAA TTAATAATAA TGGCTAAAAA TAATGGCGAT ACATATTTTA	240
10	GAGTTCAAAC AATTAAAAAG CCAGGTAATT ATATTTTTTAA CAAATATCGA TTAGAGATAC	300
	ATTCTAATTT ACCAAAATGT TTATTTCCGC TTACAGTGAG AACACGACAA AGTGGCGATA	360
	CATTTAAACT GAATGGGGCGC GATGGTTATA AGAAAGTGAA TCGCCTGTTT ATAGATTGTA	420
15	AAGTGCCACA GTGGGTTCGG GATCAAATGC CAATCGTATT GGATAAACAA CAGCGCATTa	480
	TTGCGGTAGG AGATTTATAT CAACAACAAA CAATAAAAAA ATGGATTATA ATTAGTAAAA	540
20	ATGGAGATGA ATAGCGTTAT GCATAATGAT TTGAAAGAAG TATTGTTAAC TGAAGAAGAT	600
	ATTCAAATA TCTGTAAGGA ATTGGGAGCA CAATTAACAA AGGATTATCA AGGTAAACCA	660
	TTAGTATGCG TGGGTATCTT AAAAGGCTCA GCAATGTTTA TGTcAGATTt AATTAAACGA	720
25	ATTGATACCC ATTTATCAAT TGATTTcATG GATGTTTCTA GTTATCACGG AGGCACTGAG	780
	TCAACTGGTG AAGTTCAAAT CATTAAAGAT TTAGGTTCTT CTATTGAAAA TAAAGACGTA	840
	TTAATTATTG AAGATATCTT AGAGACTGGT ACTACACTTA AGTCAATTAC TGAATTATTA	900
30	CAATCTAGAA AAGTTAATTC ATTAGAAATA GTTACTTTAT TAGATAAACc AAACCGTCGT	960
	AAAGCGGACA TTGAAGCTAA GTATGTAGGT AAAAAAATAC CAGATGaATT TgTTGTTGGt	1020
	TACGGTTTTAG ATTATCGTGA ATTATACCGA AACTTACCAT ATATCGGTAC GTTAAAACCT	1080
35	GAAGTGtATT CAAATTAAATT TTTTAATCAA TTTCAGTTAT TATTACTATG CGTTTGAGAA	1140
	ATAA [~] TAGTGT AGACTCAAAA ATATGAAAAA TGTATTTcAT ATATATTtAA TTTTAGACAA	1200
40	GACATATGTC TTGAAAAGTT GAAAAATATA GAGATTGATA AACTAATAC GGGTGTGAAT	1260
	GACATTGATG TTAAGCTCAA TTACTAGCTT ATAAACATG TCATATGTTA CAATTTTTGT	1320
	TAGTTTTTATT ATGGGAAGTA GGAGGAAATG ACGCATGCAG AAAGCTTTTC GCAATGTGCT	1380
45	AGTTATCGTA ATAATAGGCG TTATTATTTT TGGTCTATTT TCATATTtAA ACGGTAATGG	1440
	AAATATGCCG AAACAGCTTA CATATAATCA ATTTACTGAG AAGTTGGAaA AAGGTGACCT	1500
	TAAAACTTTA GAAATCCAAC CACAACAAAA TGTCTATATG GTAAGTGGTA AAACGAAAAA	1560
50	TGATGAAGAC TATTCATCAA CTATTTTATA TAACAACGAA AAAGAATTAC AAAAAATTAC	1620
	TGATGCTGCT AAAAAGCAAA ACGGTGTAAA ATTAACGATT AAAGAAGAAG AAAAACAAAG	1680

55

	TTTCTTCCTA AGCCAAGCAC AAGGTGGCGG TAGTGGCGGT CGTATGATGA ACTTTGGTAA	1800
	ATCTAAAGCA AAAATGTACG ATAATAATAA ACGTCGTGTT CGTTTCTCTG ATGTAGCAGG	1860
5	GGCAGATGAA GAAAAACAAG AATTAATTGA AATTGTTGAT TTCTTGAAAG ATAATAAAAA	1920
	ATTCAAAGAA ATGGGATCTA GGATTCCTAA AGGTGTCTTA CTTGTTGGAC CTCCAGGTAC	1980
	TGGTAAAACA TTACTTGCTA GAGCGGTTGC AGGTGAAGCT GGCGCACCAT TCTTCTCTAT	2040
10	TAGTGGTTCA GACTTTGTAG AGATGTTTGT TGGTGTGGT GCGAGCCGTG TTCGTGACTT	2100
	ATTCGATAAT GCTAAGAAAA ACGCGCCTTG TATCATCTTT ATCGATGAGA TTGATGCTGT	2160
	TGGTCGTCAA CGTGGTGCAG GTGTTGGTGG CGGTCATGAT GAACGTGAAC AAACCCTAAA	2220
15	CCAATTATTA GTTGAAATGG ATGGTTTCGG TGAAATGAA GGTATCATTG TGATAGCTGC	2280
	TACAAACCGT CCGTATATCC TTGACCCAGC CTTATTACGT CCAGGTCGTT TTGATAGACA	2340
20	AATTCAAGTT GGTGTCCTAG ATGTGAAAGG CCGTGAAGCA ATTCTTCATG TTCATGCTAA	2400
	AAACAAACCA CTTGATGAAA CGGTTGATTT AAAAGCAATT TCACAACGTA CACCTGGTTT	2460
	CTCAGGTGCT GATTTAGAGA ACTTATTTAA TGAAGCATCT TTAATTGCTG TACGTGAAGG	2520
25	TAAAAAGAAA ATTGACATGA GAGATATCGA AGAGGCAACG GATAGAGTTA TAGCCGACC	2580
	TGCTAAGAAA TCTCGAGTTA TTTCTAAGAA AGAACGTAAT ATTGTTGCTC ATCACGAAGC	2640
	TGGTCATACA ATTATCGGTA TGGTACTTGA TGAGGCAGAA GTAGTGATA AAGTTACTAT	2700
30	TGTTCCACGT GGACAAGCAG GTGGTTATGC AATGATGCTA CCTAAACAAG ATCGTTTCTT	2760
	AATGACTGAA CAAGAGTTAT TAGATAAAAT CTGTGGTTTA CTTGGTGGAC GTGTATCAGA	2820
	AGATATTAAC TTTAACGAAG TATCAACAGG TGCTTCAAAT GACTTCGAAC GTGCAACACA	2880
35	AATCGCACGC TCAATGGTTA CGCAATATGG TATGAGTAAA AAATTAGGAC CATTACAGTT	2940
	CGGTCATAGC AATGGTCAAG TATTCTTAGG TAAAGATATG CAAGGTGAGC CTAATTATTC	3000
40	AAGCCAAATC GCATATGAAA TTGATAAAGA AGTTCAACGA ATCGTTAAAG AACAAATACGA	3060
	ACGTTGTAAA CAAATTTTAT TAGAGCACAA AGAACAATTA ATTTTAATTG CTGAAACATT	3120
	ATTAACAGAA GAAACATTAG TTGCTGAACA AATTCAATCA TTATTCTACG AAGGTAAATT	3180
45	ACCTGAAATT GATTATGATG CAGCTAAAGT TGTTAAAGAT GAAGATTCTG AATTTAATGA	3240
	TGGTAAATTC GGTAAATCTT ATGAAGAGAT TCGTAAAGAG CAATTAGAAG ATGGACAACG	3300
	TGACGAAAGT GAAGATCGTA AAGAAGAAAA AGATATTGCT GAGGATAAAA AAGAAGCTGA	3360
50	TAAATCTGAT GAAAAAGATG AACCAGCACA TCGACAAGCC CCAAATATCG AAAAACCTTA	3420
	CGATCCAAAT CACCCAGACA ATAAATAATC GATTATATTC AGTACCTCTT TCTATGATAA	3480

55

	AATTGTTATA GCAGAAAATA ATTGTAAAAC AAGTTACTTC ATTATTTAGA ATGATGGGTG	3600
	TAGAATAAGT ACAATTGTTG CATTTTATGA AGTAAAGTAA TTTTTTAAAT ATAGAGTAAT	3660
5	AGAGGAGATT GAAATAATGA CACACGATTA TATTGTTAAA GCATTAGCAT TTGATGGAGA	3720
	GATTAGGGCT TATGCTGCTT TGACAACTGA AACTGTTCAA GAAGCACAAA CGAGACATTA	3780
	TACATGGCCG ACAGCATCTG CTGCAATGGG AAGAACAATG caCAGCAACA GCTATGATGG	3840
10	GCGCAATGTT GAAAGGTGAT CAAAAATTAA CTGTCACTGT AGATGGCCAA GGACCTATTG	3900
	GACGAATTAT TGCCGATGCA AATGCTAAAG GCGAGGTGCG TGCTTATGTA GACCATCCAC	3960
	AAACTCATTT TCCATTAAAT GAGCAAGGTA AACTTGATGT AAGACGAGCG GTAGGGACAA	4020
15	ATGGATCTAT TATGGTTGTT AAAGACGTTG GAATGAAAGA CTATTTCTCT GGAGCAAGTC	4080
	CaATTGTTTC AGGAGAACTT GGTGAAGATT TTAATTATTA TTATGCTACA AGTGAACAAA	4140
20	CACCTTCATC GGTAGGTCTT GGTGTATTGG TAAATCCTGA TAATACGATT AAAGCAGCAG	4200
	GAGGATTTAT CATTCAAGTT ATGCCAGGTG CCAAAGATGA AACAATTTCA AAATTAGAAA	4260
	AAGCAATTAG TGAAATGACA CCAGTTTCTA AATTAATTGA ACAAGGATTA ACGCCAGAAG	4320
25	GATTACTAAA CGAAATCTTA GGTGAAGACC ATGTGCAAAT TTTAGAGAAA ATGCCTGTTT	4380
	AATTGAATG TAATTGTAGT CATGAGAAAT TTTTAAATGC TATTAAAGGA TTGGGCGAGG	4440
	CTGAGATTCA AAATATGATT AAAGAAGATC ATGGTGCTGA AGCAGTATGT CATTTCTGTG	4500
30	GAAATAAATA TAAATATACT GAAGAAGAAT TAAACGTGTT GCTAGAAAGT TTAGCGTAAT	4560
	TTAATTTAAA TCAATACGCT AAAATGTTTA TTTTLAGCGG TTTAGTGAAA TGTAGAACTA	4620
	AATAGTTGTA TAATCCTTAG TGATTTTGTT TGCTTTCTAG AATTTATTG ATAAAATAAT	4680
35	TCTATATCCG ATAAATAAAC TAAGATTTCA ACAACTAACT AAAAAGGAGT GTTCTTAATG	4740
	GCAAAAAAAC CAGTAGATAA TATTACTCAA ATTATTGGCG GTACACCGGT AGTCAAATTG	4800
40	AGAAATGTAG TAGATGACAA TGCAGCAGAT GTTTATGTAA AATTGGAATA TCAAATCCA	4860
	GGTGGTTCTG TAAAGGATAG AATTGCTTTA GCAATGATTG AAAAAGCAGA GCGAGAAGGC	4920
	AAAATTAAAC CTGGCGATAC AATTGTAGAA CCAACAAGTG GTAATACAGG TATCGGTTTA	4980
45	GCATTTGTAT GTGCTGCTAA AGGATATAAA GCAGTATTTA CTATGCCCGA AACAATGAGC	5040
	CAAGAGCGTC GTAATTTATT AAAAGCATACT GGTGCGGAAT TAGTTTTAAC GCCTGGATCA	5100
	GAAGCGATGA AAGGTGCAAT TAAAAAGCT AAAGAATTGA AAGAAGAACA TGGTTACTTC	5160
50	GAGCCACAAC AATTTGAAAA CCCTGCGAAC CCTGAAGTTC ATGAGTTAAC TACAGGTCCT	5220
	GAGTTATTAC AACAATTTGA AGGGAAACT ATCGATGCGT TCCTAGCTGG TGTGGTACT	5280

55

GTTGCTATAG AGCCTGAGGC TTCTCCAGTA TTGAGCGGTG GTGAGCCAGG TCCACATAAA 5400
 TTACAAGGTT TAGGTGCTGG ATTTATTCCA GGCACCTTGA ATACAGAAAT CTATGACAGT 5460
 5 ATTATTAAAG TAGGAAATGA TACAGCGATG GAAATGTCTC GTCGAGTTGC TAAAGAGGAA 5520
 GGTATTTTAT CAGGTATTTT ATCAGGTGCT GCGATTTTAT CTGCCATTCA AAAAGCAAAA 5580
 GAATTAGGAA AAGGTAAAAC AGTAGTAACA GTATTGCCGA GTAATGGTGA ACGCTACTTA 5640
 10 TCAACACCTT TATATTCATT CGATGACTAA TTAATGTCAT TTAAAAGAGT GAGTTATCTT 5700
 TTTGAGATAA CTTGCTCTTT TTTTCTACCA TGTATATTTT TAAAAATATG AGCGTTAAAT 5760
 TAAACATTTT TCTGATAAAA ATATCCAGTG AATGATAAGA TAATAAACGT ACATACTAAT 5820
 15 AACTAGTAAA TAGCAGGAGT AAATTTTATT AGAGTTAAAC AATACATAAT TAAAGGGTGG 5880
 TTAACATGAC TAAAACAAAA ATTATGGGCA TATTAAACGT CACACCTGAT TcATTCTcAG 5940
 20 ATGGTGGAAG ATTTAATAAT GTTGAATCAG CTATAAATAG aGTGAAAGCC ATGATAGATG 6000
 AAGGTGCTGA CATTATAGAT GTTGGAGGTG TTTCAACGAG ACCCGGTCAT GAAATGGTTT 6060
 CATTAGAAGA TGAGATGAAC AGAGTATTAC CTGTTGTTGA AGCTATTGTC GGTTC 6115

25 (2) INFORMATION FOR SEQ ID NO: 149:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10401 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

35 TAGATACTGG GnTAAAcATc AAAAATATyT GcTtATTCaC GTGTTTAcGc TCCctCAAAC 60
 GCAACGTTAA TTGCGTGTA TcATTtAGTG TGAATTcAGA CGCTTCTTCC ATGACTATGT 120
 40 CTGATATGCC TTTTATCGAC TTTATTTTCT CTGGGTATC TAATCCTTTA AACAAAAAAA 180
 CTGCGCCGTT TGGCAATTCA ACTTTGTTAT CAGTCTTATT CCAAAGGCAC ATGTCCCAAA 240
 TACCAAAGTT TATCAAACAA TCTTTAACAT CTTCGAACAA ACTATCTTTA ATTGTTGATT 300
 45 GTACTTTTCT AAGCCACAGT ATACGCCTAG GATATTTCaC ATCTTGCAAT GCTTTGAGTA 360
 CAACTTTTTG TATAACGCCG TGAGACTTAC CGCTCGAAC TCCACCGTAA TGkACTTCAG 420
 TGAAGTcATC GTAATTGGTT AGTATTTCTG ATATGTTTCT ATTGAAAACA TTAGACGGTT 480
 50 TGTAAAGTT TAATTTAACT TTCGTATCG TACTCACCA TATTAACTC AATATTCTTC 540
 TGAGTAATTT CTTTTTTATC GATATACGCA CCATGTACTT TTAGTATGTG GTCAATAGAT 600

	TTTAAATGGT CATATTTCTT ACTGTAAGCC TCTTGAGGTT CTCCTCTAGC AATAGAAGCA	720
	GATAACGCTA AAGCTTCTGT AATACTCATT AAACGCTCTT CTTGTATCTG TTCTAATCGT	780
5	TCTTTAATAT ATTCCGAAAC ATTAACATTT CTTAACAATC GACTTGCTAA AGACTCTGCT	840
	GTTTTCTTAC TATAACCTGC TGTAAATTGCT GCTTTTTTAC CATTACATCC ATTCATTATA	900
	TATTCATCTG CGAATCTCTT TTGTTTTTCG TTCATTTTCAT TTACCACCAA CTCTCGCGCT	960
10	ATACGCTTTT TAAAATTAAA AAAGGATTGG CTATAATCAG CCAACCCACA TAGATCCTTT	1020
	ATTCCTAATT GCGATAAGGG AAACGCAGTA CGATAGTCAA TATCCTACAC TATCATAATA	1080
	TCTCATTTAA GGTATCAAAA ACTGCCACTT TACTGCCAAT TTCAGTCTTC CCCTAACTCT	1140
15	TCCGCCAATC TAGATATGAT TTTTCTTTTG ATTCTATGAG CAGTTCTATC AGAAATGTGT	1200
	ATGTCAACAC AAACCTTTCAC TAATTCCTTT TTATTAAAAT AATACTCTTG AATGAATTCG	1260
20	CGTTCTTTCC TGCTTGATGT GTTGATTATA CGTTCAATAG CGCTCTTAAA CTCAAGGATT	1320
	TTACCTCTTC GTATACTACA AAGATAATTA GTTACTGCCA TTTCTGTTTT CGATGTATTA	1380
	GACGGTACAA ACTCCCCGCC TATATTTGTA TCTGTTGGAA TCCACGGTGT CATTATTTCA	1440
25	CTTCTTAAAT CTTCAAGTTG TTTATGATAA TTAGGATAAT CACACAACCTC ATCTTCTAAC	1500
	TTTCGAACTG TTGATAATTT TAATCCGTAT TTCTTTTTAG TCATGAATAC CCTCCGTACA	1560
	AATATGTTTA ATCTTCAAAG TGTCTCAATC TACTTCTTAA TATCTCTATC TCTCGCTCTT	1620
30	TAACTTTTAC ATCACCTTTT AACTGTTCCG CTTGTAACAT CACACCAAAC AATAAGATGA	1680
	CTAGTAATAT AATTGCTATG ATTAACCACA TCATCTACTC CGACACCTCC GCCCTCATCA	1740
	AATCAGACTG ATCACTCAAC TTTGCGAAGT CACTTGGCGC CTCTACATCA TCATTAGCCG	1800
35	TCATCATAAT ATATACTTGC TCAGTTACAT ACTTACCTAA CTCATACATC GCTAGTAAGA	1860
	ATAATAGTCT CAAAATTTCT TTAACCACCA CTAAACACCC CATGTTAATT TATCGATAAT	1920
40	TTGTATAGCT TGTTTTAATG CGTCTCTTTT TTCTTTGATA TCTCTATTAT CGCCATCTTC	1980
	ATCAGCTGAC ATTAACCTCAC TGTCAATTC ATATAATAGT TCTGATATTT CATTACTAGC	2040
	TACTACTAAT AAGTTTTTCAT CTACATCAAT CGTTACCGTT TTCTTTGGCA TCTCCATCTC	2100
45	TCCTTATCTT AACTTGTGCC TCGTATTTGC GCTCAGCTTC TTCTTTACTC TCTGCCTCAA	2160
	CAACTGTAAA CGTCTGATTA TCTCTAGCAG TAGTAAAATG TTCATGTGGT TGTCTGTG	2220
	AATCTTTGAA TGTTGTGACT AAGTATTGCG TCACTTCTTA TCACTCCTTT GAATGATTCT	2280
50	AAGTTTTTCT ACGAATAAAA GTATTAGTAC AACACTCAAT GTAGCCAACA TATTTTTTTG	2340
	CTTTGCAAAA TCTACTATAA CGATTAAGAC TAATAACATT CCAATTCTGC ATGTAAATAA	2400

55

	TACAAGTATT GGAACCTAATG TAATGATGTA ACTCACTTCC CCAAAACCTC CTTGACTCGA	2520
	TCTAAGATGT CTTTACACTC CGCTACTTCC GAAGCCTTTT TCTCCACGTT CTGAAACACT	2580
5	TTCGAATTCC TCCACTTGCT TTAGTTCAGG TGTCCATATA GGCACGATAA CCAATTGAGC	2640
	TAGTTTGTCT CCTTCGTTGA TTTGATAAGT TCCGTATTGT CTTATGGCGT CACTCAAATC	2700
	GATTTCTCCT TTAATATCAA AAACACCTGG TGTGATATAA CCATTGATG CAATAGCGTC	2760
10	ATTCTTGATA TTAATCCCTA AATTGCCGTG ATATCCCGCG TCTATCTTGC CTGTTTCAAT	2820
	CACTAAATGC GTTTTACTAC TTACACCACT ACGGCTAGTT AATAGTCCGA CATAGCCCTC	2880
	TGGTATGCTT ACAGCTACAT CTGTTTTAAT CACTGCCTTT TCTTGTGGCT CAAGTACGAC	2940
15	AGTTTCAGCT GAGAATATGT CATAACCTGC ATCCGCTCTTA TGATTTCGTT CGGGCATTCT	3000
	AGCATTTTCT GATAATAGCC TTACTTGTA TGTGTTAGTC ATTTTCCTGC TCCTCCCTAG	3060
	CTGTAGCAAA CGCTATTCTC AATTTCAATC TTTCAACAAT ATGAATTAGT GCGGTATTGA	3120
20	GGAATATTTT AAATTCTTCA ATGTTCTCAT CTATAAAATC AAGTATTTCT TCCTCTTGTT	3180
	CACTGTCAAA CTCGCTTAGT ACATCCCAAA TATTTATGTC GCTTTTGCTC GTTTCTAATA	3240
25	CTCTTTTGAT TATTTCTGAA TTACTTTTAT TACTCATTTT CCTTGTTCCT CCTCATATTT	3300
	ATAGACAACCT TGACCTGCCA TAATCCCTAC TGCTTCATCA AGTTCAATAC CTTCTTTAAC	3360
	TGAATGTTGA ATAGCATTG TCATTCCCTC AAGTATTTCA TCAAACGCTT GTGCTCTCTT	3420
30	ATACACGTCC TCAATCTCTT TTAGTAATCC CTCTGTGTCA TTACCGTTAT ACGCACTAGC	3480
	ACTGATCACT GATTGTTCAA TTTGTTGCGG GTTATTCATC ATTTCCATCT CCTCTAAAT	3540
	AAAGTTAGTT GCTTCTGCTC CTCGTATTCC AAACCATGTT GCTTTATATA TGTTTCGAGC	3600
35	TCTTCCGCTG TATCAAATGT CTTTTTCACG CCTTGCCAAC CTGGCAGGAT ATGCCCATGa	3660
	AAGTAATAAG TGCCGTTTAC TACATGGATA TGTGCCACTC GTTCGTTATC CTGATACAGA	3720
	TATCTCTTAG ATCCGAAAAA TTGGTTTAAG TATCTTTTAC ATGCGCTATC GGTTTTAGGC	3780
40	ATTTATGCTT CCTGCCATTT CTAAACATT TGGTTATAAG TAGTATCAAA CCAGTACGGA	3840
	TCACGTGAAT GTTTTGTAGG CACATTAAAC AAATGTGGCT TCTTCTTACG TAGTTCAGCC	3900
45	TCTTTACGTC GTTGCCCTAGC CATTTACGC TCTTTGCTCT CTCGCTCCAT GATTTTGGAT	3960
	AACACAATTT CTTTATACTC AGCTAAGCGC ATACCATAAG GTGCATGTAA GGCTTCTAAC	4020
	AACGCCCAGC CACCTCGTAC TCTTTTGTCA ACCATTCTCG GAGTTAAACC GTTCTTTTTT	4080
50	ATCAATTCAT TTTCATGTTT GGTAAATTTA TATGGTTTAC CGTTAATCTT TACGATACTC	4140
	ATTTATTCCA CCTCTATACA TTTACTTTTT TTAATCCAAT CCTCTAATTT GTGCGTGTG	4200

55

	ACATTTAAGT TAACCATCTC AGCTTTTCCG TTTTATATC CACTAATAGT TGATCTTGAT	4320
	ACGCCAGTTT CATTGTGCAA ATCTTGGACA CTTACGTTAT CTCTAGCCAT GATTACCCTT	4380
5	AAATTAGTTG CGAATACTc GTTCAACTTC ATTTATTCCA CCTCTATATA TGCATGTCTT	4440
	ATTGTTATGT TGTCACTT TAGTAATTCG TCCGATTGT CATCTAAGCG CTTTGCCAGC	4500
	GSTATCTTTT CTTTATCCAC ATCATCGTAA TGCTGATATT CAACCTTCTGT AGGTATTCTT	4560
10	ATATCAATCG TTGCGTTTAT ATATGCTTGT TGTTCATTA GATCACTTCA TTTCTCTTTT	4620
	TCTTTTACGT CTGACTTCA CTAAGTCCTC ATATACCATC CATCTTGAC CTGTGTATTT	4680
	AGGCGCTTTA CATATCCAGG TTAAATTCAC ATCTCTATAC TGATATCTGA ATATCTTCGC	4740
15	TTTGATGTTG GCAACTTCAG TCGCCTTACC TTTAACGTCT ATAACTTCAA CCAGTTTCCC	4800
	TTCCTTCCAC AAAGAGAAAT CGGCTATATA CGTAATCGGT CTTTGTTC CGAATTTAGG	4860
20	TTGTAATTCA AATTTGGTT GTATTTCGAT ACGATCATAG TTAGTGCCAT TCATATTACT	4920
	TTCTAAATAT TGGTAATATT CGCACTCTAC TTTGCTATCA AATACAATTC CTTGTACTC	4980
	AACCTTCTTA GCATTGTATT TACTCATTGT GCCACCTCTA AATATCAAAT ATCGTTGCTT	5040
25	GCAATCCTAG CTCTTGCTCA TATAGAAGCC CGTGAGCGCC TTTGAATCGT TTTAGGTCAC	5100
	TATCAGTCAT AATTTCTTT TCGTCGCTGA AATGGGCTCC TGTGAGCGAA TAAACTTCAT	5160
	TTACGTTGTC TTTATACTTG ATGACCTTAA TATCTTCCGT GCCATCTTCT CGGTATAAGT	5220
30	AATATTTTTC TTTTCGGCATT TTTTAACACT CCTTAATGTG TGTCTTCTC CAGTTGATTT	5280
	CATTCATGAT TTTCTTTTCA ACTCTGTCGT AATCATCGAA AGGCGATAAC TCGTTATTGT	5340
	CCAACAATCT ATTGACCGCC CAACCACTCT CGATATATAC ATTTGCTACA ATCGGGTCGC	5400
35	TTTGCTTTGT CTCTTCATAC ATCGATTTCATAAAGCTTTT GAATTGCATT ATATTCATGT	5460
	GAAAACCTC TGAGTCTTCT TGTAATACTC AAATTCAATT ATTCCGGTTT CGCCGTCTTT	5520
40	GTTTTTGGCT ATGTTACATT CAACAATAGA TTTGCCAGTG ATACTGTCAT CTTGTCACG	5580
	GTTATAATAA TCATCACGGT AAAGTAGCAT CGCTAAACTC GCATCTGCTT CTATTCCGCC	5640
	TGATTCTTTC ATGTCCGATA GCATTGGTCT TTTATCCTGT CTAGACTCGA CACCACGATT	5700
45	CAGTTGTGAA AGTAGTACGA TGATTGCGCC TGTCTCGTTA GCGATTATCT TTAAGTCACG	5760
	TGATATCTTT TCTACTGCTA CACGTCTATC AACTTTTCGA TCAGTATCCA TCAGTTGAAG	5820
	ATAATCTATA AAAATAACTT GTTGCCGTGC TGAATGCCTC ATTGtTGC GC TCGCACATCT	5880
50	TGCGGTGTGA TATTACTTTT ATCAGAAATA TCGATGCCTA ATTTCATGAT TTTATCCATC	5940
	GCAATCGTTA ACTTTGTAA GTCATCCGGC GTTAAGTTCC TGATTTCTTT TATCTTTGTT	6000

	AGACTAAAGA AAGATGTTTT GATCCATTT TGTGCTATGT TCAGCATCAT GTTTAARTGCA	6120
	AAACCTGTCT TACCCACTGA GGGACGCGCT GCGATGACGA TTAATTGTGA TGGTTCTAAT	6180
5	CCCCCTATTT TGTAATCCAT TAGCTTGTA CCGTCTTAA TTTGCTTCTT AGGGCTATCG	6240
	CTGTATAACT CTTCGACAAA CTCCTCAACA AACTTCTTGG TTCCATCTTC TTTTTTGTTA	6300
	GTAATTGTTT TTAAATCCTT GAGTTCATCA ATCAAGTTGT TAAAGTTTGT GTTCGTAGGT	6360
10	TGTTGTTTGA ACTCAGTTAC CAATTCGTTA GCTTTGTTGA GCTGATAACT TTCCAATAAT	6420
	TCTTGTTGAT AACGTTCAAA GAAGCCATAT CCAATGAAAT CGGAGTTGTA AAGTTTAGTT	6480
	ATAGTATCTG CATCTAAAAA TTCTTTATCT TTAGTTGCTT TTAAATAGAT TTCTTGATGA	6540
15	TCTATCTTTC CGACGTCCAT TACATAATTG AAAAAGGTTT TAACTTTTC GTTCGTAAAC	6600
	ATGTAATCTT TAACTCTTAT CTTTTCTAAT ACGTCCGGTT GTTTAAGTAG CGTAGCGATT	6660
20	ATTGTACTTT CAATTTGAA TTGTCCGTAA TTCATTCGTT TTCGCCCCCA AATTCTGCCA	6720
	ACTTATTCAT GAACTTATCT AGCGCTATTT TTCTTTGTCT GACATATTCG GGGTCATTCT	6780
	GCATTTTCCA TTGGTGTGTA GCGGTTTCGT TATCTACTGG CTCGATAGAT ACTTTTTTAG	6840
25	GTTCCTTACG CATGATTGCT GGTAAGTTAG GCGGGTACGG GTTGTACTG TTGATATAAA	6900
	CATCTACCGC TTTTACAGTT GGTGATAAT CTCCATTTTG ACTTAATACA TCAATCCACA	6960
	TTTCTAACTT CGGTTTATCA AAATCAATGT TGTATACGTA CCTAACTTTT TTAATAATTT	7020
30	CTAATGCTTG TGTTTTGCTC ATCGGCATTA GTCATCACTC AATTCTTTTT CCATTGTGC	7080
	AATGACATCA TCAGTAGTAT TTTTCTAGG TGCTATTTTA TTTCTGCAT CTTCTTTTGT	7140
	TTTGACATTC TCTTTAGCCC AGTTGTTTAA AACTTTAATT AAATAGCCAC CATGCGCACT	7200
35	TTTGCTTTTA GTGTACTCAA CACCTACTTT TACAACCTCA AAAGCGTTTG TACCTATATC	7260
	ATCAATAGCA AACCCTAATT GTTCCATTG ATTAGGTGTT AACTTATCAT CCAAATTTGC	7320
	AATTATATAT TTTATTGAAG ATGAGAAGAC GGCTTCTCTT TCTTCTTCTT TATTCTTATA	7380
40	TTCTTCTTCT TTTTCTTCTT CTCTTCTTCT TTCTTCTTCT GTATCGTTAC GTAACGTTAC	7440
	GGTAACGTTA CGTTTTGCTT CTAGTAACTT TTTCTGTTTC TCACGATAGC GTTGTGTGCG	7500
45	CAATTTATTT TTTTCTTTAT GCTTAGCTTT GCTATCTAAG CTTTGATGCT TCTCCAGTT	7560
	TGTCACTTTT ATGACACCAT TAACTTTTTT AATCATGCCC AATGTCTCAA AAGTTTGAAT	7620
	TGCTAACCTT ATTGAGTTAA TAGGTCTATT AAATTCATTT GCTAACATTT CTTGTTGTA	7680
50	CGGCAAGTTT TCGGATAGCA TAATATAACC TTGTTTCATTG TACTTTCCTG ATAAAGTTAG	7740
	TAACTTAACC CAAATAGTTA TGATCGTATC TCTTTCGGGT AAAGCTTCGA TATATTTGAT	7800

55

	CTCCTTTCAG CATTTTGTG AGCCTCTCAT CAACTTTTAT CCACGAGTCA TGCAAGTGAT	7920
	ATTTATCATC AAACGACTTA ACGCCAATTG CGTGCTGTTC ATTATGATGT TGTCTACACA	7980
5	GTGCTAACAC ATGTTTGTG TAGTGATTCA TTTTGTCTCT GTTCATGCCT CTGCCGACTG	8040
	CTTCATAATG TGCCAGGTCT GCGTGAGGCT TTCCGCATAT TACACAGTTG CGGTTGATTG	8100
10	TAGCCCAATA TAATAACGCT TTATCTTCGC TTAACAACTT ACTCGTTTCT ACACTCATAG	8160
	GTATTTGATG ATGAAACATA AACGCTATAA TCAGTTCTAT TAACTCCCTT GCAACTTTCA	8220
	TAGAACAGTC GCGCAGACTG ATTTCTTCAT AACCTTTCAT AATTTCOAAT TCTGTTTGTA	8280
15	ATAATTTTCT AGTTGATTCT ACTGGTTTCG CCCAGTGAAG TTCTATATCT CTACACATTG	8340
	CGAATATTTT TTTGCGTTGT TCTATAGATA GTTTTTTATT GTCCGGAACC TCTACTTCTG	8400
	CTTTTAGTGG ATATCCGTTT TCTAGTAAGT CAATGTGACT TTGTTCAAGT TCAACACCAG	8460
20	TAGCAACGAC GGAATAAGTA CCGTCATTGT CTTTCTGGTA TCTTGTAATG TATTGCATTT	8520
	AAACCACGTC CTAGAACGGT AAATCATCAT CATTGATTTT TATTGGACCA TTAGCATTAG	8580
	CGAATGGGTT TGATTGTTGA CTCATTGGCG TCTGTTTCCC ATTTGCTTGC TGTCTTTTTT	8640
25	GTTTCATCTC ATCAGTTTTA GGTTCGTGTT TATTAACCTT TTCATCGTCT TTATTCCAAA	8700
	CTTTTACATA TGAGAGTCTT ACAAATACT TGCCTTGTTT CTCGTAAAT TTATTTTTAA	8760
	GTACAATAGT TCCGATTTTG TTAATTAATT GATCTGTGTC AAAAGTTAAA TCTGGTAAGT	8820
30	TCAATTTAAT TCCTAATCTA CTAAGTAACT CGATATATTG TTTTCTTGA TAATCTTGTT	8880
	GGAATGGTGG GACGAATTGG TTGTGTTTGT ATTGTTTACC TTCGTTGTTT TCAAAAACAA	8940
	TCGTGAAGTA TCTGTTTTCT CTGTCGTAA ACTCGACATT TGCAACTTTT ACTGTAAATT	9000
35	CTCCAGCTCC TAAAAAGTCC CCACCTTTCA TGAATGCCTC TTGATTAGTT TCTTGAATGT	9060
	ATTGTGTTCT ACCAGTGATT TTCATAATTT TTATACCGTC CTTTAAATTA ATTTTAATT	9120
40	ACCATTCTA ATTGCTTGTA CAACATCGTT AATACTTGA TTAATGAAAC GTTGTGTT	9180
	AATTTTGATG TTGCTTGAGT GTCTTATCTT TGTCTCGAAT AAATTTGATG GTTCAGCGTT	9240
	AAGTACATAT TGATAAGTTT TTTCGCCGTC TTGCTCATGT TCTTCTATTG TCATTCTTGC	9300
45	TAACACGTCA GATTGACTGA TGAAGCTTTT TTTTATTTGG TCTTGTGCCT CTATCGTGAT	9360
	TGTTGGATTG ATAGTACTTC CCTCATCATC TTTGTCTTTG TTAATGCCCT CGTGTCCGCT	9420
	TATAGCAAGA TGAAATTGAT AATGTTCTTG TAATTTAGAA ATATAACGAT AAATACTTAC	9480
50	AATGCGTGTA GCACACTCGC CCCAATCATT AAATGTCGGT TTCTTTGATT TACCGTCCAT	9540
	GATGTCGTCC ATAGTGATAT CACGTAACTT TTGGATTGTT TCAATCACTA CAACATCAAT	9600

55

AAAATGCTTA TAATTCTTAA TCTGCACAAC TGCCCCATCT TCTGTTACCG TTGTTCCGTC 9720
 CTCATTTATA TCTAGTACTA AGGCATTGTT ATCTTTTGTT AAAAACGTAG TTTTACCAGT 9780
 5 ACCGAACTTG CCGTATATCG CAAATTATA AAACTTGTTT GCATTTTGTT TGCTGATGTC 9840
 TTTTACACCT AGTTGCGTTA AAATATCGAC ATCTTGATTA GTTTTTTCAG TCATCTATTC 9900
 TCCCACCTTT ACCGTGTATG ACGTTGGTTT CTCCACAATG CTAGCACCTT CTAAAACTTC 9960
 10 GCCGTTTGCG TCAATCAATG TGCCGTTTTT AGTTACATTG AAATCTTTCT TAATGTCTGA 10020
 TTGGCTAAGT TTTTATGTTA CTTTACATA GTTGTCAAAA CCTCGTTGCT CAAGTTGTnT 10080
 AATGACTTCT TGCTCATTGC TAACTTGAAT GACTTTTGAA CCTTTTCTGG CTGTCACTTT 10140
 15 TCCGTAAGtG TATTCAACTT GAATTGCTA TCTTGTCTT TTTGTATTCT GTAATATTCA 10200
 ATTACAAGGC TTTGTAAATA TTCTTTGCCA CTCTGTAAAT TTTCTACTTC TTTATCTTTC 10260
 20 CATTCGTTTA TGC GTTCAAT TTCTTTATTT GCTAAATCGT TGATTTCAAT CTCTTTAGTT 10320
 GTGATTGCAT CCAGTTTCTn AAAAACCAG TTAGCACTGT CTAGATCAGT nACTTTGAAT 10380
 CGGTCGTCTT GTTCGAATGT n 10401

25 (2) INFORMATION FOR SEQ ID NO: 150:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2989 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

35 TTTCTCTCTA TTATTCTCGA TCGGTAGATA ATTGTTTAAA TTTAAGTTTA TAGTAATGTT 60
 GAGTTTATAA TTTCATATAT CTAAAAACAG GTGTTGTATA TATAATCATT CATCTAGTTA 120
 TACTTACTTT AAAAATAATA TAATTTTCATG CGATGCAATT CATTGATGGA TGTTTTTAAT 180
 40 CTTAATCAAA TCCAAATAAA GCATATATTT TTAAATTCAC TTTCTTTTGA ATCGATTTTT 240
 ATCTCTTGA TTAACCTTTT CCATTGTTTC ATTAAAGCTC TCTGTCATAT CTATTCCCAT 300
 45 TGAATTGCGT AACATAACA ACACAAATAA ATTATCACCT AATTCTGCTT TAATCGTATT 360
 TGCTTCCTCT GAATCTTTCT TCTTTTTTTC ACCATAGGTA TGATTTATTT CACGTGCAAG 420
 TTCGCCCCACT TCTTCAGTCA ATCTAGCTAA GTTAGCTAAT GGTGAAAAAT ATCCTGTTTT 480
 50 AAATTGTCCA ATATATTCAT CAACTTCAGG TTGCATTTCT ACCATTGATT TCATTTCTAC 540
 GTTCTCCTTA TATTGCATTT CTAATATAGT ATATATCAAT TTGAAGTCTC ATGCATGTTT 600

55

	AATTCAGTTT ATATAAATGT AATGCATTCC TAACTAAATT AAATCAATTG AAATTGGGAT	720
	TATAACTTTA TGATACGTAC CACTACAATA AAATAATATA GTGAATAATC TACCATTAGA	780
5	AAAATAAGCA CAAAAAACT AGCAACCACA CAAAATGTG ATTAGCTAGT TAATAAGTGT	840
	CTAATTIAAG TTAATTGTTA ATCTATAAGA TTAATCACTT GAACGCGCAA TCAAAATAAT	900
	ACGTACAAGC TCTGCTACAG CGACTGCAGT TGCTGCAACA TAAGTCATTG CTGCTGCAGA	960
10	TAATACTTTA CGCGCATGCT TGTATTCTTT TTCATTTACA ATGTTCAATG CCGTAATTTG	1020
	TTTCATCGCT CTTGAACTCG CATCAAACCTC AACTGGTAAC GTAACAATTG AGAATAATAC	1080
15	CGCTAATGAC ATTAAACCAG CACCAATCCA TAAAGCAGTT GAACCAaATG CACTACCTAT	1140
	CGCTGTTAAG ATAATACCTA ACATGATGAT CATATAACTT AATGAACTCC CTAGGTTTGC	1200
	AACAGGTACT AATGCTGCTC TGAATCTTAA GAACCAATAT CCTTGGTGAT CTTGAATGGC	1260
20	ATGACCAACT TCGTGGGCTG CAATTGCAGT TCCAGCAACT GATGGTCTGT CATAGTTTGC	1320
	AGGAGATAGT GAAACAACTT TCTTTTTAGG ATCGTAATGA TCTGTTAAGA ATCCTTCACC	1380
	TTTAACAACT TCGACATCAT AAATACCGTT TGCATGTAAA ATTTCTAATG CAACCTCAGG	1440
25	ACCCGTTTTA CCACTAGTTG ATCTAACTTG TGAATATTTT TCATAGTTAG ATTAACTTT	1500
	GTGTTGTGCC CATAAAGGAA GCACCATTAA TATTACGAAA TAAATTATCA TAGTAAAAAT	1560
	TGAAGACAAT AAACCTACTC TCCTTTATAA ATATTTTACT GTCATTTGCC GTTTTTATCA	1620
30	AATCATTAC ACTTTAATAA TTTGTTTAAT TCAATATAAA GCAAAAGTCC AAAAACACTT	1680
	AGACAACATG ATAATACACC AATTGCCCAC ACATGTGTAG TTATAAAATC ATAATATGGA	1740
	AATTGAAGGT GAAAATAGTC AATATAATCA TTCAAAAACA CCCAAATCAT yGCTACACTG	1800
35	ATTCCAATCA TAGAACGTTT AAACCTAGGA TAGAAGTAAA TTGCCTGAAC AGCCATTATA	1860
	CTGTGGGAAA ACATTAATAC CAAACCATTT ACTGTAATAT CACCTTGTTT AATAATAAAT	1920
40	AATATATTCA TTATAACTGC CCAATCCCA TATTGAATA ATGTTACAAA TGCCAGTGCA	1980
	TCGATAATAC TATTTTGTTT TTGAATTAAT ATCAATGAGA TAGAAATAAC TAAGTATAAT	2040
	ATTGCAGTTG GGCTATCTGG AACAAAAATC TTAAATGCCC AGGGCGTATG ACTTAATTGT	2100
45	TCACCATACC ATATATAACC ATAAATCATC CCTAATATAT TACAAATGAG TAGCATCATT	2160
	AACCAAGAAC GTTGATAAAG TGTATATTGC CAAATGCTT TAATTGTCAT CTGCTAAGTC	2220
	CTCAAATTGA TTATGTTTAT TTACTAGCTT GAGTGTATTT AAAATTTGCG TTAGTTGATA	2280
50	AAAACGTTGC TTTTCATTCA TCTGTAAACT TAAATCAATA TTGTGTAACA AGTAATCTAT	2340
	TAATAACGCA TGTTTATGCC GATCTATAGC CATACTATTT AAGTCATGAA GATAAGTTTG	2400

55

TGACACGTTT GCGAAGTGAA TTTGAATATC AAAAGCACAG TTATGATTAG CGATATAATC 2520
 AAATATTTCA TTTGTATTCA TTAACCTTAT ATTACGCTTA GTAAATTGAA TTGCAGAAGC 2580
 5 GTGACTTCCC ACTTCTGCAA TTTCTAATGT TTCATGATGA TTAATTTTGT TATCTACAAA 2640
 ATGAATGTTT GCCAATTTCTG CCTCATTAC TTTTATATAG TTAAGCACCC AAAGTCAAT 2700
 ACGCGACTTA AATCGATATT GAAAAAGTAA ATATTCAATA AAAGTTCTT TAATTGATT 2760
 10 GAGTGTCTCT GACATCAAAT ACCCCATTTT AAGATTGCAA TCTTGATAAT TCGTCATGCC 2820
 AATTTTCGTT ACTTGGCTCT AGTTCCAACA ATTGATTAA AATAGTAATT GCTTGTCTCT 2880
 15 TTTGACCAAT TTCAATTAAA TAGAAATAAT AATCACTCAT AAAATCAATA TTTGTTTTCA 2940
 TCGTTGGATA TGCTAATTCA AAGAAATGTT GAGCTTCTTT ATCTCGCTC 2989

(2) INFORMATION FOR SEQ ID NO: 151:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1143 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

CATCAACTCC TTAATTACAC TGTAATGAT ATGCGTCTTT TTGACAACTA TATTTGTCAA 60
 30 ATCTACACCA AAAAATATGA TTATCCACCT ATGTATGACA TTTTGAAACA AACACCTCAA 120
 CGCCTACAAG TCATAATTGT TTAATTTCTG TACACCTTCC TGCATAATTA ACAGCATTCT 180
 AATTTTAGTA TGATGCACGC ATTTTCACTA AATCAAACCA TTCAAAGGAG ACTATTATGG 240
 35 CATTTACATT ATCTGCAATT CAACAAGCAC ATCAACAATT TACTGGTGTG GACTTTCCAA 300
 AACTATTCAA AGCTTTTAAA GATATGGGGA TGACTTACAA TATCGTCAAC ATTCAAGATG 360
 GCACTGCAAC ATACGTACAT CAATCAGAAG ATGATATCGT TACGTCATCT GTAAAAAGTA 420
 40 ATCATCCTGT TGCTCAAAAA TCAAACAAAA CAATAGTTCA AGACGTCTTA ACTAGACATC 480
 AACAAAGGCA AACAGATTTT GAAACATTTT GTGATGAAAT GGCTGAAGCT GGCATTTATA 540
 45 AATGGCATAT CGATATTCMA GCGGGCACTT GTACTTATAT CGACTTGCAA GACCAAGCTG 600
 TTATTTTACA ATTAATCCCT CAATAAACTA TATTTATAGC AACATTTTAA TTATTTTATA 660
 AAATTTTATT GATAATCATT ATCGTTCGGT ATAAAGTAAA TACTATATAC TACTTATGAG 720
 50 TGAGGTTGAT TATCATGATA ACTAAGACTT TTATTTTAGG CATCACAGGC CCAACAAGTC 780
 TTGTCGTCAT TAGCATTATC GCTTTAATTA TTTTGGTCC GAAAAAATTA CCACAATTTG 840

55

AGTCTCACGA TACACCCAGT AAGGAATCGA AACACACGCG AGAGCAATAG CACTGACCAC 960
 ACCTTACTGG TTCACCTTAG CGAACTACGC CATCGGTTAG TAAAAATTTT ATTGTCGTTC 1020
 5 GTCATTACGG TCATCGTCGT ATATGTYTCA TCATTTTGGT GGATGACACC ATTCATAACG 1080
 TATATyACCC GgCACATGTG TcCTTACATG CATTTcATTc ACAGAAATGA TACAAATAAC 1140
 10 GTG 1143

(2) INFORMATION FOR SEQ ID NO: 152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7953 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

CAACGCCTGA ACGTAAACCA TATCGTTTCG CGATTTCTCTC ATCTTGACTA TTTACTAAAA 60
 ACTCTCTCAT GCGGATTAAT GTTCTTTTTT CTTCTTTAGT TAATGGTAAT TCTAACTCAG 120
 25 CTGCTTTTTG ACGCAAAGTT GGATGACCAT CTCTAATGAT GTCCTTCATT GTTAACATAT 180
 ATTGCACCTT CCTTATTTTA ATTTGTTTTA GTTGAATGAC AGTAAAAAGG TTGTTAAGAT 240
 ACTCATACAT TTTTATGTGT AAATATCTAC AAAGTTAACC AACTACTGCC AATGTTTATT 300
 30 TTAGATAGTA TATGTAAATT TTCAaGATAT GCgTAATTGC gTTAAAAAAT GaTTAAAGTG 360
 TTGGTTTCAA GCAATGaTAC TTTAGAAATT TATTTATCAT CTTGACTTTA AAAATTATAT 420
 TATAAATGAC GTAAGTGTCA ACAGATATAC TTAGTATGA AGATGTGTAA TGTAATTGTT 480
 35 TAAAATTGAT TTCCAAGCAG ATTTTATTTA TCATTTAATT TAAATAGCAA GTGGAGGTAC 540
 AAGTAATGAA ATTTGGAAAA ACAATCGCAG TAGTATTAGC ATCTAGTGTC TTGCTTGCAG 600
 40 GATGTACTAC GGATAAAAAA GAAATTAAGG CATATTTAAA GCAAGTGGAT AAAATTAAAG 660
 ATGATGAAGA ACCAATTAAA ACTGTTGGTA AGAAAATTGC TGAATTAGAT GAGAAAAAGA 720
 AAAAATTAAC TGAAGATGTC AATAGTAAAG ATACAGCAGT TCGCGGTAAA GCAGTAAAGG 780
 45 ATTTAATTAA AAATGCCGAT GATCGTCTAA AGGAATTTGA AAAAGAAGAA GACGCAATTA 840
 AGAAGTCTGA ACAAGACTTT AAGAAAGCAA AAAGTCACGT TGATAACATT GATAATGATG 900
 TTAAACGTAA AGAAGTAAAA CAATTAGATG ATGTATTAAA AGAAAAATAT AAGTTACACA 960
 50 GTGATTACGC GAAAGCATaT AAAAAGGCTG TAAACTCAGA GAAAACATTA TTTAAATATT 1020
 TAAATCAAAA TGACGCGACA CAACAAGGTG TTAACGAAAA ATCAwAAGCA ATAGAACAGA 1080

	AAGAAAAGCA AGACGTTGAT CAATTTAAAT AATTAATATA ATACAGATGG TAGGAAACAA	1200
	CTAATACAGT TCCTATTATC TGTATCTTTT TTTATTAAAA CAGAACTTTT TCAAATGGTT	1260
5	TAACAGTCCC ATTTATTTGT GGTACAATTA GTAAGGATAA AATGAATTC TATACAATTA	1320
	TGGGAAAGGT ATTGTGAATT GAATGGCTCC TAAGTTACAA GCCCAATTCG ATGCAGTAAA	1380
	AGTTTAAAT GATACTCAAT CGAAATTTGA AATGGTTCAA ATTTTGGATG AGAATGGTAA	1440
10	CGTCGTAAAT GAAGACTTAG TACCTGATCT TACGGATGAA CAATTAGTGG AATTAATGGA	1500
	AAGAATGGTA TGGACTCGTA TCCTTGATCA ACGTTCTATC TCATTAAACA GACAAGGACG	1560
15	TTTAGGTTTC TATGCACCAA CTGCTGGTCA AGAAGCATCA CAATTAGCGT CACAATACGC	1620
	TTTAGAAAAA GAAGATTACA TTTTACCGGG ATACAGAGAT GTTCCTCAA TTATTTGGCA	1680
	TGGTTTACCA TTAAGTGAAG CTTTCTTATT CTCAAGAGGT CACTTCAAAG GAAATCAATT	1740
20	CCCTGAAGGC GTTAATGCAT TAAGCCCACA AATTATTATC GGTGCACAAT ACATTCAAGC	1800
	TGCTGGTGTT GCATTGTCAC TTAACAAAACG TGGTAAAAAT GCAGTTGCAA TCACTTACAC	1860
	TGGTGACGGT GGTTCCTCAC AAGGTGATTT CTACGAAGGT ATTAACCTTG CAGCAGCTTA	1920
25	TAAAGCACCT GCAATTTTCG TTATTCAAAA CAATAACTAT GCAATTTCAA CACCAAGAAG	1980
	CAAGCAAACCT GCTGCTGAAA CATTAGCTCA AAAAGCAATT GCTGTAGGTA TTCCTGGTAT	2040
	CCAAGTTGAT GGTATGGATG CGTTAGCTGT ATATCAAGC AACTAAAGAA GCACGTGACC	2100
30	GCGCAGTTGC AGGTGAAGGT CCAACATTAA TTGAACTAT GACATATCGT TATGGTCCTC	2160
	ATACAATGGC TGGTGACGAT CCAACTCGTT ACAGAAGCTC AGACGAAGAT GCTGAATGGG	2220
	AGAAAAAGA CCCATTAGTA CGTTTCCGTA AATTCCTTGA AAACAAAGGT TTATGGAATG	2280
35	AAGACAAAGA AAATGAAGTT ATTGAACGTG CAAAAGCTGA TATTAAAGCA GCAATTAAAG	2340
	AGGCTGATAA CACTGAAAAA CAACTGTTA CTTCTCTAAT GGAAATTATG TATGAAGATA	2400
40	TGCCTCAAAA CTTAGCAGAA CAATATGAAA TTTACAAAGA GAAGGAGTCG AAGTAAGCCA	2460
	TGGCACAAAT GACAAATGGTT CAAGCGATTA ATGATGCGCT TAAACTGAA CTTAAAAATG	2520
	ACCAAGATGT TTTAATTTTT GGTGAAGACG TTGGTGTTAA CGGCGGTGTT TTCCGTGTTA	2580
45	CTGAAGGACT ACAAAAAGAA TTTGGTGAAG ATAGAGTATT CGATACACCT TTAGCTGAAT	2640
	CAGGTATTGG TGGTTTAGCG ATGGGTCTTG CAGTTGAAGG ATTCCGTCCG GTTATGGAAG	2700
	TACAATTCTT AGGTTTCGTA TTCGAAGTAT TTGATGCGAT TGCTGGACAA ATTGCACGTA	2760
50	CTCGTTTCCG TTCAGGCGGT ACTAAACTG CACCTGTAAC AATTCGTAGC CCATTTGGTG	2820
	GTGGCGTACA CACACCAGAA TTACACGCAG ATAACCTAGA AGGTATTTTA GCTCAATCTC	2880

55

	CTATTAGAAG TAATGACCCA GTCGTATACT TAGAGCATAT GAAATTGTAT CGTTCATTCC	3000
	GTGAAGAAGT ACCTGAAGAA GAATATACAA TTGACATTGG TAAGGCTAAT GTGAAAAAAG	3060
5	AAGGTAATGA CATTTCAATC ATCACATACG GTGCAATGGT TCAAGAATCA ATGAAAGCTG	3120
	CAGAAGAACT TGAAAAAGAT GGTATTCTG TTGAAGTAAT TGACTTACGT ACTGTTCAAC	3180
	CAATCGATGT TGACACAATT GTAGCTTCAG TTGAAAAAAC TGGTCGTGCA GTTGTAGTTC	3240
10	AAGAAGCACA ACGTCAAGCT GGTGTTGGTG CAGCAGTTGT AGCTGAATTA AGTGAACGTG	3300
	CAATCCTTTC ATTAGAAGCA CCTATTGGAA GAGTTGCAGC AGCAGATACA ATTTATCCAT	3360
	TCACTCAAGC TGAAAATGTT TGGTTACCAA ACAAATGA CATCATCGAA AAAGCAAAG	3420
15	AACTTTAGA ATTTTAATAC ATTTTAAAAG TTAACGAAGT TAGCGTATTT TAGTCTCATT	3480
	GATTAATG AAATGTTTAA TTTACGAAAT CTTAGGAGGG CAAAAACGTG GCATTTGAAT	3540
20	TTAGATTACC CGATATCGGG GAAGGTATCC ACGAAGGTGA AATTGTAAAA TGGTTTGTTA	3600
	AAGCTGGAGA TACTATTGAA GAAGACGATG TTTTAGCTGA GGTACAAAAC GATAAATCAG	3660
	TAGTAGAAAT CCCATCACCA GCATCTGGTA CTGTAGAAGA AGTTATGGTA GAAGAAGGTA	3720
25	CAGTAGCTGT AGTTGGTGAC GTTATTGTTA AAATCGATGC ACCTGATGCA GAAGATATGC	3780
	AATTTAAAGG TCATGATGAT GATTCATCAT CTAAGAAGA ACCTGCGAAA GAGGAAGCGC	3840
	CAGcAGaGCA AGCACCTGTA GCTACTCAA CTGAAGAAGT AGATGAAAAC AGAACTGTTA	3900
30	AAGCAATGCC TTCAGTACGT AAATACGCAC GTGAAAAAGG TGTTAACATT AAAGCAGTTT	3960
	CTGGATCTGG TAAAAATGGT CGTATTACAA AAGAAGATGT AGATGCATAC TTAAATGGTG	4020
	GTGCACCAAC AGCTTCAAAT GAATCAGCTG CTTAGCTAC AAGTGAAGAA GTTGCTGAAA	4080
35	CTCCTGCAGC ACCTGCAGCA GTAACATTAG AAGGCGACTT CCCAGAAACA ACTGAAAAAA	4140
	TCCCTGCTAT GCGTAGAGCA ATTGCGAAAG CAATGGTTAA CTCTAAGCAT ACTGCACCTC	4200
40	ATGTAACATT AATGGATGAA ATTGATGTTT AAGCATTATG GGATCACCGT AAGAAATTTA	4260
	AAGAAATCGC AGCTGAACAA GGTACTAAGT TAACATTCTT ACCTTATGTT GTTAAAGCAC	4320
	TTGTTTCTGC ATTGAAAAAA TACCCAGCAC TTAACACTTC ATTCAATGAA GAAGCTGGTG	4380
45	AAATCGTTCA TAAACATTAC TGGAAATATCG GTATTGCAGC AGACACTGAT AGAGGATTAT	4440
	TAGTACCTGT TGTTAAACAT GCTGATCGTA AGTCTATTTT CCAAATTTCA GATGAAATTA	4500
	ATGAATTAGC TGTTAAAGCA CGTGATGGTA AATTAACAGC CGATGAAATG AAAGGTGCTA	4560
50	CATGCACAAT CAGTAATATC GGTTCAGCTG GTGGACAATG GTTCACTCCA GTTATCAATC	4620
	ACCCAGAAGT AGCAATCTTA GGAATTGGCC GTATTGCTCA AAAACCTATC GTTAAAGATG	4680

	ATGGTGCAAC TGGCCAAAAT GCAATGAATC ACATTAAACG TTTATTAAAT AATCCAGAAT	4800
	TATTATTAAT GGAGGGGTAA AACATGGTAG TTGGAGATTT CCCAATTGAA ACAGATACTA	4860
5	TAGTAATCGG AGCAGGTCCT GGTGGATACG TTGCAGCAAT TCGTGCAGCT CAATTAGGAC	4920
	AAAAAGTAAC AATCGTTGAG AAAGGTAATC TTGGTGGTGT TTGCTTAAAC GTAGGATGTA	4980
10	TTCCTTCAAA AGCATTACTA CATGCTTCTC ACCGTTTTGT TGAAGCACAA CATTCTGAAA	5040
	ACTTAGGTGT TATTGCTGAA AGTGTCTTCTT TAAACTTCCA AAAAGTTCAA GAATTCAAAT	5100
	CATCAGTTGT TAATAAATTA ACTGGTGGTG TTGAAAGCTT ACTTAAAGGT AACAAAGTTA	5160
15	ACATCGTTAA AGGTGAAGCA TATTTCTGAG ATAACAATAG CTTACGTGTT ATGGACGAAA	5220
	AGAGCGCACA AACATACAAC TTTAAAAATG CAATCATTGC AACAGGTTCA AGACCAATTG	5280
	AAATTCCTAA TTTCAAATTC GGTAAACGTG TTATCGACTC AACAGGTGCT TTAAACTTAC	5340
20	AAGAAGTACC AGGTAAATTA GTTGTAGTTG GTGGAGGATA CATTGGATCA GAATTAGGTA	5400
	CAGCATTGTC TAACTTTGGT TCAGAAGTAA CCATCCTTGA AGGTGCTAAA GATATCTTAG	5460
	GTGGCTTCGA AAAACAAATG ACACAACCTG TTAATAAAGG TATGAAAGAA AAAGGTGTTG	5520
25	AAATCGTTAC TGAAGCTATG GCTAAATCAG CTGAAGAAAC AGATAACGGA GTTAAAGTTA	5580
	CTTATGAAGC TAAAGGCGAA GAGAAAACAA TCGAAGCTGA TTATGTATTA GTAAGTGTAG	5640
	GTCGTCGTCC AAACACAGAC GAATTAGGCC TAGAAGAATT AGGTGTTAAA TTCGCTGACC	5700
30	GTGGATTATT AGAAGTTGAT AAACAAAGCC GTACGTCTAT CAGCAATATC TATGCAATTG	5760
	GTGATATCGT TCCAGGTTTA CCACTTGCTC ACAAAGCTAG CTATGAAGCT AAAGTTGCTG	5820
35	CTGAAGCAAT TGATGGTCAA GCTGCTGAAG TTGATTACAT TGGTATGCCA GCAGTATGCT	5880
	TTACTGAACC AGAATTAGCT ACAGTTGGTT ATTCAGAAGC GCAAGCTAAA GAAGAAGGTT	5940
	TAGCAATTAA AGCTTCTAAA TTCCCATATG CAGCAAATGG TCGTGCATTA TCATTAGATG	6000
40	ATACTAACGG ATTTGTTAAA CTTATTACAC TTAAAGAAGA TGATACTTTA ATCGGTGCTC	6060
	AAGTAGTTGG TACTGGTGCA TCAGATATTA TCTCTGAATT AGGTTTAGCA ATTGAAGCTG	6120
	GSTATGAATGC TGAAGATATC GCATTAAACAA TCCATGCACA TCCAACATTA GGTGAGATGA	6180
45	CTATGGAAGC AGCAGAAAAA GCTATCGGAT ACCCAATCCA TACAATGTAA TAACTGATTA	6240
	TCTATAAAGA TTCAGTCATT AAAAGCTGTA GCATATGCTA CGGCTTTTTT GTTTTAGGTA	6300
	AAGTAATGTA AGGAAATTGA TTTGAGATAT CGTTAACATG TGACATGCAT GTTATACTAG	6360
50	CGATGCTAAT AAAAGAATTG AAATGGAGGG TTCAACAATG GAATATGAGT ATCCAATTGA	6420
	TTTAGACTGG AGTAATGAAG AGATGATTTT AGTGATAAAT TTCTTTAATC ATGTAGAGAA	6480

55

AATTGTGCCT GCTAAAGCAG AGGAAAAACA AATTTTAAAT ACTTTCGAAA AAAGTAGTGG 6600
 CTATAATAGT TACAAAGCAG TTCAAGATGT AAAAAGTCAC TCTGAAGAAC AAAGAGTAAC 6660
 5 AGCTAAAnAA TAATTCGTTT GAAATTAACA CAATTTAATA GGAATTTTTC TTTAAAACTA 6720
 TTGCTAATAA AGCTATATTT TGATACCTTT ATCAAGTGTT AAACAAAATG TTTGATAAAA 6780
 GTAAACTTAA TATAGCTTTT TTAGGTGGAA AAATAAATGA ACATAGGTAA TAAAATTAAA 6840
 10 AATCTTAGAA GAATTAAAAA TTTAACGCAA GAAGAACTTG CTGAACGTAC AGACTTATCG 6900
 AAAGGCTACA TTTACAAAAT AGAAAGTGAA CATGCCTCAC CAAGTATGGA AACTTTCTTA 6960
 AATATTATAG AGGTGTTAGG AACGACGCCA AGTGAATTTT TTAAAGACAG TGAAAATGAA 7020
 15 AAAGTATTAT ACAAGAAGGA AGAACRAAGT ATTTATGATG AGTATGATGA AGGTTATATA 7080
 TTAAATTGGT TAGTTTCAA GTCAAATGAA TATGATATGG AGCCATTAAT ATTAACTTTA 7140
 20 AAGCCTGGAG CATCATATAA AAATTTTAAAT CCATCAGAGT CTGATACGTT TATTTATTGT 7200
 ATGTCAGGTC AGATAACACT TAATTTAGGC AAAGAGATAT ATCAAGCACA AGAAGAAGAC 7260
 GTTTTGATT TTAAAGCAG AGATAATCAT CGTTTGTCAA ACGAATCAAA CAATGAAACA 7320
 25 CGAATACTTA TTGTAGCGAC AGCTTCATAT TTATAGGGGG GATCTTATTT GGAACCGTTA 7380
 TTATCATTAA AATCAGTTAG TAAAAGCTAT GATGATCTTA ATATCTTAGA TGACATAGAT 7440
 ATTGATATG AATCAGGATA CTTTTATACA TTATTAGGTC CTTCAGGTTG TGGTAAAACA 7500
 30 ACAATTTTAA AATTAATTGC AGGGTTTGAA TATCCTGACA GTGGTGAAGT GATTTATCAA 7560
 AACAAACCAA TTGGTAATTT ACCACCAAAT AAACGTAAAG TGAATACAGT CTTTCAAGAT 7620
 TATGCATTAT TTCCACACTT AAACGTCTAT GATAATATCG CTTTGGTTT GAAATTAAAA 7680
 35 AAATTATCAA AAACCGAAAT TGATCAAAAA GTAACGAGG CATTAAAATT AGTAAAACCTT 7740
 TCAGTTATG AAAAAAGAAA TATTAATGAA ATGAGTGGCG GACAAAAGCA ACGTGTGCA 7800
 ATTGCACGTG CTATCGTAAA TGAACCAGAA ATATTATTGT TAGATGAATC TTTATCCGCA 7860
 40 TTAGATTGA AATTGCGTAC TGAATGCAA TATGAATTAC GAGAATTGCa ATCTAGATTA 7920
 GGtATTACAT TTATATTTGT aACACATGAT CCA 7953

(2) INFORMATION FOR SEQ ID NO: 153:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2347 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	GGCGTGATCA TACGACCGTC ATTCTGCTC ATGAAAAAAT ATCTAAAGAT TTAAAAGAAG	60
	ATCCTATTTT TAAACAAGAA GTAGAGAATC TTGAAAAAGA AATAAGAAAT GTATAAGTAG	120
5	GAACTTTGG GAAATGTAAT CTGTTATATA ACAGCACTAA TGATnACAAT CATTTTTTAC	180
	ATTTCTATAT GCTAATGTGG CAAGATGAGC AAAACTCATT TTGTGGATaA TGTTTtAAAG	240
	TCATACACAC CATAACAAG TTATCAACAT GTGTATAAyT cGcCAAATCT ATGTTTTTTAA	300
10	GACTTATCCA CCAATCCACA GCACCTACTA CTATTACTAA GAACTTAAAA CCTATATAAT	360
	TATATATAAA CGACTGGAAG GAGTTTTAAT TAATGATGGA ATTcACTATT AAAAGAGATT	420
15	ATTTTATTAC ACAATTaAAT GACACATTAA AAGCTATTTT ACCAAGaACA ACATTACCTA	480
	TATTAAGTGG TATCAAAATC GATGCGAAAG AACATGAAGT TATATTaACT GGTTCAGACT	540
	CTGAAATTTT AATAGAAATC ACTATTCCTA AAAGTGTAGA TGGCGAAGAT ATTGTCAATA	600
20	TTTCAGAAAC AGGCTCAGTA GTACTTCCTG GACGATTCTT TGTTGATATT ATAAAAAAT	660
	TACCTGGTAA AGATGTTAAA TTATCTACAA ATGAACAATT CCAGACATTA ATTACATCAG	720
	GTCAATCTGA ATTTAATTTA AGTGGCTTAG ATCCAGATCA ATATCCTTTA TTACCTCAAG	780
25	TTTCTAGAGA TGACGCAATT CAATTGTCGG TAAAAGTGCT TAAAAACGTG ATTGCACAAA	840
	CAAATTTTGC AGTGTCCAcC TCAGAAACAC GCCCAGTACT AACTGGTGTG AACTGGCTTA	900
	TACAAGAAAA TGAATTAATA TGCACAGCGA CTGACTCACA CCGCTTGGCT GTAAGAAAGT	960
30	TGCAGTTAGA AGATGTTTCT GAAACAAAA ATGTCATCAT TCCAGGTAAG GCTTTAGCTG	1020
	AATTAATAA AATTATGTCT GACAATGAAG AAGACATTGA TATCTTCTTT GCTTCAAACC	1080
	AAGTTTTATT TAAAGTTGGA AATGTGAACT TTATTTCTCG ATTATTAGAA GGACATTATC	1140
35	CTGATACAAC ACGTTTATTC CCTGAAAAC ATGAAATTAA ATTAAGTATA GACAATGGGG	1200
	AGTTTTATCA TGCGATTGAT CGTGCCTCTT TATTAGCGCG TGAAGGTGGT AATAACGTTA	1260
40	TTAAATTAAG TACAGGTGAT GACGTTGTTG AATTGTCTTC TACATCACCA GAAATTGGTA	1320
	CTGTAAAAGA AGAAGTTGAT GCAAACGATG TTGAAGGTGG TAGCCTGAAA ATTCATTCA	1380
	ACTCTAAATA TATGATGGAT GCTTTAAAAG CAATCGATAA TGATGAGGTT GAAGTTGAAT	1440
45	TCTTCGGTAC AATGAAACCA TTTATTCTAA AACCAAAAGG TGACGACTCG GTAACGCAAT	1500
	TAATTTTACC AATCAGAACT TACTAAAAAT AAATATAAAT AAAGGATGAC GTGATTAATT	1560
	AAAACGTCAT CCTTTATTTT TTGGCAAAAA TAATTCTAGG TGCGTATGTA AAATAAATTT	1620
50	GGCAGCATTT TAAACAGCAA ATAAAGACG CCAATTAAAT TTATGACAAA TGTATCCAAA	1680
	ATTTAATAAG TGTGCTTATA TGCCCTTTAA ATTTAAAATT TTAATAGTCA ATAACAAGTT	1740

55

5 AAAAATAAGA ATTAATTATT TATATGTAAA CGGTTTCTAC CTCTATTTTA AATGAAATTT 1860
 GTGACAAAAA AAGGTATAAT ATATTAATGA CATACAAAGA AATGGAGTGA TTATTTTGGT 1920
 TCAAGAAGTT GTAGTAGAAG GAGACATTAA TTTAGGTCAA TTTCTAAAAA CAGAAGGGAT 1980
 TATTGAATCT GGTGGTCAAG CAAAATGGTT CTTGCAAGAC GTTGAAGTAT TAATTAATGG 2040
 AGTGCGTGAA ACACGTCGCG GTAAAAAGTT AGAACATCAA GATCGTATAG ATATCCCAGA 2100
 10 ATTACCTGAA GATGCTGGTT CTTTCTTAAT CATTCAATCA GGTGAACAAT GAAGTTAAAT 2160
 ACACTCCAAT TAGAAAATTA TCGTAACTAT GATGAGGTTA CGTTGAAATG TCATCCTGAC 2220
 GTGAATATCC TCATTGGAGA AAATGCACAA GGGAAAGACA AATTTACTTG GAATCAATTT 2280
 15 ATACCTTAGC TTTAGCAAAA AGTCATAGAA CGAGTAATGG ATAAGGGACT CCATACCGTT 2340
 TTAATGC 2347

(2) INFORMATION FOR SEQ ID NO: 154:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13542 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

30 ACAAGACGTT TCTATACTT ATCTGAAATC GCTCGTCAAG ATAAAGATTA TGCAACTATC 60
 TCATTCTTAA ACTGGTTCTT AGATGAACAA GTCGAAGAAG AATCAATGTT TGAAACTCAC 120
 ATCAATTATT TAACTCGTAT CGGCGATGAC AGCAATGCAT TATATCTTTA CGAAAAAGAA 180
 35 CTTGGCGCTC GTACATTGCA CGAAGAATAA TTAAACATCA CTACAATAGA CAGATAAATA 240
 TCATACGACA TGATAGGCAT TTGGGTCACT TACAATAACC CAATGTCTAT ATTATTTTGC 300
 TTTACGGAGA TCACTAGATT CATTCTCTGA ATCATTGATC TGCCTTTTTT CATTCTCAAG 360
 40 GCTAATTATT GTATTTTATG TCATTATTTT TTTAACTAC TAATGTTAAT AACTCTAAAT 420
 TTGATGTTGA ATTAATTTGA CGATTTTAAA GCATATCATC ATTTACTTTT TAATCAGAGT 480
 45 TACATCCAAA TGATAGATTT CACGTTATAC CTTACGTAT AATATTATGT ATCGTTTGTA 540
 AGCAATGAC TAAAAGTCTA TTAATATATA CATTTAATTA ATTGAAAGGA TTGACTACAT 600
 GATACAAGAT GCGTTTGTTG CACTTGATTT TGAACAGCA AATGGTAAAC GTACAAGTAT 660
 50 TTGTTCTGTC GGAATGGTTA AAGTCATTGA TAGTCAAATA ACAGAAACAT TTCATACTCT 720
 TGTGAATCCG CAAGACTATT TTTCAACAACA AATATTAAA ATTCATGGCA TACAACCAGA 780

55

	aGATTTACCT GTTGTGCGAC ATAACGCGGC ATTTGATATG AACGTCTTAC ATCAAAGCAT	900
	TCAAAATATT GGTTTACCAA CTCCAAATTT AACTTACTTT TGTAGTTATC AACTTGCTAA	960
5	AAGAACCGTT GATTCGTATC GATACGGTTT AAAACATATG ATGGAGTTTT ATCAATTAGA	1020
	TTTTCATGGT CATCATGATG CATTGAATGA TGCCAAAGCA TGCGCAATGA TTACTTTTAG	1080
	GCTACTGAAA AATTATGAAA ATTTAACATA TGTAACATA ATTTATGGTA AAAATCTAAA	1140
10	AGATAAAGGC TAGGACTAAA TAAAATACTC CCTTCAAAAG TAAGCATTGT AAAATGTAA	1200
	ACTTTCAGG GAGCTTTATT TTATATAAAG TCATATATCG TCATATTTTT ATAAGTTGAT	1260
15	TGTTCTAAAT TACCTACAGT GACACCAATA AGTCGAATTG GTACATCAGG GTCTTTTAAA	1320
	TCGTTATAAA GTAAATATGC AATATTATAA ATATCTTCTT CAGAACTAAC CGAATCTCTT	1380
	AAACTCATCT GTTTAGATAG CGTTTCAAAT TGATAAGTTT TAATTTTAAAC CGTTACAGTT	1440
20	TTAGCTGACT TCTGTAATTT ATTTAGACGT TCAGCTGTTT TACCTGnACA ATTCCCATAC	1500
	TTTTCTTAAA ATCTCTTCAT CATCATTCAC GTCTGTTGCA AATGTGCGTT CAGTCCCTAC	1560
	TGATTTTCTT ACTCTTGATG ATTTCACTTC ACTATGGTCA ATACCGCGTG CCTTGTTATA	1620
25	TAAACCCCGA CCTCTTTTTT CAAACAAACG TATTAATTCA AATTCGGTTT TCTCATATAA	1680
	ATCTCTACCG TTAAAATAC CATTATCATG CATTACTTTT TTGGAAGCTT TACCTACGCC	1740
	TGGaAAATCT CCAATATCCA ATGTCATCAA AATATCATGG aCATTTTGAT AATCAATCAC	1800
30	AGTCATACCA TCAGGTTTAT TCATACCACT CGCTAATTTA GCTAAAAATT TGTTATAAGA	1860
	AACACCTGCA GATGCTGTTA AATGTGTCTG CTCTAGAATA TCTTTTCTAA TATACTGAGC	1920
	AATTTTCGAA GCAGGAAGGT CTGGTCTCAC TAATTCTGTA ATATCTAAAT ACGCTTCATC	1980
35	CAATGACATC GGTTCACCT TATCTGTATA ACTTCGGAAA ATAGACATAA TCTGCGCAGA	2040
	TGTTTCTCGG TAAGCACCAA AATTACTTGT GACAAAGTAT CCATTTGGAC ATAATTTATG	2100
	CGCTTGAGAC ATAGGCATTG CTGAATGGAC GCCGTATTTT CGTGCTTCAT AGGATGCCGT	2160
40	AGAGACAACA CCCCTACTGC TTGCTTTACC ACCAACAATG ACTGGTTTCC CTTTCAATTT	2220
	GGGGTTATCT CTCATTTTGA CTTGTGCAAA AAAATAGTCC ATATCTATAT GAATAATTCTG	2280
45	TCTCTCAGTC AAGTGCTCAC CTCCCTACTA ATTTTACTT TTATAACGCA CAAAAATATC	2340
	TCAACATAAT TATACGCTGT GTACGATTTT TTTACATAAA TCTTGCACTT AGCGATAACT	2400
	ATATTGaGAT AACTACAAGT TGTTATaAAA TCAATTGCTA TTTAAGCATG ATGATGAAGA	2460
50	CGATTGAGTA AGAAAACATA GGTAACTCTGA AATAATTCAA GCAAATTCAT TTTGTTGGTA	2520
	TCATCATATT AAAATTTATT ATTGAGTCGG CTTTGTATGA TACAAATAAA TACTATCTTC	2580

55

	AAAGCAATAA	GCGGTATGCA	TACTAAACAT	AAAAATAAGT	GATGAATAAC	CAAATACCTT	2700
	AATTAAAATA	AGCAAGCCAG	TACTTAATAG	GATTAGTGGT	GACAGCATAA	TAATTGAGAA	2760
5	TTGCCATTTG	TTGAAGCAAG	CATCTGCTGT	TTGGAATAAG	ATTCTGTCTT	TTTTTATATT	2820
	AAACATAGGT	TTGCTATCTT	TTTTAAATAA	AAGAAATAAT	GCTCTATGGA	TAAGTTCATG	2880
	TAAAATCAAT	AAAATAATGA	ATCCAGCAAA	CCCATATACA	AGATTGATGA	TGATATTTTG	2940
10	ATCGACAACC	GCTGTGACAC	CTAACGCCCA	CTTATACGTA	AATAAAATCA	CGAATAACGC	3000
	AATAACAAGT	TGCAAGATAA	TAAACCTTCG	CATTTGAAAA	TTATTTGTCTG	TTAAATCAAT	3060
15	TTTATGCATT	ACCAACCCCTC	CCGATCATGA	CATTCTTATT	CTTCTTTAAA	TATAGTATAC	3120
	AATGTCACAT	TTAATTTTAAA	AAGTTCATAT	CAAGAAAGTA	AATTGGCTGT	AATAAAATTT	3180
	TAATATACGA	CTTCTTTCTT	CACTTATTAA	GGCGAAATTT	TATCTCAAAT	CATGTGCGCT	3240
20	ATTTCAAATT	GAATAATGCC	ACTGTCTCAA	CATGTGTTGT	TTGTGGAAAC	ATATCTACCG	3300
	GTGTTACCTC	TTCAAGTTGA	TATTTTTTCAG	CTAATAATAA	TGCATCACGT	TGCTGTGTTG	3360
	CGGGATTACA	TGAAATATAG	ACAATACGCT	TAGGTTCTAA	TGTAAGCAAA	GTCTGAATAA	3420
25	ACGTTTCGTC	ACAGCCCTTT	CTTGGCGGAT	CAACCATTAC	AACATCTGGT	TTAATCCCTT	3480
	GTGCTTTCCA	TTGTAAAATA	ACTTCTTCAG	CTTTCCCA	GACAAAAGTT	GTATTATTGC	3540
	ATTGGTTTAT	AGTCGCATTT	TGTTGTGCGT	CTTCAATTGC	AGAAGGTACT	ACTTCAACAC	3600
30	CGTATACATG	TTTTGCAAGT	GGTGCCATAT	ATAGCCCTAT	TGTTCCAATA	CCACAATAGG	3660
	TATCTAATAC	AAC TTCATTA	CCTGTCAATT	GCGCATACTC	AATTGCTTTA	TTATATAATT	3720
	TCTCTGTTTG	TTCAGAATTA	ATTTGGTAGA	ATGACTGATC	ACTTATTTTA	AATGTACTAT	3780
35	CTGTTAATTG	ATCAATAATT	GTATCTTTAC	CATATAGCGT	TATAGATTGA	CGTCCCATAA	3840
	TAACATTAGA	GTGGCTATCA	TTAATGTTTT	GTTTAAATGCT	TGTCACATTA	GGAAATGCAT	3900
40	CTAATATCTT	CTCAACAACA	GCATTTTTTT	GTGGCCACTT	TTTACCATTA	GTTACAAAAA	3960
	TAATCATCAT	TTCGTCTGTA	TGATATCCTG	TTCTTACAAC	CAAATGTCTC	ATTAAACCTT	4020
	TTTTCAATTG	TTCTTGATAA	ATACTTACAT	TTAAATCTTT	TAAAATAGAT	TTAACTTCAT	4080
45	TCATCACTTC	TTGATGTTGT	GAATCTTGTA	TTAAACAAC	TTCCATGTCA	ATAATGTCAT	4140
	GGCTTCTTTG	ACGATAAAAG	CCCATATAAA	CTTCATTCTG	TTCAATTCTTA	CCAACTGGAA	4200
	TCTGGGACTT	GTTTCGATAT	CTCCAAGGAT	CTGTCATGCC	AACTGTATCG	TTAATCTTAG	4260
50	AATTATCAAA	ATGCGCTTTT	CGCTGAAACA	AATTAATCAC	TTGTTCCCTT	TTCAATTTCAA	4320
	GTTGTGCTTC	GTATGATAAG	TGTTGAAGTT	GGCACCCACC	ACAACGTTCA	TAATATATAC	4380

55

	AGTTCTTTTT TACTTTGATA ATTTTATATT CAATTTGTTT ATTAATTAAA GCTTGTGGTA	4500
	TGAAAATAGG AAAGCGATCT ATTTTACGA CACCATGGCC TTCATGCGTT AAATCAACAA	4560
5	CTGTTCCCGT TTTTATGTCA TTTTtagCTA TTGCTTGCAA AATTTTACCT CCAAAATGAA	4620
	CAGGTTAGGA ACAAATTAT GCGCTTCCTA ACCTGCCATT ATATATTTCA CTATTTCTGT	4680
	TTATTCTTCG ATTAAATTGT CATCAACATG ATCATTATTT ATTAACTCTT CATTTACAAT	4740
10	ATCATTAGGT GCAAAGACAT CTATATGACG TTCTAGATTT AAGAAATTCG CTGGTAATTT	4800
	ACCACCATAT TCTCCATCTA CATTTAGTTG TAAGTCTGTG AATGATGAAA TATTAATTGC	4860
	CTTTGCTTTT TCATAAATAA CTTTAGGATG CTTAGTATGT TCTCCTCTTG AAGCTAAAGT	4920
15	CATAATATGA CCAAGTTCTG CAAGGTTTGA TTTTCAACT ATAATTAACG TAAATAGCC	4980
	GTCATCTAAC TTAGCGTCCG GCACTAATTT TTCAAATCCT GCCATTGAAT TTGTTAAACC	5040
20	TAAAAAGAAT AATAATGCTT CTCCTTGGAA AACATTACCA TCATATTCAA TTCTTAAATC	5100
	TACAGCTTTC ATTTGAGGTA ACATTTGCAA ACCTTTGATG TAATAAGCAA ATGGACCAAC	5160
	AATAGATTTC AATTTACTCG GTGTTTCATA AGAGACTTGC GTCAATTGTC CGCCTGCAGC	5220
25	TAAATTAATA AAGTATCGAT TATTCATTTT ACCAATATCT ACTTTAGTAG AATGACCTTC	5280
	AATGATGACA TCAAGTGCCC CCATGATGTC ATTAGGTATA TGCAATGCAC GTCCAAAGTC	5340
	ATTAACAGTA CCCATAGGAA TGACACCTAG CTTAGGACGA TTAGGCTTTT CTGCGATACC	5400
30	ATTAACTACT TCATTTAATG TTCCATCACC ACCTGCAGCG ATTAATACAT CATAATTTTC	5460
	ATGCATAGCT CTTTCTGCTT CAAGTGTGGC ATCACCTATT TTCTCGGTTG CATATGCACT	5520
	CGTTTCATAT CCCGCTTTTT CTAATTTTAT TAAGGCATCA GGTAATTCTC TTTTAAATAG	5580
35	CTCTTTACCT GATGTCGGGT TATAAATGAT TCTAGCACGT TTCCTCATAT CTTATCCCTC	5640
	TACTTAAAAT TCATATATTT TAACTTCATC TTTGTTTCGT CTAATAGGGA GTGGGACAGA	5700
	AATAATATTT AACAAAATTT ATTTGCTTCT ACCCCAATT GCATTGTCTG TAGAATTTCC	5760
40	TTTCGAAATT CTCTATGTTG GGGCCCCACC CCAACTTGCA CATTATTGTA AGCTGACAGA	5820
	AAGTCAGCTT CTTTGTGTTG GGGCCCCGCC AACTTGCACA TTATTGTAAG CTGACAGAAA	5880
45	ATCAGCTTCT ATGTTGGGGC CCCACTAGAA TTGAAAAAG CTTGTTACAA GCGTATTTTC	5940
	TTTCAGTCAA CTACAGCCAA TATAACATTG TAGTGCCTAG GACATTGAAT TTATGACCCA	6000
	GGCTCAGTCT TATTTTCATCA TTCTTAATAT CGTTAAAGAC CAACTTGTAT CTTAAACAAA	6060
50	TACTATCTCA ATATGTACAA AGCTTGTTAT TTATTCAGCA TTTTTGCGG TTCTTCATTA	6120
	TATAGCTTCG TCAGTTATGC TATTTTACCT TTAAAATGAT GTTGTAATA TAATGTTGTC	6180

55

	AACGCATTAA TAAAATTAAT ATTTTACCA TTAACATGTA CAATGAATAA AGTTAAAAGT	6300
	AATTTGACTT CTATAGATAT AAATAAACCC TCGATTGCAT CTAAGTCAGC AATCAAGGGT	6360
5	TTATTTTTTA AATCTTCATA GTTTGATGAT TTAAATTATC TTTTATCTAA TTCTTGTTTT	6420
	AATAGTTGAT TTAATAATTG TGGATTAGCT TGACCTTTAG ACGCTTTCAT AATTTGACCA	6480
10	ACTAGAAGC CCATAGCTTT GCCTTTACCA TTTTGTAAAT CTTCAACTGA TTGTCGTTA	6540
	TTGTCTAATG CTTCAATTAC AAATTTTAGA AGTGTGCTT CATCAGAAAT TTGAACCTAG	6600
	CCATTATCTT CCATAATCTG TTTAGCATTG CCACCTTTAG CTGCTAACTC TGGGAAGACT	6660
15	TTCTTCGCAA TTTTACTGCT CATTGTTCCG TCTTCGATAA GTTTAATCAT ACCTGCTAAA	6720
	TTTTCTGGTG TTAATTTAGT ATCTAATAAT TCTACTTGAT TTTTATTTAA ATATTCGTTT	6780
	ACGCCACCCA TTAACCAGTT AGATGTTAAT TTAACATCTG CACCGTGTTT AATTGTTGAT	6840
20	TCAAAGAAAT CTGACATTTT TTTAGTCAAT GTTAATACGT GTGCATCGTA TGCAGGTAAA	6900
	CCTAATTCAT TTACATACTT AGCTTTACGT TCATCTGGTA ATTCAGGAAT TGTCTGACGA	6960
	ACACGCTCTT TCCAAGCATC ATCAATATAT AAAGGTACAA TGTCAGGCTC TGGGAAGTAA	7020
25	CGGTAATCAT CAGAACCTTC TTTAACACGC ATTAAAATTG TTTTACCTGT AGATTCATCA	7080
	AATCGACGTG TTTCTTGTCG GATTTCTCCA CCATTTAACA ATTCTTCTTC TTGGCGTTTT	7140
	TCTTCATATT CTAAACCTTT ACGTACATAG TTAATGAGT TTAAGTTTTT CAATTCGGCT	7200
30	TTAGTACCAA ATTTTCTTG ACCATATGGA CGTAAAGAGA TGTTAGCATC ACAACGTAAA	7260
	GATCCCTCTT CCATCTTAAC GTCTGATACA CCAGTGTATT GAATAATTGA ACGCAATTTT	7320
	TCTAAATATG CATATGCTTC TTTAGGTGAA CGAATATCTG GTTCAGATAC GATTTCAATT	7380
35	AGCGGTGTAC CTTGACGGTT CAAGTCAACT AATGAATACT CACCTTTATG TGTTGACTTA	7440
	CCAGCATCTT CTTCCATGTG AAGACGAGTA ATACCGATTG GTTTTGTTTT ACCGTCGACT	7500
40	TCGATATCGA TATATCCATT TTCACCAATT GGTGATCAA ATTGAGAAAT TTGATATGCT	7560
	TTTGGATTAT CTGGATAGAA ATAGTTCTTA CGGTCAAACCT TAGATTCTGT TGCGATTTCC	7620
	ATATTTAGTG CCATTGCAGC ACGCATTGCC CAGTCTACTG CACGCTTATT AACAACTGGT	7680
45	AAGACACCTG GATATGCTAA GTCGATAACA TTTGTATTTG AGTTAGGTTT TGCTCCAAA	7740
	TGCGCTGGTG ATGGAGAAAA CATTTTTGAG TCCGTTTTTA ACTCTACGTG AACTTCAAGT	7800
	CCTATAACTG TTTCAAAATG CATGATTTCC ACTCCTTATA ATTTTTCATA AACGTCATGT	7860
50	AAATTGTATT GTGTTTCATA TTGATAAGCG ACACGATATA ACGTTTTTTC ATCGAATGGT	7920
	TTACCAATGA ACTGTAAACC GATTGGTCGG CCATTTGATT GTCCACAAGG AACAGAAATA	7980

55

	GGATCATCAA TTTCTTCACC TAAATTAAAC GCaGTgTnAG GCGCTGTTGG ACCAACTACT	8100
	ACATCATAAT TTTGAATAC TTTATCAAAG TCATTTTAA TCAATGTTCT AACTTTTTGA	8160
5	GATTTTTTAT AGTAAGCATC ATAGTAACCT GAACTTAATG CAAATGTACC TAAGAAAATA	8220
	CGACGTTTTA CTTCTTTACC GAAACCTTCA GATCTTGACA TTTTATATAA TTCTTCTAAT	8280
	GAATGAGCTT CTTTAGAATG ATAACCATAA CGAATTCCGT CAAAACGAGA AAGGTTTGAC	8340
10	GAAGCTTCTG ATGATGCAAT CACGTAATAT GATGGAATAC CAAATTTAGT ATTTGGCAAT	8400
	GATACTTCCT CAACGACAGC ACCTAAAGAT TTTAAAGTTT CTACAGCGTT TTGAACTGCT	8460
15	TCTTTTACGT CATCAGCTAC ACCTTCACCT AAGTATTCTT TAGGTAATGC AACTTTTAAT	8520
	CCTTTAATAT CTTTACCAAT TTCAGATGTA AAGTCTACAT CATCAACTGG TGCACCTGTA	8580
	GAGTCATTAA CATCTGCACC AGAAATAGCT TCTAATACGA TTGCATTATC TTTTACATTT	8640
20	CGAGTCAATG GACCAATTTG GTCTAATGAA GATGCAAAAG CAACTAATCC AAATCGAGAT	8700
	ACACGACCGT ATGTTGGTTT CATACCGACA ACGCCACAAT ATGCAGCCGG TTGTCTAATT	8760
	GAACCACCTG TGTCTGAACC TAAGCTAAAT GGTACTAAGC CAGCTGCAAC TGCTGCTGCA	8820
25	GATCCACCTG ATGAACCACC TGGCACTGCT TTATGGTCAA ATGGGTAAAC TGTTTTTTTG	8880
	AAATAAGATG TTTCTGTTGA ACCACCCATT GCAAACCTCAT CCATATTTAA TTTACCGATT	8940
	AAAACGGCAT TTTCAATTATG TAGTTTTTCC ATTACAGTAG ATTCGTAAAT TGGCACAAAA	9000
30	CCTTCTAACA TTTTACTTGC ACATGTTGTT TCTAATCCGT TTGTAATAAT GTTATCTTTT	9060
	ATACCCATTG GAATACCAA TAATTGCCA TCCATTGAT CTTTGGCTTG TAATTCATCC	9120
	AATTCTTGGC CTTTTTTGAT TGCATTTTCT TTATCCAGCG CTAGAAAAGA CTTAATTGTT	9180
35	GGATCAGTCT CTTCAATTGC ATCATATATA TCTTTAACA CATCAGATGG TTTGATTTTT	9240
	TTGTCTTTTA TTAAAGTTAA TAAATTCTCA ACCGATTCTG AGCGAATGCT CATCTTACGC	9300
	GTCTCCTCA TTCATGATTG TAGGCACTTT AAATTGTCCA TCTTCTGTTT CTTTGGCATT	9360
40	TTTCAAAGCT AATTCTTGTG GAATACCTTT AATTGCTTTA TCTTCACGTA AAACGTTTTG	9420
	TAAATCTAAA ACGTGATATG TAGGTTCAAC GCCTTCTGTA TCAGCGCTAT CATTTTGTTT	9480
45	TGCAAAATCT AAAATGCTTT CTAATGTGTT GGCCATTTCT TCCGTTTCTT CAGGAGAAAT	9540
	TTGAAGTCTT GCAAGATTCG CGATATGCTC AACTTCTTCA CGTGTTACTT TTGTCATTAA	9600
	TAAAAGCCTC CTTTAAGTCA TTCATCACTA AATTGTATCA AATTTCCAAT TAAAAATCTA	9660
50	AGTATTTATG AGGTGCTACT TTAATTCAT ATAACTGTA TAAACATTAT CATTCGTTTA	9720
	TCAAATCATT TTTTATGAAA ACAACACTCT TTTAATATTA GACAACCCAA TTCAATATTA	9780

55

TATATTGGTA TGCAAGTATT TCAAAAAGAA TAAATTTAAT TTTCCTACTT TTCTAAACAT 9900
 TTATCTTTAT GTATAATGTT TTCAAGTAAC TAAATTATAA ATTAAATAAA GGGAGTGTTT 9960
 5 ATCATGCTTA CAATGGGGAC AGCATTAAAGT CAACAAGTAG ATGCCAATTG GCAAACCTTAT 10020
 ATTATGATTG CCGTCTACTT CTTGATACTA ATCGTTATTG GCTTTTACGG TTACAAGCAA 10080
 GCAACTGGTA ACCTAAGCGA GTACATGTTA GGTGGACGTA TATTGGACCG TATATTACTG 10140
 10 CATTATCAGC TGGAGCTTCA GATATGAGTG GATGGATGAT TATGGGGCTA CCTGGTTCTG 10200
 TCTATAGCAC TGGTCTATCA GCTATGTGGA TTACAATCGG TTTAACATTA GGTGCTTATA 10260
 TAAATTACTT TGTGTGTTGCT CCTAGACTTC GTGTTTATAC CGAATTAGCT GGAGATGCAA 10320
 15 TTACATTACC AGATTTCTTT AAAAATCGTT TAAACGATAA AAATAATGTG TTAAAGATTA 10380
 TTTCTGGATT GATTATCGTA GTATTCTTTA CATTATATAC ACATTCTGGT TTCGTATCTG 10440
 GTGGTAAACT ATTTGAAAGT GCTTTTGGAT TAGATTATCA TTTCCGTTTA ATATTAGTTG 10500
 20 CTTTCATTGT CATTTTCTAT ACTTTCTTTG GTGGATATTT AGCTGTATCA ATTACAGATT 10560
 TCTTCCAAGG TGTCATTATG TTAATTGCGA TGGTTATGGT CCCTATTGTT GCTATGATGA 10620
 ATTTAAACGG CTGGGGAACG TTTCATGATG TAGCAGCTAT GAAACCTACA AATTTAAATT 10680
 TATTTAAAGG GTTATCATTT ATAGGAATTA TCTCTCTATT TTCATGGGGA TTAGGTTATT 10740
 TCGGTCAACC TCATATCATT GTAAGGTTTA TGTCTATTAA ATCACACAAG ATGCTACCTA 10800
 30 AAGCTAGACG TTTAGGTATT AGCTGGATGG CTGTTGGTTT ATTAGGCGCT GTGGCTGTTG 10860
 GTTTAAACAGG TATTGCATTG GTACCTGCTT ATCATATTAA ACTAGAAGAT CCTGAGACAT 10920
 TATTCATCGT GATGAGTCAA GTACTCTTCC ATCCTCTTGT AGGTGGTTTC TTACTTGCTG 10980
 35 CGATTCTAGC TGCAATTATG AGCACGATTT CTTCACAATT ACTTGTAACA TCTAGTTCAC 11040
 TAA CGGAAGA CTTTTATAAA TTAATTCGTG GTGAAGAAAA AGCTAAAACG CACCAAAAAAG 11100
 AATTGTTTAT GATTGGAAGA TTATCTGTAT TAGTTGTAGC AATTGTTGCC ATCGCGATTG 11160
 40 CATGGAATCC AAACGACACA ATTCTAAACT TAGTAGGTAA CGCTTGGGCC GGATTTGGTG 11220
 CATCGTTTCA TCCACTTGTG CTATTTGCAC TTTACTGGAA AGGTTTGACA CGTGCCGGTG 11280
 CTGTAAGTGG AATGGTTTCA GGTGCCTTAG TCGTTATCGT TTGGATTGCA TGGATTAAAC 11340
 45 CATTGGCACA TATCAACGAA ATATTGGGCT TATATGAAAT TATTCCTGGA TTTATTGTAA 11400
 GTGTAATCGT TACATATGTT GTAAGTAAAC TTAATAAAAA ACCTGGTGCA TTTGTTGAAA 11460
 50 CTGACTTAAA CAAAGTTCGT GACATCGTTA GAGAAAAATA ATTCATAAGT CTTAACAAAT 11520
 TAAAAAGGTA CTAATGTAA TCAAAATTAT GACTAACATT GGTACCTTTT TATTATCTTT 11580

55

	AATTAAAGCA CGTGGTTGGT TACCATCTTT AATACGAATT TCATAGTTAT CGATTTTATC	11700
	GAAATATTTA TTCGCTTGTT CAGTAACGTA CTGTGTAATA CCAATTGTTT CAGCTTGTC	11760
5	ATAGTAATCG ATTGGTAAAT CTACTACTAA TCGTTGTGGC TTTTATCAA CAAATTTAAC	11820
	TTTCCCTACT GCTTGTGTGA AATTAGAAAA ATATGATTGC AAATTATCAT TAAATTGCTT	11880
	GAAATTATTA TTTAAATTTT CATCATAATC TGCTGCTGTT GAAGAAGGTA ATAAAGCTGA	11940
10	TTTTTCATTG ATATTATGCC ATTCATTAAG CTTTGTGTTGA CTCTTTTCTG CAGTCGCTTG	12000
	AGTGATAAAT TCACCTGGTG TGATTGAATC TTCACTTGAT TGCTTATAAA TTGCAAAATG	12060
	AATTGGTATA TCTTTTAAAT CATCATTTTC ACGTAACCTT GATAATATCT CACTAGCCAT	12120
15	TTGTTTACCT TGCTTTTAA CTcGCTATCA TCTAGTTTTT TACTAAAAGT CGATCCATCT	12180
	TTTTCTTTTT TATAGTAATA AACACTATTC ATAGCTAAAC CAATCGTCAT ACCTTTAATA	12240
20	TTCTTACCTT TTGTATCTCC ACCACCATAA AAATCTTGCT CTAAAATGTT AGATAAATAG	12300
	GCTGGTGATT TTTCTGCAAT CTTTTCAGGA TCTGTTTCAC CTcCGTGTGA TGGATTAAGT	12360
	CCTAAATTTT CATTCGCTTT CTGTCTTTTT TTATCTTTTT CAGACATTTT ATCGATTTC	12420
25	CGTTTTGTAT ACTTAGGATT TAAATAGGCA TTAATTGTTT TCTTGTCCAA AAATTGACCA	12480
	TCTTGATACA AATATTTATC TGTGGAAAT ACTTCTTTAC TTAAGTTCAA TAAACCATCT	12540
	TCAAAGTCGC CGCCATTATA ACTATTTGCC ATGTTATCTT GTAAAAGTCC TCTTGCTGG	12600
30	CTTCTTTTAA ATGGTAACAA TGACGATAG TTATCACCTT GTACATTTTT ATCCGTTGCA	12660
	ATTTCTTTTA CTGATTGTA ACTATTGTTA TGTTTTTGAT TATCTTTTCC AGCCTGGTCA	12720
	TCCTTATGGT TACCACAAGC AGCGAGTATA AAGATAGCTG TAATCAATAA TACTAATGTA	12780
35	CGCTTCATCG ACATACCCCT CTAACATTTT AATTCATTTT GCTTATCTAC AAATTGTTGC	12840
	TCTGTCCAAA TTTCAATACC TAAACTTTGT GCTTTTGTTA ATTTTGAACC TGCATCTTCA	12900
	CCAGCAATAA CGACATCTGT ATTTTTAGTA ACGCTACTTG TAACTTTAGC ACCTTGATG	12960
40	GCAAGCCATT TAGATGCTTC ATTGCGTGTC ATTTGATGTA GCTTACCAGT CAGTACTATC	13020
	GTTTTACCAC TAAATTCAGG ATGTCCTTCA ATATCTGATG TTTTGATACC TTTATAAATC	13080
45	ATATTAACAT GTTTATCTTT TAATTTTTGA ATTAAAGCAC GAATATCTTC ATTTTCTAAA	13140
	TAAGTAACTA CAGATTGTGC TACTTTATCA CCTATATCAT GAATTTCTAC TAATCCGCT	13200
	TCAGTTACCG TTAGTAATCG ATCTATCGTT TCATATTTTT CTGCTAACAC TTGGCTCGCT	13260
50	TTAACACCTA AATGCCTAAT ACCTAGACCA AATAATAAAT TTTCTAAAGA GTTGTCTTAA	13320
	GCTTGTTGAA TGGCAGCTAA TAAATTATCA ACTTTTTTCT GCCCATTTCT GTCTAAAGGT	13380

55

TAAAGCTGTT GAATAATTTT AGTGCCTAAA CCATCAATAT TcATGGCTTG TCTTGaTACA 13500

AAGTGnATCa ATCctTcAAC AAGTTGTGCT TGGTCATTTT GG 13542

5

(2) INFORMATION FOR SEQ ID NO: 155:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1893 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

CAGTAAACAC CTCTGATTAC GAATATTTAT ACATTTATTT TAACACATGC ACTGATTTAC 60
 GACTACTAAA CACCTTTACG TAAAAAGGGT AAACATGGTT TATCTATCTT GGTATCTAT 120
 20 TTATAAATAT TtTtCATATT ACGCATAACA ATTGCTTAAA ATATGTATAA AAATGAATAT 180
 ATGTGTAATA AACTTGCTAA TTATTAGATT TAATAAGCGT CAATTGTTTG AACATATTtA 240
 ATTAAAATCA CATTGATATC ACAGATACGA ATATTGTCGT ATAGAAATTG AAAATTCTAT 300
 25 TTTTAAATG AAAGTCTTCA ACATAATTTT AAGTTTCAAC ATGAGAAAAA TCGATTAACA 360
 AACAACTGCA GTTGAATATG CCTTTTGAGA CATTTCAAAC TTTACAATTG TTGCTAATCG 420
 ATATATTTGC TTTTAGTGAT CCCTGCTATA AAATAATCA ACGATTTCTA ATAAGTGTtT 480
 30 TGTATTGAAT TGTTCACTAA TTTGCGTTAG TTCATCCACT GCTGCGTCTC TATGATAAGT 540
 CAATTTATCT TCTGCGCCAT CTTTCCCTAA TAACTCACG TACGTACTTT TATTATTTTC 600
 AAGATCGCTG CCCACTTTTT TACCTAACTT TGCTTCATCA CCATAGCAGT CTAATAAATC 660
 35 ATCTTTAATC TGGAACATCA TACCTAAATG ATAACTATAA CTTTCTAAAT GTTCTTTAGT 720
 TGTATCATCG ACATTAGCGA TATCTGCTGC ACTCATAACC GCAAAAGTTA ATAATGCTCC 780
 TGTTTTTGTT TTGTGTATCA TTTCCAAAGT TTCAAGATCA ATTGGTTGGC CTTGCTTTG 840
 CATATCTAAC ATTTGACCGC CGACCATTC AACATGACCA CTTGCTATTG ACAGCCGTTG 900
 TAGAACTTTT ATTTTACTT CATCAGTTAA TCTATCATCA CTTGAAATAA GTTCAAATGC 960
 45 TTTAGTTAAT AAAGCATCAC CTGCTAATAT CGCAGTCCAC TCACCATATA CTTTATGATT 1020
 TGTTAATTTT CCTCGTCGAT AATCATCATT ATCCATCGCT GGTAGGTCAT CATGAATAAG 1080
 TGAATATGTA TGAATCATTT CTAGTGCAAT TGCGTCTTTC ATACCTAACT CATACTCGGT 1140
 50 ATTTAGTGAA TCTAAAGTGA GTAATAACAG AACTGGTCGG ATGCGTTTAC CTCCAGCATT 1200
 TAATGAATAC AACATACTTT CTTCTAGCTG AGTATCCATT ACTGATTTAT TTATCGCAAC 1260

55

CATCCTCAGC TTCTTCTTTT ATTAAGTCAT TCACCTTTTT TCGGCATT TTTAAAGTTG 1380
 TGTCACAAGC TGCTGATAGT TTCATACCAC GTTGATATAA ATCTAATGAT TCCTCTAAAG 1440
 5 ATACTGTTTC ATTATCTAAT TTTTGAACAA TTTGCTCTAA TTCTTGCATC ATTTCTTCAA 1500
 AACTTTGCGT TTCTTTAGTC ATTATTACAC CTTACTTTTC TAACTTTTGC ATCTACTAAG 1560
 CCATCTTTCA TTGTTAACGT CAATTGATCA TTTTCTGTAA AATCTTTAGT ACTCGTAATG 1620
 10 ACTTCGTCTT TTTTATTAAC AATTGCATAT CCACGCAACA TTGTATTAGT TGGACTTAAA 1680
 TTGTTTAAAGT TTTCTACTTT ATTTTCAAA TCATTTTAT AACTTAATAT CTTAGAATTC 1740
 15 AATAATTTAA CAAGTTGGTT TGTCAATTGA AGATTATnTT GTTGTCTTG ATTAACACTA 1800
 CTTAGTAATG CTTTTAAATn ATAACGTTGG TGCAACAGCA TTAAATCGAG GCCCGGTGG 1860
 TCCAAAGTTG CCCGAATTnG TGGTTTCAGG CCC 1893

20 (2) INFORMATION FOR SEQ ID NO: 156:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 821 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

30 AAAATATATT CCTTCACTTA ATATTCAATT AGAGAAAAAC ATGGTGATTG TAATATGTTG 60
 TGCAATATTT CTGGGTGTTT TAATACTTTT TTTATTTCTG AATCGTAAGC TAAGGTTGGA 120
 AATTTATAAT AATAACTCTA GTAAAGGGAA AATAATTTTA TTTCTTCAT TAAAAACTT 180
 35 TTGTTTCACA ATATTTTATT ATTTTTTATT TGGCGGTCTT TCAATAATGG CTCTAAGTAT 240
 GTTATTAACT TTAAATCCTC AAAATATAAT AGGCTTTTATT GGTGTTGGTGG TAATGACTGC 300
 AGGTTTCTTT CTGTTAAACA TGTCATCGAT TATTGACAAA AAAATTTATG TATTATCTAA 360
 40 AACTAACACG GTGGAAAAAT GATGGTTTAG CTGGATTTC TGCAGGTTCT ATTTGGGCAA 420
 TACTTGATA TTGGACCAAT CAAAAAATG AATTTGGAAT AAAAGATAAA AACGATTGGA 480
 TAGGACATAA ACTAGACGTT GGTATAGATG CTGTAGAAAA ATCTGCAGAA AAAACAGTAG 540
 45 ATGGTGTTGA AAATGTCATG GTGAAGCTTC AAAAAGTATT TCTAATCATA TAAGCCCTAA 600
 GAAATGGAGC TGGTAAATGT TGCTATGCGA ATCTAAAATC ATCAATAAAA ACCCAAAATA 660
 TAGAATTATT AAATATAATG ATGAATACTT AATGGTCGAT ATAATAAGCA CTTGGATTAG 720
 50 TTTATTTTTT CCTTTTATTA ATTGGTTCAT CCCaAAAGaA TACGTCAAAA TTAGTAGAGA 780

55

(2) INFORMATION FOR SEQ ID NO: 157:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2343 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

AGTAAGATAA TTTTCAATTA GAAAATATCT TACTGCTGTT CTCTATTTAT ACAATACTTC 60
 GTATTGAATG GcTTTCGCTTT CCTAGGGTGC CGTCTCAGCC TTGGTCTTCG ACTGGCACTG 120
 CTCCCTCAGG AGTCTCGCCA TTAATACTAC GTATTAAACAT GTAATTTTAC TTTGAAATAC 180
 TTTTAAAAAA TAAGACACTT TGCCCAACTT GCACATAAAT GTAAAATTCA ATAAAATGAA 240
 TTTTCTGTGT TGGGTCCCTT CTTATAATTT AATAAATACC ACTAACTAA ATTAACGAGG 300
 TGCCTTATGT ATAAAAATTA TAACATGCCC CAACTACACT ACCAATAGAA ACTTCTGTTA 360
 GAATCCCTCA AAATGATATT TCACGATATG TTAATGAAAT TGTGAAACr ATACCTGATA 420
 GCGAATTCGA TGAATTCAGA CATCATCGTG GCGCAACATC CTATCATCCA AAAATGATGT 480
 TAAAAATCAT CTTATATGCA TATACTCAAT CTGTTTAATT ATGTTCAAAG CATTAAAGTA 540
 ACAAGACAAT ATCTAAGATA TCAAAGATAG AAATTTTTTG ACGTTGTTGC TGATTGTAA 600
 CATAACCATC AATTTTCATAA TTAATAGCAT CAATACGATA AATGGTTAAG CGTACTGAAT 660
 CTACAAAGCC ATTATTATAA AATTTAACTT CTACAGGTTG GGCATATTGT AGCGCCTCGT 720
 GTAGCCGAAT GTTTAGCTCA GCCAATTGAT CATCTGATAA TACAGGACGT GTAATTTTGT 780
 TTTGGTCGAT AATGTATTGT TGAATCGTTT CGAATTGTTT GGGTAATGTT GCAAAAGGAG 840
 CCCATTTAAT CATGCCTCTT CCCATAGGTA TATTGTTATC TAGTAATTCT CTTGGAACGT 900
 TACGATAATC AGTTTCTTCT TCATAACTTG TCATCCTTAA TTCACCCCAA TCTGATAATT 960
 ACATTATACG AACATGTGTT CTATTTTGCA ACAAAAATTT TGTGGaAGCA TAAACGCGTT 1020
 AATAATTAAT GCTCGTGtAA GTAAAAAGA GGGATTAATT AAAATCGAAT AATGaCATAT 1080
 CACaGCAAAT AGTTCTTTTA AAGTAGTTAA ATAGTTTtag CTTTAAGGAA aTGATAAaTG 1140
 ATTGTWAATT CTAGCTAAAA TTTAATAAAA TGAAATAAG ACTAACATGG AGGGGTAAAA 1200
 GTAATGACAA ATGGATATAT TGGTTCTTAC ACTAAAAGA ATGGTAAAGG GATTTATCGT 1260
 TTTGAATTAA ACGAAAATCA GTCACGTATT GATTTATTAG AAACAGGATT TGAATTAGAA 1320
 GCGTCTACAT ATTTGGTGCG TAATAATGAA GTTTTATATG GAATCAACAA AGAAGGAGAA 1380

TGT TTTGTCTT CAAAAGCTGG TACAGGTTGT TATGTATCGA TTTCAGAAGA TAAACGATAT 1500
 TTATTTGAAG CGGTATATGG TGCTGGCATC ATACGTATGT ATGAATTAAA TACGCACACA 1560
 5 GGTGAAATTA TACGTCTAAT TCAAGAACTT GCACATGATT TTCCAACAGG TACACATGAA 1620
 AGACAAGATC ATCCACACGC ACATTATATT AATCAAATC CAGATGGTAA GTACGTTGCA 1680
 GTAACAGATT TAGGTGCTGA TCGTATCGTT ACTTATAAAT TTGATGACAA CGGGTTTGAA 1740
 10 TTTTATAAAG AATCTTTATT TAAAGATAGT GATGGGACAA GACATATTGA ATTTTCATGAT 1800
 AATGGAAAAT TTGCTTATGT CGTACACGAA TTATCAAATA CTGTGAGTGT TGCAGAATAT 1860
 AATGACGGTA AATTTGAAGA GCTCGAGCGT CATTTAACAA TTCCTGAAAA CTTTGATGGA 1920
 15 GATACTAAAC TTGcAGCAGT GCGTTTATCT CATGaTCAAC AATTCTTATA TGTATCTAAT 1980
 AGAGGGCATG ATAGCATTGC AATTTTTTAA GTTCTTGATA ATGGTCAACA CTTAGAACTA 2040
 GTAACaATTA CTGAAaGTGG TGGTCAATTC CCAAGAGATT TTAATATTGC CTCATCAGAT 2100
 20 GACCyTTTAG TTgTGCTCA kGaGCaAGGA GATTcAGTTG TAACTGTTTT CGAAAGAAAT 2160
 AAAGAAACAG GTAAAATTAC GCTATGTGAT AACACTCGTG TAGCATCTGA AGGTGTATGT 2220
 25 GTCATATTTT AATCTTTAAT TAATCATGAT AAAAAGAAAA CCATGTTTCC AAAAAATTG 2280
 TGTATACCTT GAAATTTATT GnTTTCCAGn ACATCAATTA TGGGAAGCAT GgNTTATTTT 2340
 TGT 2343

30 (2) INFORMATION FOR SEQ ID NO: 158:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4837 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

AAATTGCCAG TTGGTATCGC TTCTGGTGCA GTAGTCGAAG GTTTCTTCCA AGGTATCATT 60
 CCGATTGGCT ATATCGTTAT GATGGCAGTA TTGTTATACA AAATTACTGT TGAATCTGGA 120
 45 CAATTTTTAA CAATTCAAGA TAGTATTACA AATATTTAC AAGACCAACG TATTCAAGTT 180
 TTACTTATTG GATTTCATT CAACGCATTT TTAGAAGGTG CAGCAGGATT TGGTGTACCA 240
 ATTGCAATTT GTGCACTTTT ATTAACACAA TTAGGATTTA ATCCATTAAA AGCTGCGATG 300
 50 TTATGTTTAG TCGCAAATGC AGCGTCTGGT GCTTTTGGTG CGATTGGTAT CCCTGTAGGT 360
 GTTGTAGAAA CGTTGAAATT ACCTGGAGAT GTTTCATAT TAGGTGTTTC TCAATCAGCA 420

55

	GGTTTTAGAG GTGTTAAAGA AACATTACCA GCAATTTTAG TAGTTTCAAT CACTTATACA	540
	CTTACTCAAG GATTATTAAC TGTATTCAGT GGACCTGAAT TAGCAGATAT TATTCCACCG	600
5	TTATTAACAA TGTTAGCATT AGCAGTATTT TCTAAAAAAT TCCAACCAAA ACACATTTAT	660
	CGTGTTAATA AAGATGAAGA AATTGAACCT GCAAAAGCAC ATTCTGCAA AGCAGTATTA	720
	CATGCATGGA GCCCATTGAT TGTATTAACA GTCATTGTAA TGATTGGAG TCGGCCATTG	780
10	TTTAAAACT TATTCTTACC AAATGGTGCT TTATCATCAT TAGTATTTAA ATTCAACTTA	840
	CCTGGaACAA TCAGCGAAGT TACGCATAAA CCATTAGTAT TGACTTTAAA TATTATTGGA	900
	CAAACAGGTA CAGCTATTTT ATTAAGTATT ATTATTACAA TTTTAATGTC TAAAAAGGTT	960
15	AACTTTAAAG ATGCAGGTAG ATTATTCGGC GTTACATTTA AAGAGTTGTG GTTACCAGTT	1020
	CTTACAATTT GTTTCATCTT AGCAATTTCT AAAATCACAA CTTATGGTGG TTTAAGTGCA	1080
20	GCAATGGGTC AAGGTATTGC TAAAGCAGGT AATGTCTTCC CAGTTCTATC ACCAATTTTA	1140
	GGTTGGATAG GTGTGTTTAT GACAGGATCA GTTGTAATA ACAACTCATT ATTTGCACCA	1200
	ATTCAAGCTT CTGTTGCACA ACAAATTGGA ACAAGTGGTT CACTTCTTGT ATCTGCTAAT	1260
25	ACAGTTGGTG GTGTAGCGGC AAAATTGATT TCACCACAAT CAATTGCAAT TGCAACTGCA	1320
	GCAGTAAAC AAGTTGGTAA GGAATCAGAA TTATTAAAA TGACATTGAA ATACAGTGTA	1380
	TGTTTACTAA TATTCATCTG TATTGGACT TTCATCTTGT CATTATTATA AAAAAACGTA	1440
30	TTTCAAATA TAAATATACA GAAGGTGAGA TGTTTTCTAA CATCTCATCT TTTTTTTATG	1500
	GATCATTAAAT GAAAGAAGTT TGACATTATA ATAATGGTAG CGCTTTATGT TAAATGAAT	1560
	AGTGAGTAAT CAGCAATCAA ATTAAATTGG TTGATAGCTG TTAAGGTTTG TGGTTTTGTC	1620
35	TTTGTGCTAT CGCnCATAAA GTATATAATT AAAGTAGTTT CGTTATTATA AAATATTAAT	1680
	ATACATAGTA GATAGTAATA GAGCATCACC ATGGGAACCT ATTGAGACAC TTATTGATTT	1740
	AAAGTGGTAT TAATATGTCG TATTTCTCGA ACGTTCCATT ATTCATTTTA AAAAGGGGGA	1800
40	CTGTATTTGT TATGACAACA CAACATAGCA AAACAGATGT CATCTTAATT GGTGGCGGTA	1860
	TTATGAGTGC aCATTAGGAA CATTACTTAA AGAATTATCA CCTGAGAAAA ATATTAAAGT	1920
45	GTTTGAAAAA TTAGCACAAC CTGGCGAAGA GAGTTCAAAT GTATGGAATA ATGCCGGTAC	1980
	AGGGCATTCA GCACTTTGCG AGTTGAACTA TACAAAAGAA GGTAAGGATG GCACAGTTGA	2040
	TTGTAGTAAA GCAATTAAGA TAAATGAGCA GTACCAAATT TCAAAACAGT TTTGGGCATA	2100
50	TTTAGTTAAA ACAGGACAAT TAGATAACCC AGATCGCTTT ATTCAAGCGG TGCCACACAT	2160
	GAGTTTGTG ATTGGCGAAG ATAATGTAGC TTTTATAAAA AGTCGTGTTG CAACGTTAAA	2220

	GGTACCGTTA ATGATTGAAG GTCGTAAGTC TGATGAACCA ATTGCTTTAA CTTATGATGA	2340
	AACTGGTACa gATGTAACT TTGGTGCGTT AACTGCAAAG TTATTTGATA ATTTAGAGCA	2400
5	ACGTGGTGTG GGAATTCAAT ATAAGCAGAA TGTATTAGAC ATCAAGAAAC AGAAATCTGG	2460
	GGTATGGCTA GTTAAAGTTA AAGATTTAGA AACTAATGAA ACGACAACAT ATGAATCTGA	2520
	TTTTGTATTT ATTGGTGCTG GCGGTGCGAG TTTACCATTa CTCCAAAAGA CTGGGATTAA	2580
10	ACAAATCAAAA CATATTGGTG GTTTCCCGGT AAGTGGATTa TTCCTGCGCT GTACAAATCA	2640
	AGAAGTGATT GATCGTCATC ATGCTAAAGT GTACGGAAAA GCAGCAGTGG GTGCGCCACC	2700
	AATGTCAGTG CCGCACTTAG ATACACGTTT TGTAGACGGC AAGCGTTCAT TGTTATTTGG	2760
15	TCCATTTGCA GGTTCCTCAC CTAAATTTTT AAAAACAGGT TCACATATGG ATTTAATTAA	2820
	ATCGGTTAAA CCAAATAATA TCGTGACGAT GTTATCTGCA GGTATCAAAG AAATGAGTCT	2880
20	TACGAAGTAT TTAGTGTCAC AATTGATGTT ATCTAATGAT GAGCGTATGG ATGATTTAAG	2940
	AGTCTTTTTTC CCAAATGCTA AAAATGAAGA TTGGGAAGTG ATTACAGCAG GGCAACGTGT	3000
	CCAAGTAATC AAGGATACTG AGGATTCTAA AGGTAACTTA CAATTTGGTA CTGAAGTTAT	3060
25	TACGTCAGAT GATGGCACAT TAGCTGCATT ACTTGGTGCA TCACCTGGTG CGTCAACAGC	3120
	TGTAGATATT ATGTTTGATG TTTTACAGAG ATGCTATCGT GATGAATTCA AAGGATGGGA	3180
	ACCAAAGATT AAAGAAATGG TGCCGTCATT TGGTTATCGC tTAACAGATC ATGAGGATTT	3240
30	ATATCATAAA ATTAATGAAG AAGTAACTAA GTATTTACAA GTTAAATAAT AAACGAAACG	3300
	GTAATGTCTT TTTTAATGTG ATAGACATTA CCGTTTTTTT GTGGTTAATA AAAATCATTT	3360
	TAATTGTTTC AGTTGCTTGT TAATAGTGTC TACGTAGTTC TTGTTTTTAA AGAATTGAAT	3420
35	TATCCAAATT AATACATAAA CCACAATGAA GATAATTGTG AATATGATTA GATAATGCAC	3480
	TGTTAGTGGA AACCAACCGG CAAGCATTGC TAAAGGCAAG AATCCGACAT ACGTTGTTAT	3540
	GAAATGCATT ATAGTTGCTT TAGTAATGCT CCAATCTGTG TATTTAAAGA TAAAATCTCC	3600
40	AAGGAAAAAG ACGACGCCTA TGAGTAACCA TAAAATGATA GAAATCAACA TTACGGTAGT	3660
	TTCTGTGAAA TGCGTATAAT ACAATATGCC AATAGTTGAT TGTGGGTTCa GTGGATAATA	3720
45	TTTGCCGTCT GCAAATAACA TACTAAAGAA CAGTGAAAGG GACAAACCAA TGATTAAGCT	3780
	AATAAATAAT GAGTTTTTCA AATTTTTTCAT ATTGATAAGC GCTCCTTTAT AGATTTTAAA	3840
	TAACGTCTAG AAGAATAGGT GTAGTGTGCA TCTTTAAGAT ACATACGTAT AAGTCCATTT	3900
50	GGCTCTAATA ATAATTTTTT AATGTAATAC TTGTTGACGA TTTCTGATTT GGAAATGCGA	3960
	ATGAAATGTT GTGGTAACTG TTTTCTAGT TCATAAAGTC GTAATTTTAG TTTGAATTTT	4020

55

ACATTAATGA TATGGATTTC TTTGTCTATG TATCCGACTA ATGTATGTGA TTTGTCTAAA 4140
 TCATTGACTG CATTATAAT ACTTTGAACG TTATCATTCA TTTTAGGTGC ATGTATATCA 4200
 5 ATATAAGATT CCGTCTCATT TGCATTGATA AATAAATTGA GTTTCATCAT AGGTTAATGC 4260
 CTCCTTCAAA ATTATTAAAC CATAAATGAC CATCGATATA TTTAAATTTT GTTGAATGGT 4320
 AGAAATTAAA TGTTAAGTGG CTAGAAAGCG CTAATCAATA TAAAAGATAC CTCCTGAAAT 4380
 10 AAAAACAGAA ATGTTTTTTC AGGAGGTAGA GATTAAAGTG AATTATTTGG CAGTGTAATA 4440
 GTAAAGGTGG TTACATACTC GTTACTTTGT GTGAATTGGA TTGTACCATG ATGCAATTCA 4500
 ATGATGGATT TTGTAATTGC AAGACCTAAA CCATTGCTAT TATCATGTTT GCTCACTTTA 4560
 15 TAAAAACGTT CAAATAAACG TGCTTCAGCT TGTGGACTAA TTGGTGAACC ATCATTACTT 4620
 ATTGTGAAAA TGATATTGTT GTGACTATGT TGCAAAGCGA TGCAATGGC ACCACCAACA 4680
 20 TCTGTATACT TAATAGCATT TATTAATAAA TTACTCAATG CTGATGTAA CAAACGTTGA 4740
 TTTCTAGGA AATTGATGAT TCTAGGTCAG CTAANATGAT TAACGACTTT TCATCAGCAG 4800
 CANATTGTTT ATGTCGAATG ATATCnTTAA TGAGCTG 4837

25 (2) INFORMATION FOR SEQ ID NO: 159:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1600 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

35 ACAATTATTG GATTATTATC AAGCAACGTT AATGGATGAC TTCCACTTAC AACAGAAATG 60
 CCCATAGATT CTAAATCTtT TGCATGAGCA TCTTGTGATA AGTCTTTTCC ATCATTGACA 120
 40 GTTACATTCTG CACCTAATTT ACTTAATAAT TTAGCTGCTT CATAACCACT TTTTGCCAAA 180
 CCGACAACCTA ATACATTTTT ATTTTCTAAC CCTGTATAAT TAAGCATCTT AATGCACTCC 240
 AATCCATAAA CCGATTAAAC CTGAAATCAG ACCAACAGCC CAAAATACTG TAACTACTTT 300
 45 CCATTGCGTC CATCCTATCA ATTCAAAATG ATGATGAATC GGACTCATTT TAAATATACG 360
 CTTTCCAGTC AATTTAAAGC TAGCGACTTG TAACATAACA GATAATGTTT CAATTACGAA 420
 TACTAAACCT ATAAAAATTA ATGATAATTC CTGATTAAGC ATGATTGAAA TGGTAGCAAA 480
 50 TATACCACCT AAAGCTAAGC TACCTGTATC TCCATAAAC ACTTTAGCAG GGTTAATGTT 540
 ATATGGTAAA AATCCTAAAA GTGCAAACAA CATAATGATA CAGAAAATAC CAATTGCCGT 600

55

TGCTAATCCA TCTAAACCAT CTGTTAAATT TACTGCATTA GAAAAACCTA CTTGCCAAAA 720
 AACAAATGAAA ATAACATATG CAAATGATAG TGGGATTGCT ACATTCGTAA ATGGAATATG 780
 5 TATGCTCGTA GAAAAATTCA CCAAATGAAA CACATTACTT AAAACAAAGA ATATAATCGC 840
 AATACCAATT TGCGCCAAAA ACTTCTGTTT ACTTGTTAAA CCTTGGTTAT TCTTTTAAAC 900
 AACAAATAATA TAATCATCTA TAAAACCAAT TAACCCAAAA CCAATCGTCA CAAATAATAA 960
 10 CAGTATGATT GGATTAGCTT GATCTACAAA TATAATAGCC ACCAAAGACG TTATCACAAT 1020
 ACTTAATAGA AATGTTAGTC CACCCATCGT TGGTGTACCA GTCTTCTTCA TATGGCTTTG 1080
 TGGACCTTCT TCTCGAATAC TTTGACCAAA TTTTCATCCTT TTTAATGTAG GTATTAAAAAC 1140
 15 AGGTACCAAA ACAAAATGTAA TCACTAGCGC TAATAACGCA TATACAAAAA TCATAACTAT 1200
 CTCCTCTTCT TAATCCAGAC TTTTTTAAAC ACTAATATAT TATCAATTTT TCAATTAAAT 1260
 AAACAAAGTT GTAATCAAAA TTTATAATTT TTCTTTTTTA CGGCATAAGA GGCCAGTATA 1320
 20 AAAAGTTTGC CTATAACAAA CAAGTTAATC TGACCTCGTC TACCTTAAAA TTCTCTATCA 1380
 ACACCTATTT ATAAAGATTA AATGAAGATG TTGTTTTCTA TCACAGCATT ACTTTAGTAA 1440
 25 AAACAAATAG TGACAATACA TCCTAATTTA ATGTAGCCAT TCTTGTTAGT CCGACTTATC 1500
 CTTGTCAAGT TTAATGTCAG ATTTCTCTCT ATCATCTGAA TTTGAATCAG AATTATTCGT 1560
 CGAATTGCTG TCTACATTCT CTGGATGGAA AATTCTACGT 1600

(2) INFORMATION FOR SEQ ID NO: 160:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1186 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

40 ATTGCCTTTG TTTTAATTTT AAATCAAAAT mGCCTATGAA AGATTTAAAT CAATTAATTT 60
 CTATAATATT ATCATTTTTA AAGCATATCA TTGTTTAGTT TTTTATAAT TGGATAAATA 120
 45 CTAATAGTTA CTTTATAAAA CATTACATAG AGAAAGGTTA AGGAGTGCAC ATGTCGAAAA 180
 AGGATCACTC TTCTTCAAAA TACCTTAATT CTGTTAAGGA AGCGCAAGAG GAGTCAAAAA 240
 AGAAAAATAA AAGTAATCCC AAAATTGATG TTGATCGTAC ATATATTGAA CCTCAACAAT 300
 50 TCCAATCTAA GAAACCTAAA AAAGATGATC AGGTTTTCTT CTTATCAAGA TTAAATAAAC 360
 CTGCAAAATA TAAGAAAGAC TCTAATTTCT TATCATATCT CATCTATCGC ATAGGAAAAG 420

5 TGTGCTTTT CCTATTAACA TTATTACCAT TTTTCAATAT TAAGCAGAGT CAAATTACTA 540
 ATATGTTAAG CAATGCACCC GCTGAAACAT CTA CTACTCTAAT TAAGAGTGTA ATTGGTGATA 600
 10 TAACTCAAAA CTCCAGTGGT GGCTTATTAT CTATCGGTTT GATTTTAGCA ATTTGGTCAG 660
 CTTCAAATGG AATGACTGCA ATTATGAATT CTTTCAATGT TGCTTACGAT GTAGAAGATA 720
 GCCGTAATGG AATCGTATTA AAACCTACTAA GTGTGTCTT CACTGTAGTT ATGGGCGTTG 780
 15 TGTTTGTAGT TGCTCTAGCA TTACCAACGC TTGGTCTGT AATTAGTCAT TTCCTATTTCG 840
 GTCCACTTGG aTTTGACGAA CAAGTGAAAT GGATTTTAA CTTATTAGA ATTGTGTTAC 900
 CAATCATTAT TATATTTATC ATATTTATCG TGTATATTC GGTTGCACCT AACGTTAAAA 960
 CGAAGCTTAA GTCAGTATTA CCAGGTGCAG TATTTACTTC AATTATTTGG TTAGCTGGTT 1020
 CATTGGTTT TGGTTGGTAT ATTTCAAATT TTGGTAACTA TTCTAAAACA TATGGCAGTA 1080
 20 TCGCGGGTAT CATCATTTTG TTA CTATATGGT TATATATCAC AAGTTTTATT ATAATTGTCC 1140
 GnGCTGAAAT CAATGCAATC ATTCATCAGC GTAGTGTAAT TAAAGG 1186

(2) INFORMATION FOR SEQ ID NO: 161:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7872 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

35 TCTTGAGCCA TCTTTTGAGC TAACTGACTA GATTGATACC CAAAAATCAT AGTTACCAAC 60
 ATAAACTTTA ATTTTACCGA AGTCTAAATC AGCGATATGA GTACATACAT TATTTAAGAA 120
 ATGATCGGTCA TCGGATACTA CGATAACAGT ATTATCAAAG TTAATTAAGA AATCTTCTAA 180
 40 CCAACTGATT GCTGGAATAT CGAGACCGTT AGTAGGCTCA TCCAGTAATA GTACGTCTGG 240
 TTCACCGAAT AAACCTTGCG CTAATAATAC TTTAATTTTT TGGTTGTTTT CTAATTCAGC 300
 CATTTTTTTA TCGTGTAAG TTGGATCGAT ACCTAAACCA GATAAAAGGT TAGCAGCATC 360
 45 AGCTTCAGCA TTCCAACCAT TCATTTCTGC AAATTCACCT TCAAGTTCAG CAGCACGGAT 420
 ACCATCTTCA TCACTGAAAT CTGGCTTCAT ATAGATTTCA TCTTTTTCTT TCATAACCTC 480
 ATAAAGACGT TCGTGACCTT TAATTACAAC ATCAAGCAGC CGTTCATCTT CATAAGCATA 540
 50 GTGGTCCTGT TTTAAACAG CTAGACGTTT ATTTTCCCT AATGAAACAT GTCCTGTTTG 600
 AGAATCTAAT TCACCAGATA ATATTTTAA GAATGTTGAT TTACCTGCAC CATTGCGACC 660

55

	ATCTCCAAAA CGTAAACTCA CATCAGTTAC TTGTAACATG CATTTTCTCC TTTTTTTCAT	780
	TCGATATTCT AACGGAAGAA TTATATCATA TTATCGTCAC AGTTTCGACC TCATATAAGT	840
5	TGTAATGATA GAATGACTCA CACATGTTAT AATAATAAAG AATACAAGAA TCGAAGGAGA	900
	ATAACATGGC ATTAGACAAA GATATAGTAG GTTCTATAGA ATTCCTTGAA GTAGTAGGGT	960
	TACAAGGTTC AACTTACCTT TTAAGAGGAC CAAACGGTGA AAACGTAAAG TTAAACCAAT	1020
10	CAGAAATGAA CGATGATGAT GAATTAGAAG TAGGTGAAGA ATATAGTTTC TTCATTTATC	1080
	CAAACCGTTC AGGTGAATTA TTTGCAACTC AAAATATGCC TGATATTACG AAAGATAAAT	1140
15	ATGACTTTGC TAAAGTACTT AAAACGGATC GCGATGGGGC ACGTATAGAT GTTGGATTAC	1200
	CCCGTGAAGT GTTAGTACCA TGGGAAGATT TACCAAAAGT GAAATCACTA TGGCCACAAC	1260
	CTGGTGATTA TTTGCTAGTT ACATTACGAA TTGACCGTGA GAATCATATG TATGGACGTT	1320
20	TAGCGAGTGA ATCTGTTGTA GAAAATATGT TTACACCTGT ACACGACGAT AATTTAAAAA	1380
	ACGAaGTCAT TGAAGCCAAA CCTTACCGCG TATTACGAAT TGGTAGCTTT TTATTAAGCG	1440
	AATCAGGTTA CAAAATTTTC GTACATGAAT CAGAACGTAA AGCTGAACCA AGATTAGGTG	1500
25	AATCTGTTCA AGTTAGAATT ATCGGGCATA ATGATAAAGG TGAGTTAAAT GGTTCATTTT	1560
	TACCACTTGC ACATGAACGT TTAGACGATG ACGGCCAAGT CATCTTTGAT TTACTAGTTG	1620
	AATATGATGG TGAATTACCA TTCTGGGACA AATCAAGCCC TGAAGCGATT AAAGAAGTAT	1680
30	TCAATATGAG TAAAGGTTCA TTCAAACGTG CAATCGGTCA CTTATATAAA CAGAAGATTA	1740
	TTAATATAGA AACAGGTAAA ATCGCTTTAA CTAAAAAAGG TTGGAGTCGA ATGGACTCAA	1800
	AAGAATAATC ATTTTACAC GTGTCGTAGG ATGCGTGTTC TTTTATTCA ATATTAAATC	1860
35	GGACAGATGA AGTAGTTTTT TAAACATTCC TTTCAAAGTA AAAAATTAAA TAATTCAAAC	1920
	GAATAGGCTG GGaCATTAAAG TTCTTAGGCA ATGTAAAAAA GCTGATTCTT ATTAATTATT	1980
40	TGATGGAAAT CAGCTTTTTT GATATGTATT TTATAATGTA CAGCTCGTTG AGCTGCTATT	2040
	TTCTTATAT TAAGTGCCAT TAATACAAA CCTAGCTCTC GTTTAACTTT ATTTATTCCT	2100
	CGAACTGACA TTCGAGTGAA aCCCAAAATA GCCTTCATAA ATCCAAAAAC AGGCTCTACA	2160
45	TAAATTTTTT TATGACTATA GATTTTTTTC GTTCTGGTT CAGAAAGCTT TTGaTTAATT	2220
	TGGGCTTTAA TGTATTTCAA AGTAAATTA CATGTTAATA CGTAGTATTA ATGGCGAGAC	2280
	TCCTGAGGGA GCACTGCCAG TCGAAGACAG GGGCCCCAAC ACAGAAGcTG ACATATAGTC	2340
50	AGCTTACAAC AATGTGCCGG TTGGGGTGGC TGAGACGGCA CCCTAGGAAG GGACCCGTCA	2400
	TCAAAAATTC TATTTATAGA ATTTTACAGT AATGTGACAG ACGGGCAAAG CGAAgCCATT	2460

55

	CTTACTGCTG TTTTTTTAGG GATTATGTC CCAGCCATTT TTGTATTCAT ATTTAAATTT	2580
	CGATAATTTT TCAGGAAGCA TTTTAATTTT ACTAATGAAG CAATATTTT TAGATTAAAC	2640
5	AAAATTAATA TTTACATTTT CTTAACAATT TTTTATGTAA CATTTACAGT TTCTAAAAAT	2700
	GAGGTTAATA ATTCAAGGTT AAGATAAAGA TGTAATCAAT ACAAATACTA TTTGTTGTTC	2760
	ATACAGGGAG GATATTTCAA TGAAAAATG GCAATTTGTT GGTACTACAG CTTTAGGTGC	2820
10	AACACTATTA TTAGGTGCTT GTGGTGGCGG TAATGGTGGC AGTGGTAATA GTGATTTAAA	2880
	AGGGGAAGCT AAAGGTGATG GCTCATCAAC AGTAGCACCA ATTGTGGAGA AATTAAATGA	2940
	AAAATGGGCT CAAGATCACT CGGATGCTAA AATCTCAGCA GGACAAGCTG GTACAGGTGC	3000
15	TGGTTTCCAA AAATTCATTG CAGGAGATAT CGACTTCGCT GATGCTTCTA GACCAATTAA	3060
	AGATGAAGAG AAGCAAAAAT TACAAGATAA GAATATCAAA TACAAAGAAT TCAAAATTGC	3120
20	GCAAGATGGT GTAACGGTTG CTGTAAATAA AGAAAATGAT TTTGTAGATG AATTAGACAA	3180
	ACAGCAATTA AAAGCAATTT ATTCTGGAAG AGCTAAAACA TGGAAAGATG TTAATAGTAA	3240
	ATGGCCAGAT AAAAAAATAA ATGCTGTATC ACCAAACTCA AGTCATGGTA CTTATGACTT	3300
25	CTTTGAAAAT GAAGTAATGA ATAAAGAAGA TATTAAAGCA GAAAAAATG CTGATACAAA	3360
	TGCTATCGTT TCTTCTGTAA CGAAAAACAA AGAGGGAATC GGATACTTTG GATATAACTT	3420
	CTACGTACAA AATAAGATA AATTAAAAGA AGTTAAAATC AAAGATGAAA ATGGTAAAGC	3480
30	AACAGAGCCT ACGAAAAAAA CAATTCAAGA TAACTCTTAT GCATTAAGTA GACCATTATT	3540
	CATTTATGTA AATGAAAAAG CATTGAAAGA TAATAAGTA ATGTCAGAAT TTATCAAATT	3600
	CGTCTTAGAA GATAAAGGTA AAGCAGCTGA AGAAGCTGGA TATGTAGCAG CACCAGAGAA	3660
35	AACATACAAA TCACAATTAG ATGATTTAAA AGCATTATTAT GATAAAAATC AAAAATCAGA	3720
	CGACAAGAAA TCTGATGATA AAAAGTCTGA AGACAAAAAA TAATAAGACG CAATTTCAAA	3780
40	TGTGTCTTGA AACATGATTT TGATGGTGAA TCATTATTTA GAGTACAAAG CTTGATTAT	3840
	CGAGACGCTG ATTTTGACAT TCAGTTAGTC TACAAGCTTA TCAACTTAAA ATAGTGGTTC	3900
	ATCATTATTT TACAAATCTA ATTATTTTGG GAGTAATAGA AAGAGGTTTG ATTATGACTT	3960
45	CATCTACTAA TGTTAAAGCT TTAATCGAAA AAAATAATAA TAAAAAAGGA AAGCATAATG	4020
	ACAAAATTAT ACCAGTTATT TTAGCCGCAA TTTGAGCGAT TTCCATTTTA ACAACACTAG	4080
	GTATATTAAT CACATTGCTT TTAGAAACCA TCACTTTTTT CACCAGAATT CCAATAACTG	4140
50	AATTTCTATT TTCTACTACT TGGAATCCTA CCGGTTGAGA CCCTAAGTTT GGTATCTGGG	4200
	CATTGATAAT AGGGACTTTA AAAATCACAG TTATTGCGAC TATATTTGCA GTTCCAGTCG	4260

55

	AACCGATATT AGAAATTTTA GCAGGAATAC CAACAATTGT GTTTGGTTTC TTTGCATTAA	4380
	CCTTTGTTAC ACCAGTATTA AGATCTTTCA TACCAGGTCT TGGAGAGTTT AATGCTATAA	4440
5	GTCCCGGCTT AGTTGTCGGT ATTATGATTG TCCCTCTCAT CACAAGTTTG AGTGAGGATG	4500
	CAATGGCATC TGTACCAAAT AAAATTCGAG AAGGTGCCTA TGGACTTGGA GCAACTAAAT	4560
	TAGAAGTAGC AACTAAAGTC GTACTTCCCG CAGCAACATC AGGTATTGTA GCTTCAATCG	4620
10	TTCTCGCGAT TTCAAGAGCA ATTGGAGAAA CGATGATTGT ATCATTAGCG GCAGGTAGTT	4680
	CGCCAACAGC TTCATTAAGT TTAACAAGTT CGATTCAAAC AATGACTGGA TATATTGTTG	4740
	AGATAGCGAC AGGTGATGCA ACATTTGGAT CAAATATTTA TTACAGTATT TATGCTGTAG	4800
15	GGTTCACACT ATTTATCTTT ACCTTAATCA TGAATTTACT TTCTCAGTGG ATTTCTAAGC	4860
	GTTTTAGGGA GGAGTATTAA TATGGAAACG ACAGATAATA ATAGACAATC ACTCGTCGAT	4920
20	CAACAACCTG TCCAAAAACA TTTATCATCC AGAACGGTTA AAAATAAAGT GTTCAAACCTC	4980
	ATATTTTTAG CATGTACATT ATTAGGACTT GTCGTACTTA TTGCGTTGTT AACTCAAACA	5040
	TTGATTAAAG GGGTAAGTCA TTTAAATTTA CAGTTTTTCA CTAATTTTTT TTCTTCAACA	5100
25	CCATCTATGG CTGGCGTTAA AGGCGCGTTA ATCGGTTTAC TTTGGTTAAT GTTAAGTATC	5160
	ATTCCATTAT CAATCATCCT AGGAATAGGT ACAGCTATAT ACTTAGAAGA ATATGCGAAA	5220
	AACAACAAAT TTAATCAGTT TGTAAAAATC AGTATTTCCA ATTTAGCTGG TGTACCATCA	5280
30	GTTGTATTTG GGTATTATAGG TTATACTTTG TTCGTTGGTG GTGCAGGGAT TGAAGCCTTG	5340
	AAAATGGGTA ACAGTATATT GGCAGCAGCG CTAACAATGA CCTTACTGAT ATTACCAATT	5400
	ATTATTGTTT CAAGTCAGGA AGCAATTAGA GCTGTACCTA ACTCAGTACG CGAACTTCTT	5460
35	ACGGCTTAGG TGCTAATAAA TGGCAAACGA TAAGACGTGT TGTCTTACCA GCAGCGTTAC	5520
	CTGATATTTT AACTGGATTC ATTTTGTCTC TTTCAAGAGC ACTGGGAGAA ACAGCGCCAC	5580
	TTGTGCTAAT CGGTATACCG ACTATATTAT TGGCAACACC TAGAAGTATA TTGGATCAAT	5640
40	TTTCAGCATT ACCTATCCAA ATATTTACTT GGGCGAAAAT GCCTCAAGAA GAATTCCAGA	5700
	ATGTTGCATC GGCAGGCATT ATCGTTTTAC TAGTTATCTT AATCTTAATG AATGGCGTTG	5760
45	CGATTATTTT ACGTAACAAA TTTAGTAAAA AATTCTAATT TAAACAATCA ATCTCATTTA	5820
	TCTATTAAAA AGGGAGTTTT AAATATGGCG CAAACACTTG CACAACTAA ACAAATATCT	5880
	CAAAGTCATA CGTTTGATGT CTCACAAAGT CATCATAAAA CACCAGATGA TACAACTCA	5940
50	CATTCTGTTA TATATTCAAC ACAAATTTA GACTTATGGT ATGGCGAAAA TCATGCATTA	6000
	CAAAATATTA ATTTAGATAT TTATGAAAAC CAAATTACTG CCATTATAGG TCCATCTGGT	6060

55

	AAAACAGCTG GTAAATATT ATATCGAGAT CAAGACATTT TTGATCAAAA ATATTCTAAA	6180
	GAACAATTAC GTACAAATGT GGGCATGGTC TTTCAACAAC CTAATCCATT TCCAAAATCA	6240
5	ATATACGATA ATATTACTTA CGGTCCAAAG ATTACCGGTA TTAAAAATAA AAAAGTTCTT	6300
	GATGAAATCG TTGAGAAATC ATTACGTGGC GCTGCAATTT GGGATGAATT AAAGGATAGG	6360
	TtGCACACAA ATGCATATAG TTTATCCGGT GGGCAACAAC AACGTGTTTG TATCGCGCGT	6420
10	TGTTTAGCAA TTGAACCTGA AGTCATTTTA ATGGATGAAC CGACATCAGC ATTAGATCCA	6480
	ATCTCAACAT TAAGAGTAGA AGAGTTGGTT CAAGAATAA AAGAAAAGTA TACAATTATT	6540
	ATGGTtACAC ATAATATGCA ACAAGCAGCT CGTGATCAG ATAAACTGC ATTTTCTTA	6600
15	AATGGTTATG TCAATGAATA TGATGATACT GATAAAATTT TCTCTAACCC ATCAAACAAG	6660
	AAAACAGAAG ATTATATTTT AGGAAGGTTT GGTGATATA TAATGGCAAT AATTAGACAA	6720
20	CGATATCAGG AGCAACTTGA TGATTTAATA AAAGAATTAC GTCGGTTAGG TGCaAATGTC	6780
	TATGTGAGTA TTGaAAATGG TATAAAAtCA TTAAGTATTG aCGATAGAGG cTTTGACGA	6840
	CAAACAGTTA AAAACGATAA ACATATCAAT CAATTAAATT ATGATATTAA TGAGCGAGTT	6900
25	ATCATGTTAA TTACAAAGCA ACAGCCCAT TCGAGTGATT TCGGTATGAT GATTTCTTCA	6960
	TTAAAAATCG CCTCCGATTT AGAAAGAATA GGAGATAATG CCTCGAGTAT TGCCAATATT	7020
	CGATTGCGTA CAAAGATTAC AGATGATTAT GTGTTAACCC GTTTAAAGAC AATGGGTAAA	7080
30	TTAGCTATGT TAATGTTAAA GGACTTAGAT CAAGCATTTA AAAAGAAAGA TACCGTATTA	7140
	ATAAGAGAAA TAATTGAGCG TGATGAAGAT ATCGATGACT TATATAGTCA TATTATTAAC	7200
	GCAACGTATC TTATTGATAA CGtCCATTTG TCGCTGCACA AGCTCATTTA GCAGCAAGAC	7260
35	ATTTAGAACG TATTGGTGAT CATATTATTA ACATCGCTGA AAGTGTttAT TTTTATTTAA	7320
	CAGGTACACA TTACGAACAA TAACTTAAAG TTATTACTAT AAAATCCCTT ACGATAAATA	7380
	TATATTTCTA TTATTCATAA ACCCTCAAAA AAACCAAGAT TCTCACAATT AGTAATGTGA	7440
40	AAATCTTGGT TTATATTGTT CTACTATAAA TTGTCTCGCA TCTTAGTTAT TTGCTTGCTC	7500
	AATTTCACTT GTTAATTTTT CAACTTCATC GACTAAATCA GAAATATATT GAATTGTAGA	7560
45	TTTAAGTGGC TGTTCTGTAG TAATGTCTAC ACCTGCAATG TTTGCAAGTT CGACAGGTGA	7620
	TACACTACCA CCTTTTTTCA ATGTTTCTAA CCAAGCATCA ACAGCTGGTT GGCCTTCATT	7680
	TTTAATCTTT TGAGAAACGA CAGTTCCGAT TGTTAAGCCA GCAGAATACG TATACGAATA	7740
50	TAATCCCATa TAGTAATGAG GTTGACGCAT CCATGTTAAT TCAGCACCCCT CAGTCATGTC	7800
	TACTGCATCT CCAAAAAAATT GTTTATAAAC ATTTAGCATT ATTTCAATTA ATGThCGGCG	7860

55

(2) INFORMATION FOR SEQ ID NO: 162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

```

TTTTTCTTT TCTTCATTG AAAATTGATC ATTCAGCAAT ATAAGCGTAT TTGTTAATGA      60
TTTAGGTGTT CCAATTTTCAT AATCCCACCA ATTTAAGTTG GTATTCTTGC CAGTTGTTTT      120
AGTAAAATTC TCACTTAATT CTTTACTTTT TTTATCTGGT TCTTTTCCAT ATGCATTTTT      180
ATGCAGCCAC TCAAGGGCAT CTTTCACTTT CTTCTTATTT TCGTCAGTAT TTAAAGTGGT      240
TTTAGGATTC CTCATCGCTT CTGCGATTTT CTCAATATTA CGATAGGTAC GAGTCATATG      300
AGAAGAATTA GTTTCAGGG TTTCCGCTCC TGACCACAAG TATTCCTAC CACTTTCAGT      360
TTTCATTTCC TTGAGTAAAT TCGTCGCCTC TTTCTCTGTA GCATCAAAC TCTTCTTCAT      420
ATCTGGATTA TTCTCATCAT ACTTATCATA ACCATAGTTA ACGTCCAGCC ATGTGTTTCT      480
CAATTTTTCA TAATCTGGCG TTTGAACATT CGTATCAGCC ACAGCGATTT GATGTTTATC      540
AACACTTCTG AATTCACCAC CATTCAAAGT AATCACACCA GCCATTAATA ACGTAATGGT      600
GGATAATTTT TGCCATTTCT TTAATTCTATA TGTCATTGac ATGTCTCCTT TTTGTGTTGC      660
GCGTGCGCAA TGAATATTAT GATTAAATAA TGATTCAATT TTTCAAAATT CGTTAACGTA      720
TACAAATGAC TGTCTACTGT CAAACAATCC ACAAAGAATG TTGATGtCAT ATaACAATC      780
GATCACCCAA ATTTTCCG

```

(2) INFORMATION FOR SEQ ID NO: 163:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5132 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

```

TACAGGTTTT ACTATAATGG ATGGTATTTT GGCTAAACGA CATTGGTTTA GTCTTCTTTT      60
TTTnACTTCC TanATTTTACA ATGGTATAAA TAATAATGCT ATATTTAGAA TGATGAGTAT      120
ACTTACTGAA ACTAAATTAA AAGTGTCTGG TTCTTTACTA AAGATAGCTG CTATCCTTGC      180

```

	AATACAAGTT	CCAATGAGCG	CAATTA AAAAG	TACTAACCCA	ACGATGAAAC	TCTGTTTGTC	300
	ACTTAACTCA	AAGAACTAT	AGATAGGATA	TTTTTTAATA	ATCAAGCCAC	CTAAAATCAT	360
5	CCATAAAAAT	ACGATAATTC	CATAAGTCAC	ATTTATAACA	TACGTTATTT	TTTGGTCACC	420
	AAATCGGACT	AATGTATTTT	GTAGAATCAG	CATACCAATG	ACAACACCTA	AAATAACGAT	480
	ACTAGCTATA	TAAAGTAAAA	ATGCAATTGT	CACATCAAAT	GTACCCAAAT	CTAAAAACCT	540
10	AGGAATTAY	AyGACTGCTA	AAATAAAAGC	GAAGyACAAA	GTAATATAKT	TATACAAACC	600
	GGTAGTAAGA	CTTATCTCAG	GTGATAATTG	ATCAGCCATT	GACTTAATCG	GTGTATTAAT	660
	AATTGAACCT	GTATCTTCGT	TATTTTTTTC	AGCCATAGTT	AAATGATCTT	CGAGCTCTTC	720
15	CAATAACTCT	TCTACTTCTG	CTTCAGTCTT	ACCTCTAAAT	AACAATTCAA	CACGTAATTT	780
	TTCTAAAAAA	TCTTGAGATT	GTTTACTTAA	CATCGTTTTT	CCCTCCAAAC	AAGTTAATCA	840
20	TCCCTTTATT	CAAAACTTGC	CATTTTCGATT	TAAATACTTT	TAGTTCCTTT	AAACCTGAAT	900
	CGGTAATCGT	ATAGTATTTT	CGCCTCGGGC	CGCCATTACT	AGATTTTTTT	ATTGTCGTAT	960
	CAACGTATCC	TTTTTTGTTT	AAACGCATTA	AAACTGGATA	AATACTACCC	TCACTTATCT	1020
25	CTGGAAACTC	TTGATTCTTA	AGTTTCGTCA	TAATTTTATA	TCCATACGTT	TCGCCTTGGG	1080
	CAATGAGACC	TAATATCGCC	CCATCTAAGA	GACCTTTTCT	AATCTGATCT	GACACTGACA	1140
	TTTTAATCAC	CTACTATCTT	ACATAATAAG	ATAGTACATT	GAGAACTTTT	CGTCAACTAT	1200
30	CTTTTATTGT	AAGGTAGTTG	TTGTACACAT	TCCTTAAATG	ACTAACAAC	TTGTTAATAG	1260
	GGTAATACTT	ACGGAAGTAT	ATTTTATTTA	TGGGGGAGGA	ATTAATAATG	ACTACAAAA	1320
	CAGTATTTGA	TGTCATTGAT	ATGGGGTTAG	GATATTTAGT	AAATGTGTAT	GATGCTTGGA	1380
35	AAGTTGAAAA	GGTACTTGAT	GATTATCATA	AGCCTTTTTT	TAATACCATT	CATTGGCAAT	1440
	TTGGCATGT	ATTAACAATT	TTGAATCGG	CCTTAGCTGT	TGCTGGTAAA	GAGAATATTG	1500
	ATTTAAATAT	CTATAGACCT	TTATTCGGAA	ATGGTTCGTC	TCCAGATGAA	TGGAAGGATG	1560
40	AAGTACCGAG	TATTGAAAGG	ATTTTAGAAG	GTCTCCAAAC	TTTACCTGAA	CGTGCACGAA	1620
	ATCTAACTGA	AGATGATTTA	GCAATTGAAT	TGAAACAGCC	AATTGTCTGGT	TGTAATAACT	1680
45	TAGAAGAGTT	ATTAGTATTA	AATGCCATTC	ACATCCCACT	TCATGCTGGT	AAAATTGAAG	1740
	AGATGTCTCG	TATATTAAAA	AATTTAAAT	AAATATGTGC	TTATTAACCG	TTAACAACAC	1800
	GTTAACGGgT	TTTTTATTTG	TTTAAAAGGT	CACTTTTTTG	AATTTAATAA	ACACCATCTA	1860
50	TACCAGTTCT	TCACCGATTC	TCGAAAAATA	ATTATATTAA	TGATTTCTGTT	AATTTAATTT	1920
	TATATTTAAT	TATTACTGTA	CATCTTTTGT	AGTTAGCTTT	ATTCTTAAAT	TGAAATATGT	1980

55

	TACTCCCTAT CGTTGTAGGT CTCCTTATTT GGGCACTTAC ACCTTTTAAA CCGGATGCTG	2100
	TGGATCCAAC AGCATGGTAT ATGTTGCGAA TATTCGTCGC GACAATCATT GCTTGTATTA	2160
5	CACAACCGAT GCCAATTGGG GCCGTCTCTA TAATTGGATT TACAATCATG GTACTCGTTG	2220
	GCATTGTTGA CATGAAAACG GCTGTCGCTG GTTTTGGTAA TAATAGCATT TGGTTAATTG	2280
	CTATGGCATT TTTCAATTCTG AGAGGATTTG TGAAAACAGG TCTTGGTAGA CGTATCGCAC	2340
10	TTCATTTCTG CAAATTATTT GGTAAAAAAA CATTAGGATT AGCATATTCT ATCGTCGGTG	2400
	TAGATTTAAT TCTAGCGCCT GCTACACCAA GTAATACCGC GCGTGCTGGT GGAATCATGT	2460
	TCCCAATTAT CAAATCACTT TCTGAATCAT TTGGTTGCGA ACCGAAAGAC GGATCAGCAC	2520
15	GCAAAATGGG TGCATTTCTT GTTTTCACAG AATCCAAGG TAATTTAATT ACTGCGGCTA	2580
	TGTTTTTAAC TGCAATGGCC GGTAACCCCC TTGCACAAAA TTTAGCATCT AGCACATCTA	2640
20	ATGTTACAT TACATGGATG AATTGGTTTC TAGCTGCTTT AGTTCCTGGA CTGTTTCCT	2700
	TAATTGTTGT ACCTTTTATT ATTTATAAAA TTTATCCACC AACTGTTAAA GAAACACCAA	2760
	ATGCTAAGAG TTGGGCTGAA AATGAATTAG CGACTATGGG TAAAATCGCT TTAGCTGAAA	2820
25	AATTTATGAT TGGTATTTT GTGCTGCGT TAACACTATG GATTGTCGGA AGTTTCATTC	2880
	ATATTGATGC AACTTTAACG GCCTTTATTG CGCTAGcATT gTTATTATTG ACAGGCGTCT	2940
	TAACATGGCA AGACATTTTA AACGAAACAG GTGCTTGGA CACATTAGTA TGGTTCTCAG	3000
30	TATTAGTGTT AATGGCCGAC CAATTAAACA AGCTTGATT TATTCCTGG TTAAGTAAAT	3060
	CCATTGCTAC AAGTCTTGGT GGCTTAAGCT GGCCTATAGT CCTGGTCATT TTAATATTGT	3120
	TCTACTTCTA TTCACATTAC TTATTGCAA GTTCTACAGC ACATATCAGT GCGATGTATG	3180
35	CAGCATTACT AGgCGTTGCC ATCGCAGCCG GTGCACCACC ATTATTCAGT GCATTAATGT	3240
	TAGGTTTCTT CGGTAACCTA TTAGCTTCAA CAACACACTA TAGTAGTGGT CCAGCGCCGA	3300
	TTCTATTCTC TTCAGGTTAC GTGACTCAAA AACGTGGTG GACAATGAAC TTAATATTAG	3360
40	GTTTCGTCTA CTTTATTATC TGGATTGGTT TAGGATCACT TTGGATGAAA GTAATTGGTA	3420
	TATTTTAAAA TATTTAAATT AGCGCTCGAA TCTCATTGAT TTGGGCGCTT TTAAATTTGT	3480
45	ATTTAAATC AACCTTTGCT AAATCAAGAC TCCCTTTTAA AAATACGTTT ATCCTTTAAA	3540
	TCATTGCGTG CTTCACTGAA AATTTGTATA AAGATTTAAG TCATTACGTA ACATCACATA	3600
	AAATACATTT CTATACTATT CCGCTTCATT GATTAAACATT ACGTATGCCC TCATAAATCA	3660
50	TCATACAAAA AACACCTTCG TTTAAATTCA TTTTAATTGC GAATTCAACG AAAGTGCCTT	3720
	ATTTCATATT TAATGTTTCA AATTTATACG TCTGTCACTG TTACTGCACA CATACCTCAG	3780

55

TTATAGGGTT TTTGCGACCG GATGTTTCTT CAATTTAATG TATTGAGAAA GACTATATAA 3900
 CACAATACCT GTCCAAATAA ATATAAACGT AATTAAATGA TCTATACTAA AAGGCTCTTT 3960
 5 GAAAACAAAT ATGCCGAGTA CAAACATTAT TGTGTGTCCTA ACGTATTGAA TAAATCCTAT 4020
 TAGCGAAAGT GGAATACGTT TTGCCCCGGC TGAGAATAGG ATTAGTGGTA TTGCCGTAAT 4080
 AGCACCAGAA AATAACAACC AAAATGATGA CATGTTCAAT CCAAATGACA TCTGATGTTG 4140
 10 CTGCCATAAA TAAATAACGT ATATTAGTCC AGCAGGTGCG GTAACAATAC ATTCAATCGT 4200
 AATACTGCTG ATGGCATCAA TATGTACTAC TTTTTCAT AATCCGTATG TACCAAAGGA 4260
 TAACGCTAAT ATAATAGAGA CGATTGGGAA TTCTCCAATC TTGAGCGTCA TATATAATAC 4320
 15 ACCGATGAAT GCGAATAAAA TGGCTAGCCA TTCAAATTTA TTGAATCTTT CTTTTAAAAA 4380
 GATAAGTGGC AGCAAATGC TAACAAGTGG ATTTATATAA TAACCTAAAC TTGTTTGTAG 4440
 20 GACGTGACCG TTCGTTACAG CCCAAATAAA TGTACCCCAA TTTAATGTAA TGACATAGCC 4500
 TGCTACGACA ATCGCTAATA GCTGAATGGG CTTGCCTAAC AATTGATTCA TATCTCGTTG 4560
 AAATGCATTG CGTTGTTTTT GTCCAACCGC GAGTATGAAA ATCATGAATA TTGCTGAAAA 4620
 25 TATAATACGA AAGGCTAAAA TTTCAAATGC GCCTATTGCA TCAACGAACT GCCAATATAT 4680
 AGGTAGTATT CCCCACAGAA TGTATGCACT GAGTGCTAAA AATATGCCTT TTTTATACTC 4740
 TGAATTCACC TTCAAACCTC CTTACTTTCC TAATTTTAA TTTACTGCAT ACGCTCACTT 4800
 30 GGTATGCTA ATATAACGAT TTTACTAATA ATATTTGAT AAAGATATCA TTTTGTATTAT 4860
 ATTTCCCA CA TTTATTCACC AACCACTAAA CAATATTAAT TTTATAAATA ATTCTGTACA 4920
 AATCAGGTA TATTGCCAGA AAGACTACCA TACAACATAA AGGATGGATA CAAATGACTT 4980
 35 TACCTAAAAT TGGAAAGCCT GCAACACGCG CGCTAAATTC ACAAGGTATA TACACATTAG 5040
 AAGCAGTATC ACAATATACG AAGTCATCTC TAATGGAGAT GCATGGCGTT GGTCTTAAAG 5100
 CTATATCAAT ATTGGAACAA GCTTTATTTT AG 5132
 40

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22243 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

AAGTAAATTA TATTATGAAT TTGCTGTCA ATTTCTTAAA GACATTCTTA CCGGAACATAA 60

	TAGAAGCAAT TAATAATGcy mAAGAAAAGA CAGCTAATAA TACCGGCTTA AAATTAAATAT	180
	TTGCAATTAA TTATGGTGGC AGAGCAGAAC TTGTTTCATAG TATTAAAAAT ATGTTTGACC	240
5	AGCTTCATCA ACAAGGTTTA AATAGTGATA TCATAGATGA AACATATATA AACAAATCATT	300
	TAATGACAAA AGACTATCCT GATCCAGAGT TGTTAATTCG TACTTCAGGA GAACAAAGAA	360
	TAAGTAATTT CTTGATTGG CAAGTTTCGT ATAGTGAATT TATCTTTAAT CAAAAATTAT	420
10	GGCCTGACTT TGACGAAGAT GAATTAATTA AATGTATAAA AATTTATCAG TCACGTCAAA	480
	GACGCTTTGG CGGATTGAGT GAGGAGTAGT ATAGTATGAA AGTTAGAACG CTGACAGCTA	540
15	TTATTGCCTT AATCGTATTC TTGCCTATCT TGTTAAAAGG CGGCCTTGTTG TTAATGATAT	600
	TTGCTAATAT ATTAGCATTG ATTGCATTAA AAGAATTGTT GAATATGAAT ATGATTAAAT	660
	TTGTTTCAGT TCCTGGTTTA ATTAGTGCAG TTGGTCTTAT CATCATTATG TTGCCACAAC	720
20	ATGCAGGGCC ATGGGTACAA GTAATTCAAT TAAAAAGTTT AATTGCAATG AGCTTTATTG	780
	TATTAAGTTA TACTGTCTTA TCTAAAAACA GATTTAGTTT TATGGATGCT GCATTTTGCT	840
	TAATGTCTGT GGCTTATGTA GGCATTGGTT TTATGTTCTT TTATGAAACG AGATCAGAAG	900
25	GATTACATTA CATATTATAT GCCTTTTTAA TTGTTTGGCT TACAGATACA GGGGCTTACT	960
	TGTTTGTTAA AATGATGGGT AAACATAAGC TTTGGCCAGT AATAAGTCCG AATAAAACAA	1020
	TCGAAGGATT CATAGGTGGC TTGTTCTGTA GTTTGATAGT ACCACTTGCA ATGTTATATT	1080
30	TTGTAGATT CAATATGAAT GTATGGATAT TACTTGAGT GACATTGATT TTAAGTTTAT	1140
	TTGGTCAATT AGGTGATTTA GTGGAATCAG GATTTAAGCG TCATTTTCGGC GTTAAAGACT	1200
	CAGGTCGAAT ACTACCTGGA CACGGTGGTA TTTTAGACCG ATTTGACAGC TTTATGTTTG	1260
35	TGTTACCATT ATTAAATATT TTATTAATAC AATCTTAATG CTGAGAACAA ATCAATAAAC	1320
	GTAAGAGGA GTTGCTGAGA TAATTTAATG AATCTCAGAA CTCCTTTTGA AAATTATACG	1380
40	CAATATTAAC TTTGAAAATT ATACGCAATA TTAACTTTGA AAATTAGACG TTATATTTTG	1440
	TGATTTGTCA GTATCATATT ATAATGACTT ATGTTACGTA TACAGCAATC ATTTTTAAAA	1500
	TAAAAGAAAT TTATAAACAA TCGAGGTGTA GCGAGTGAGC TATTTAGTTA CAATAATTGC	1560
45	ATTTATTATT GTTTTTGGTG TACTAGTAAC TGTTTCATGAA TATGGCCATA TGTTTTTTGC	1620
	GAAAAGAGCA GGCATTATGT GTCCAGAATT TGCGATCGGT ATGGGGCCAA AAATTTTATG	1680
	TTTTAGAAAA AATGAAACAC TTTACACTAT TAGGTTATTG CCTGTTGGTG GATATGTTTCG	1740
50	TATGGCAGGA GATGGCTTAG AAGAGCCACC AGTCGAGCCC GGTATGAACG TTAAATTTAA	1800
	ACTTAATGAA GAAAATGAAA TAACACATAT CATATTAGAT GATCATCATA AGTTTCAACA	1860
55		

	CACTGCTTAT GATAATGAAA GACATCATTT TAAAATTGCT AGAAAGTCTT TCTTTGTTGA	1980
	AAATGGTAGC TTAGTTCAAA TTGCTCCGAG AGACAGACAA TTTGCACATA AAAAGCCATG	2040
5	GCCGAAATTT TTAACATTAT TTGCGGGACC GTTATTTAAC TTTATATTAG CTTTAGTCCT	2100
	ATTTATTGGT ETTGCATATT ATCaAGGcAC GCcTACGTCT ACTGTAGAAC AAGTCGCAGA	2160
10	TAAGTATCCA GCTCAACAAG CAGGATTACA AAAAGGTGAT AAGATCGTCC AAATTGGCAA	2220
	ATATAAAATA TCTGAATTG ATGATGTTGA TAAGGCGTTA GATAAAGTTA AAGATAATAA	2280
	GACGACTGTT AAATTTGAAC GTGATGGTAA AACAAAGTCA GTTGAATTAA CACCTAAAAA	2340
15	GACTGAAAAA AAAGTACTA AAGTAAGTTC AGAGACGAAG TATGTTCTCG GATTCCAACC	2400
	AGCGAGTGAA CATACTTTT TTAAACCAAT TGTATTCCGA TTTAAAAGCT TTTAATCGG	2460
	TAGTACTTAT ATTTTACAG CTGTAGTAGG TATGTTGGCT AGTATATTTA CGGGCGGATT	2520
20	CTCATTTGAT ATGTTAAATG GTCGGTTGG TATTTATCAT AACGTCGACT CAGTTGTTAA	2580
	AGCGGGTATC ATTAGCTTAA TTGGTtncAC TGCCTTATTA AGTGTAAGT TAGGTATTAT	2640
	GAATTTAATT CCTATTCCTG CACTAGACGG TGGTCGTATT TTATTTGTTA TATATGAAGC	2700
25	GATTTTCAGA AAACCAGTTA ATAAAAAGC GGAAACAACG ATTATTGCTA TTGGTGCCAT	2760
	TTTCATGGTC GTTATAATGA TATTAGTAAC GTGGAATGAT ATTCGACGAT ATTTCTTATA	2820
	ATTTAGGAGG ATAAATAATT ATGAAGCAAT CCAAAGTTTT TATACCAACG ATGCGTGACG	2880
30	TGCCATCAGA AGCAGAAGCA CAAAGTCATC GTTTATTATT GAAATCGGGT TTGATAAAAC	2940
	AAAGTACAAG TGGGATTTAT AGTTATTTAC CGCTAGCAAC ACGTGTGTTA AATAATATTA	3000
35	CTGCAATTGT GCGACAAGAA ATGGAACGTA TCGATTCTGT TGAAATTTTA ATGCCAGCGT	3060
	TACAACAAGC TGAATTATGG GAAGAATCAG GACGTTGGGG TGCATATGGC CCAGAATTAA	3120
	TGCGTTTACA AGATAGaCAT GGAaGACAA TTTgCATTAGG TCCaACACAT GAAGAATTAG	3180
40	TTACATCAAT AGTAAGAAAT GAATTGAAAT CATACAAACA ATTACCGATG ACATTATTCC	3240
	aAATTCAATC TAAATTCGGT GATGAAAAGA GACCACGTTT TGGTTTAYTC GTGGGCGTGA	3300
	ATTTATTATG AAAGATGCAT ATTCATTCCA TGCTGACGAG GCATCATTAG ATCAAACGTA	3360
45	TCAAGATATG TATCAAGCGT ATAGCCGTAT TTTTGAGAGA GTTGGCATT ACGCAAGACC	3420
	AGTAGTTGCA GATTCAGGTG CTATAGGCGG TAGCCATaCA CATGAATTTA TGGCATTAA	3480
	TGCTATCGGT GAGGATACAA TCGTTTACAG TAAAGAAAGT GATTATGCTG CTAACATCGA	3540
50	AAAAGCAGAA GTCGTTTACG ArcCAaATcA TaAGCATACT ACTGTGCAAC CTTTAGAAAA	3600
	AATTGAAACA CCAAATGTTA AGACTGCGCA AGAATTGGCA GACTTCTTAG GTAGACCAGT	3660

55

	GCGTGGCCAT CATGAAATTA ATGACATTAA ATTAAATCT TATTTCCGCA CAGATAATAT	3780
	TGAATTAGCA ACACAAGACG AAATTGTTAA TTTAGTTGGT GCAAATCCTG GTTCACTAGG	3840
5	TCCTGTAATT GATAAAGAAA TCAAAATTTA TGCAGATAAT TTTGTGCAAG ATTTAAATAA	3900
	TTTAGTTGTC GGTGCTAACG AAGATGGTTA TCACTTAATT AATGTAAATG TAGGTAGAGA	3960
	CTTCAACGTT GATGAATATG GCGATTTCCG TTTTATTTTA GAAGGCGAAA AGTTAACTGA	4020
10	TGGTTCAGGC GTTGCACTT TTGCTGAAGG TATTGAAGTT GGTCAAGTAT TCAAATTGGG	4080
	TACTAAGTAT TCAGAATCAA TGAATGCTAC ATTCTTAGAT AACCAAGGAA AAGCTCAATC	4140
	TTTAATTATG GGTGTTTACG GAATTGGAAT TTCTAGAACG CTAAGTGCGA TTGTTGAACA	4200
15	AAATCACGAT GATAATGGAA TTGTTTGGCC TAAATCAGTT ACTCCGTTTG ATTTACATTT	4260
	AATTTCTATT AATCCTAAGA AAGATGATCA ACGAGAACTA GCAGATGCAC TATATGCTGA	4320
20	ATTTAATACT AAATTTGATG TGTGTACGA TGATCGTCAG GAACGTGCAG GTGTTAAATT	4380
	TAATGATGCC GATTTAATTG GTTTACCACT GCGAATTGTT GTTGGTAAAC GTGCATCGGA	4440
	AGGTATTGTA GAAGTTAAAG AACGTTTAAAC AGGTGATAGC GAAGAAGTTC ACATTGATGA	4500
25	CTTAATGACT GTCATTACAA ATAAATATGA TAACCTAAAA TAATTAAGAT CGAATGAATT	4560
	ATAAGAGTAG GAAAAAGCTG AAAGAAATCT GATGCTTATG TCCTGCTCTT ATTATTTTTG	4620
	ATATAATGAT TATTCGATGA AAAATGACTG AAGACATAGT ATAATTAAAG ATAAATTGT	4680
30	TTTAACAATA TAATGATTAG CCAAATATAA AGCATTTAAT TTTCTATCAT TACTATGCTC	4740
	ACATAATCTA AATATTGTTT GAACACGTAA AAGTAATTTT TATTTAAGGT GGTAATTGTC	4800
	TTGGCAATGA CAGAGCAACA AAAATTTTAA GTGCTTGCTG ATCAAATTAA AATTTCAAAT	4860
35	CAATTAGATG CTGAAATTTT AAATTCAGGT GAACTGACAC GTATAGATGT TTCTAACAAA	4920
	AACAGAACAT GGGAAATTTCA TATTACATTA CCACAATTCT TAGCTCATGA AGATTATTTA	4980
	TTATTTATAA ATGCAATAGA GCAAGAGTTT AAAGATATCG CCAACGTTAC ATGTCGTTTT	5040
40	ACGGTAACAA ATGGCAGCAA TCAAGATGAA CATGCAATTA AATACTTTGG GCACTGTATT	5100
	GACCAAACAG CTTTATCTCC AAAAGTTTAA GGTCAATTGA AACAGAAAAA GCTTATTATG	5160
45	TCTGGAAAAG TATTAAAAGT AATGGTATCA AATGACATTG AACGTAATCA TTTTGATAAG	5220
	GCATGTAATG GAAGTCTTAT CAAAGCGTTT AGAAATTGTG GTTTTGATAT CGATAAAATC	5280
	ATATTCGAAA CAAATGATAA TGATCAAGAA CAAAACCTAG CTTCTTTAGA AgCACaTATT	5340
50	CAAGAAGAAG ACGAACAAAG TGCACGATTG GCAACAGAGA AACTTGAAAA AATGAAAGCT	5400
	GAAAAAGCGA AACACAAGA TAACAACGAA AGTGCTGTCTG ATAAGTGTCA AATTGGTAAG	5460
55		

GCAATAGAGG GTGTCATTTT TGATATAAAC TTAAAAGAAC TTAAAAGTGG TCGCCATATC 5580
 GTAGAAATTA AAGTGACTGA CTATACGGAC TCTTTAGTTT TAAAAATGTT TACTCGTAAA 5640
 5 AACAAAGATG ATTTAGAACA TTTTAAAGCG CTAAGTGTG GTAAATGGGT TAGGGCTCAA 5700
 GGTCTGATTG AAGAAGATAC ATTTATTAGA GATTTAGTTA TGATGATGTC TGATATTGAA 5760
 GAGATTAAAA AAGCGACAAA AAAAGATAAG GCTGAAGAAA AGCGTGTAGA ATTCCACTTG 5820
 10 CATACTGCAA TGAGCCAAAT GGATGGTATA CCCAATATTG GTGCGTATGT TAAACAGGCA 5880
 GCAGACTGGG GACATCCAGC CATTGCGGTT ACAGACCATA ATGTTGTGCA AGCATTTCOA 5940
 GATGCTCACG CAGCAGCGGA AAAACATGGC ATTTAAATGA TATACGGTAT GGAAGGTATG 6000
 15 TTAGTTGATG ATGGTGTTC CATTGCATAC AAACCACAAG ATGTCGTATT AAAAGATGCT 6060
 ACTTATGTTG TGTTGACGT TGAGACAAC GGTTTATCAA ATCAGTATGA TAAAATCATC 6120
 GAGCTTGCAG CTGTGAAAGT TCATAACGGT GAAATCATCG ATAAGTTTGA AAGGTTTAGT 6180
 20 AATCCGCATG AACGATTATC GGAAACGATT ATCAATTGA CGCATATTAC TGATGATATG 6240
 TTAGTAGATG CCCCTGAGAT TGAAGAAGTA CTTACAGAGT TTAAGAATG GGTGGCGAT 6300
 25 GCGATATTG TAGCGCATAA TGCTTCGTTT GATATGGGCT TCATCGATAC GGGATATGAA 6360
 CGTCTGGGT TTGGACCATC AACGAATGGT GTTATCGATA CTTTAGAATT ATCTCGTACG 6420
 ATTAATACTG AATATGGTAA ACATGGTTTG AATTTCTTGG CTAAAAATA TGGCGTAGAA 6480
 30 TTAACGCAAC ATCACCCTGC CATTATGAT ACAGAAGCAA CAGCTTACAT TTTCATAAAA 6540
 ATGTTTCAAC AAATGAAAGA ATTAGGCGTA TTAAATCATA ACGAAATCAA CAAAAAATC 6600
 AGTAATGAAG ATGCATATAA ACGTGCAAGA CCTAGTCATG TCACATTAAT TGTACAAAAC 6660
 35 CAACAAGGTC TTAAAAATCT ATTTAAATTT GTAAGTGCAT CATTGGTGAA GTATTTCTAC 6720
 CGTACACCTC GAATTCACG TTCATTGTTA GATGAATATC GTGAGGGATT ATTGGTAGGT 6780
 ACAGCGTGTG ATGAAGGTGA ATTATTTACG GCAGTTATGC AGAAGGACCA GAGTCAAGTT 6840
 40 GAAAAAATTG CCAATATTA TGATTTTATT GAAATTCAAC CACCGGCACT TTATCAAGAT 6900
 TTAATTGATA GAGAGCTTAT TAGAGATACT GAAACATTAC ATGAAATTTA TCAACGTTTA 6960
 ATACATGCAG GTGACACAGC GGGTATACCT GTTATTGCGA CAGGAAATGC AACTATTTG 7020
 45 TTTGAACATG ATGGTATCGC ACGTAAATTT TTAATAGCAT CACAACCCGG CAATCCACTT 7080
 AATCGCTCAA CTTTACCGGA AGCACATTTT AGAACTACAG ATGAAATGTT AAACGAGTTT 7140
 50 CATTTTTTAG GTGAAGAAAA AGCGCATGAA ATTGTTGTGA AAAATACAAA CGAATTAGCA 7200
 GATCGAATTG AACGTGTTGT TCCTATTAAA GATGAATTAT ACACACCGCG TATGGAAGGT 7260

55

	CTGCCTCAAA TCGTAATTGA TCGATTAGAA AAAGAATTAA AAAGTATTAT CGGTAATGGA	7380
	TTTGCGGTAA TTTACTTAAT TTCGCAACGT TTAGTTAAAA AATCATTAGA TGATGGATAC	7440
5	TTAGTTGGTT CCCGTGGTTC AGTAGGTTCT AGTTTTGTAG CGACAATGAC TGAGATTACT	7500
	GAAGTAAACC CGTTACCGCC AACTATATT TGTCCGAACGT GTAAAACGAG TGAATTTTTC	7560
10	AATGATGGTT CAGTAGGATC AGGATTTGAT TTACCTGATA AGACGTGTGA AACTTGTGGA	7620
	GCGCCACTTA TTAAAGAAGG ACAAGATATT CCGTTTGAAA CATTTTTAGG ATTTAAGGGA	7680
	GATAAAGTTC CTGATATCGA CTTAAACTTT AGTGGTGAAT ATCAACCGAA TGCCCATAAC	7740
15	TACACAAAAG TATTATTTGG TGAGGATAAA GTATTCCGTG CAGGTACAAT TGGTACTGTT	7800
	GCTGAAAAGA CTGCTTTTGG TTATGTTAAA GGTTATTTGA ATGATCAAGG TATCCACAAA	7860
	AGAGGTGCTG AAATAGATCG ACTCGTTAAA GGATGTACAG GTGTAAACG TACAACGGA	7920
20	CAGCATCCAG GGGGTATTAT TGTAGTACCT GATTACATGG ATATTTATGA TTTTACGCCG	7980
	ATACAATATC CTGCCGATGA TCAAAATTCA GCATGGATGA CGACACATTT TGATTTCCAT	8040
	TCTATTCATG ATAATGTATT AAAACTTGAT ATACTTGGAC ACGATGATCC AACAATGATT	8100
25	CGTATGCTTC AAGATTTATC AGGAATTGAT CCAAAAACAA TACCTGTAGA TGATAAAGAA	8160
	GTTATGCAGA TATTTAGTAC ACCTGAAAGT TTGGGTGTTA CTGAAGATGA AATTTTATGT	8220
	AAAACAGGTA CATTTGGGGT ACCAGAATTC GGTACAGGAT TCGTGCGTCA AATGTTAGAA	8280
30	GATACAAAGC CAACAACATT TTCTGAATTA GTTCAATCT CAGGATTATC TCATGGTACA	8340
	GATGTGTGGT TAGGCAATGC TCAAGAATTA ATTAACCCG GTATATGTGA TTTATCAAGT	8400
	GTAATTGGTT GTCGTGATGA TATCATGGTT TATTTAATGT ATGCTGGTTT AGAACCATCA	8460
35	ATGGCTTTTA AAATAATGGA GTCAGTACGT AAAGGTAAAG GTTTAACTGA AGAAATGATT	8520
	GAAAAGATGA AAGAAAATGA AGTGCCAGAT TGGTATTTAG ATTCATGTCT TAAAATTAAG	8580
40	TACATGTTCC CTAAAGCCCA TGCAGCAGCA TACGTTTTAA TGGCAGTACG TATCGCATAT	8640
	TTCAAAGTAC ATCATCCACT TTATTACTAT GCATCTTACT TTACAATTCG TGCGTCAGAC	8700
	TTTGATTAA TCACGATGAT TAAAGATAAA ACAAGCATTC GAAATACTGT AAAAGACATG	8760
45	TATTCTCGCT ATATGGATCT AGGTAAAAA GAAAAAGACG TATTAACAGT CTTGGAAATT	8820
	ATGAATGAAA TGGCGCATCG AGGTTATCGA ATGCAACCGA TTAGTTTAGA AAAGAGTCAG	8880
	GCGTTCGAAT TTATCATTGA AGGCGATACA CTTATTCCGC CGTTCATATC AGTGCCTGGG	8940
50	CTTGCGGAAA ACGTTGCGAA ACGAATTGTT GAAGCTCGTG ACGATGGCCC ATTTTATCA	9000
	AAAGAAGATT TAAACAAAA AGCTGGATTA TCTCAGAAAA TTATTGAGTA TTTAGATGAG	9060

55

	GAAATAATCA AGGTATTTAT TTAATGCGTA TGGCGTAGTC AAAGAAATAC AAAATTGTTG	9180
	CTGGACACAA AATTATGCCC GTATTTCTTT TCAATGTCTT ACGAGTCTAT TCAAATGTAA	9240
5	TGGTGAAATA AAGGAACAAA CTTTACAAAG AATCTCTGAT TAATAGTGAA GTCATTGTGTT	9300
	TCAAGCATAA ACTTATGCTA TAATTAAGTT GCTTAAAAAT TAGTGAACTC AGGCAGAAGA	9360
	GTGGGAGATT CCCGCTCTTT TCTATTTGCC AAAAAGGGAG GCCTGTATGA GTAAAATTAC	9420
10	AGAACAAGTA GAAGTGATTG TTAAACCAAT TATGGAAGAC TTGAATTTTG AACTTGTAAG	9480
	CGTTGAATAT GTCAAAGAGG GTAGAGATCA TTTTCTTAGA ATCTCTATTG ATAAAGAAGG	9540
	TGGCGTAGAT TTAAATGATT GTACGCTAGC TTCTGAAAAA ATAAGTGAAG CTATGGATGC	9600
15	AAATGATCCT ATTCTTGAAA TGTATTATTT AGACGTAGCG TCACCTGGTG CAGAACGTCC	9660
	AATTAATAAAA GAACAAGATT TCCAAAATGC AATAACTAAA CCTGTATTTG TTTCTTTATA	9720
20	TGTACCAATT GAAGGTGAAA AGGAATGGTT AGGCATTTTA CAAGAAGTCA ATAATGAAAC	9780
	AATTGTAGTA CAAGTTAAAA TCAAAGCAAG AACGAAAGAT ATAGAGATAC CGAGAGACAA	9840
	AATAGCAAAA GCACGTCACG CAGTTATGAT TTAACGTGAT GAGGAGGAAA AAACGTGTCA	9900
25	AGTAATGAAT TATTATTAGC TACTGAGTAT TTAGAAAAAG AAAAGAAGAT TCCTAGAGCA	9960
	GTATTAATTG ATGCTATTGA AGCAGCTTTA ATTACTGCAT ACAAAAAGAA TTATGATAGT	10020
	GCAAGAAATG TCCGTGTGGA ATTAAATATG GATCAAGGTA CTTTCAAAGT TATCGCTCGT	10080
30	AAAGATGTTG TTGAAGAAGT ATTTGACGAC AGAGATGAAG TGGATTTAAG TACAGCGCTT	10140
	GTTAAAAACC CTGCATATGA AATTGGTGAT ATATACGAAG AAGATGTAAC ACCTAAAGAT	10200
	TTTGGTCGTG TAGGTGCTCA AGCAGCGAAA CAAGCAGTAA TGCAACGTCT TCGTGATGCT	10260
35	GAACGTGAAA TTTTATTTGA AGAATTTATA GACAAAGAAG AAGACATACT TACTGGAATT	10320
	ATTGACCGTG TTGACCATCG TTATGTATAT GTGAATTTAG GTCGTATCGA AGCTGTTTTA	10380
	TCTGAAGCAG AAAGAAGTCC TAACGAAAAA TATATTCCTA ACGAACGTAT CAAAGTATAT	10440
40	GTTAACAAAG TGAACAAAC GACAAAAGGT CCTCAAATCT ATGTTTCTCG TAGCCATCCA	10500
	GGTTTATTAA AACGTTTATT TGAACAAGAA GTTCCAGAAA TTTACGATGG TACTGTAATT	10560
	GTTAAATCAG TAGCACGTGA AGCTGGCGAT CGCTCTAAAA TTAGTGCTTT CTCTGAAAAC	10620
45	AATGATATAG ATGCTGTTGG TGCATGTGTT GGTGCTAAAG GCGCACGTGT TGAAGCTGTT	10680
	GTTGAAGAGC TAGGTGGTGA AAAAATCGAC ATCGTTCAAT GGAATGAAGA TCCAAAAGTA	10740
50	TTTGTAAGAA ATGCTTTAAG CCCTTCTCAA GTTTTAGAAG TTATTGTTGA TGAACAAAT	10800
	CAATCTACAG TAGTTGTTGT TCCTGATTAT CAATTGTCAT TAGCGATTGG TAAAAGAGGA	10860

55

	GATGCGCGTG AAGCGGGTAT CTATCCAGTA GTTGAAGCTG AAAAAGTAAC TGAAGAAGAT	10980
	GTTCCTTTAG AAGATGCTGA CACAACAGAA TCAACCGAAG AGGTAAATGA TGTTTCAGTT	11040
5	GAAACAAATG TAGAGAAAGA ATCTGAATAA TAGGTTGGAG TGAAGTATCT ATGAAAAAGA	11100
	AAAAAATTCC GATGCGAAAA TGTATTCTTT CAAATGAAAT GCATCCCAA AAAGATATGA	11160
	TTCTGTTTGT TGTTAATAAA GAAGGCGAAA TCTTTGCGGA TGTTACTGGA AAGAAACAAG	11220
10	GCCGTGGCGC ATATGTTTCT AAAGATGTTG CTATGGTTGA AAAAGCACAA CAAAAAGAAA	11280
	TTTTAGAAAA ATATTTTAAA GCATCTAAAG AGCAATTGGA TCCTGTTTAC AAAGAAATTA	11340
	TTAGATTAAT TTATAGAGAA GAGATCCCAA AATGAGTATA GATCAAATAT TAAACTTTTT	11400
15	AGGATTAGCA ATGAGAGCTG GTAAAGTAAA AACAGGTGAA TCAGTCATTG TTAATGAGAT	11460
	TAAAAAAGGA AATTTGAAGC TCGTTATTGT TGCAAATGAT GCGTCTGATA ATACAGCTAA	11520
20	ATTAATTACA GATAAATGTA AGAGTTACAA AGTTCCATTG AGAAAGTTTG GAAATCGAAA	11580
	TGAATTGGGA ATAGCACTTG GAAAAGGTGA GCGTGTTAAT GTAGGGATTA CTGACCCAGG	11640
	CTTTGCTAAA AAGTTGCTAT CAATGATAGA TGAATATCAT AAGGAGTGAT TATATGAGTA	11700
25	AACAAAGAAT TTACGAATAT GCGAAAGAAT TAAATCTAAA GAGTAAAGAG ATTATAGATG	11760
	AGTTAAAAAG CATGAATATT GAGGTTTCAA ATCATATGCA AGCTTTGGAA GATGACCAA	11820
	TTAAAGCATT AGATAAAAAG TTCAAAAAG AACAAAAGAA CGACAATAAA CAAAGCACTC	11880
30	AAAATAATCA CCAAAAATCA AACAAATCAA ACCAAAATAA AGGGCmACAA AAAGATAACA	11940
	AAAAGAATCm ACAACAAAAT AATAAAGGCA ACAAAGGCAA TAAAAAGAAT AATAGAAATa	12000
	ATAAGAAAAA TAACAAGAAT AATAAACCAC AAAATCAACC AGCTGCTCCA AAAGAAATAC	12060
35	CATCAAAAGT GACATATCAA GAAGGTATTA CAGTAGGCGA ATTTGCGGAT AAATTAAATG	12120
	TTGAATCATC AGAAATTATC AAAAAATTAT TCTTACTTGG TATTGTTGCT AATATCAATC	12180
40	AATCATTAAA TCAAGAAACA ATCGAATTAA TTGCCGATGA TTATGGCGTT GAGGTTGAAG	12240
	AAGAAGTTGT GATTAATGAA GAAGACTTAT CAATCTATTT CGAAGACGAA AAAGATGATC	12300
	CAGAGGCAAT TGAGAGACCA GCAGTTGTAA CAATTATGGG ACATGTTGAC CATGGTAAAA	12360
45	CGACTTTATT AGATTCAATT CGTCATACAA AAGTTACAGC AGGTGAAGCA GGCGGAATCA	12420
	CTCAACATAT TGGTGATAT CAAATTGAAA ACGATGGCAA AAAAATCACT TTCTTAGATA	12480
	CACCGGGACA TGCTGCATTT ACAACGATGC GTGCGCGTGG TGCaCAAGTA ACAGATATTA	12540
50	CTATTTTAGT AGTAGCAGCT GACGATGGTG TTATGCCACA AACAAATTGAA GCAATTAACC	12600
	ATGCTAAAGA AGCagAAGTA CCAATTATTG TTGCAGTAAA TAAAAITGAT AAACCAACTT	12660

55

	GCGGCGAAAC AATTtTCGTC CACTTTCTGC ATTAAGTGGT GATGGTATCG ACGATTTATT	12780
	AGAAATGATA GGATTAGTTG CAGAAGTTCA AGAACTTAAA GCAAATCCTA AAAACCGTGC	12840
5	TGTTGGTACA GTTATCGAAG CTGAATTAGA TAAATCACGT GGTCCCTCTG CATCATTATT	12900
	AGTACAAAAC GGTACATTAA ATGTTGGTGA TGCATTGTA GTTGGTAATA CTTACGGCCG	12960
10	TATCCGTGCA ATGGTTAATG ACTTAGGTCA AAGAATCAAA ACGGCTGGTC CATCAACGCC	13020
	TGTTGAAATT ACAGGTATTA ATGATGTGCC ACAAGCTGGG GATCGCTTTG TTGTATTTAG	13080
	TGATGAAAAA CAAGCTCGTC GTATTGGTGA ATCAAGACAC GAAGCTAGCA TTATACAACA	13140
15	ACGTCAAGAA AGTAAAAATG TTTCAATTAGA TAACCTGTTT GAACAAATGA AACAAGGTGA	13200
	AATGAAAGAT TTAAACGTTA TTATTAAAGG TGATGTTCAA GGTCTGTTG AAGCTTTAGC	13260
	TGCATCATTa ATGAAAATTG ATGTTGAAGG CGTAAATGTT CGTATCATTc ATACAGCGGT	13320
20	TGGTGCAATT AATGAGTCAG ACGTGACACT TGCTAATGCC TCAAATGGTA TTATCATTGG	13380
	TTTCAATGTT CGTCCAGACA GTGGTGCAA ACGTGCTGCA GAAGCTGAAA ATGTTGATAT	13440
	GCGTTTACAC AGAGTTATTT ATAATGTTAT CGAAGAAATT GAATCAGCGA TGAAAGGTTT	13500
25	ACTTGATCCA GAATTTGAAG AACAAGTTAT CGGACAAGCT GAAGTTCGTC AAACATTCAA	13560
	AGTTTCTAAA GTTGGTACTA TTGCTGGATG TTATGTTACT GAAGGTAAAA TTACGCGAAA	13620
	TGCTGGTGTA CGTATTATTC GTGATGGTAT TGTTCAATAT GAAGGCGAAT TAGATACACT	13680
30	TAAACGTTTC AAAGATGATG CTAAGGAAGT TGCAAAAGGT TATGAATGTG GTATTACAAY	13740
	TGAAAACTAC AATGACCTTA AAGAAGGCCA TGTTATCGAA GCATTTGAAA TGGTTGAAAT	13800
	TAAGCGTTAA TTAAATAAAT TACAAGCTAA AAGTATAGTT AAGATTGATA TGCTCCCTAT	13860
35	AAATATTGCA CTTTTTAAGT GTCTACTTTA TAGGGAGCAT ATTTGATACT AGCTTTTGGT	13920
	TTTPIATTAG AATAGATTAC CTATTAAAAG TTACGTTATA TGGACATGAT TTTGTATAAA	13980
40	ATTTTGTGGT GGCCTAGAAT GATTTTAAAT GACAAAATAT AATGTCGACT ATTATTGGAA	14040
	AATTTTCTGT TGaAATGCCT ATCTTACGGC AAACCTTTATT TGATTTTATA GGCTTAATTT	14100
	ATTAAAATAA CGTGTGAGCT AAAATAATTG TTTAAGCATT GTTACACTAA AAAATGCAAA	14160
45	TAACAATTGA ACTTAAAGAT AAAGAGGTGA CAAGAATGAG CAGTATGAGA GCAGAGCGTG	14220
	TTGGTGAACA AATGAAGAAG GAATTAATGG ATATCATCAA CAATAAAGTC AAAGATCCTC	14280
	GAGTTGGTTT TATTACAATT ACAGATGTTG TTTTAACAAA TGATTTATCG CAGGCTAAAG	14340
50	TATTTTTAAC TGTATTAGGT AACGATAAAG AAGTAGAAAA TACATTTTAA GCACCTTGATA	14400
	AAGCAAAAGG CTTCAATTAAG TCTGAATTAG GTTCTAGAAT GCGATTACGT ATTATGCCGG	14460

55

	AAGATTIACA CAAACAAGAT AGATAATTTA GTGTTAGGTA TCTGGAAAAT GTTTGATAAT	14580
	TTCTTAATAT CGGTATATTA ACATTAAACA GTTAATACAT AGATGTGTAG AAATAGTTAA	14640
5	CATTTTCCAG TTTTTTTATG AATAAATTTA GTTGATACGC TATTAAAATA TATTTTAAAA	14700
	AAGAAGGTGA CTATATGTAT AATGGGATAT TACCAGTATA TAAAGAGCGC GGTTTAACAA	14760
	GTCA TGACGT TGTATTCAAA TTGCGTAAAA TATTAAAAAC TAAAAAATA GGTCAACACGG	14820
10	GTACGCTTGA TCCCGAAGTT GCAGGCGTGT TACCGGTATG TATAGGTAAT GCAACGAGAG	14880
	TTAGTGATTA TGTATGGAT ATGGGCAAAG CTTATGAAGC AACTGTATCG ATAGGAAGAA	14940
	GTACAACGAC TGAAGATCAA ACGGGTGATA CATTGGAAAC AAAAGGTGTA CACTCAGCAG	15000
15	ATTTTAATAA GGACGATATT GACCGATTGT TAGAAAGTTT TAAAGGTATC ATTGAACAAA	15060
	TTCCGCCGAT GTACTCATCC GTCAAAGTAA ATGGTAAAAA ATTATATGAA TATGCGCGTA	15120
	ATAATGAAAC AGTTGAAAGA CCAAAGCGTA AAGTAAATAT TAAAGACATT GGGCGTATAT	15180
20	CTGAATTAGA TTTTAAAGAA AATGAGTGTC ATTTTAAAT ACGCGTCATC TGTGGTAAAG	15240
	GTACATATAT TAGAACGCTA GCAACTGATA TTGGTGTGAA ATTAGGCTTT CCGGCACATA	15300
25	TGTGCGAAAT AACACGAATC GAGTCTGGTG GATTTGTGTT GAAAGATAGC CTTACATTAG	15360
	AACAAATAAA AGAACTTCAT GAGCAGGATT CATTGCAAAA TAAATTGTTT CCTTTAGAAAT	15420
	ATGGATTAAA GGGTTTGCCA AGCATTAAAA TTAAAGATTC GCACATAAAA AAACGTATTT	15480
30	TAAATGGGCA GAAATTTAAT AAAAATGAAT TTGATAACAA AATTAAAGAC CAAATTGTAT	15540
	TTATTGATGA TGATTCAGAA AAAGTATTAG CAATTTATAT GGTACACCCT ACAAAGAAT	15600
	CAGAAATTAA ACCTAAAAAA GTCTTTAATT AAAGGAGATA GAATTTATGA AAGTCATAGA	15660
35	AGTGACACAT CCTATACAAT CTAAACAGTA TATTACAGAG GATGTTGCAA TGGCATTCCG	15720
	ATTTTTCGAT GGCATGCATA AAGGTCATGA CAAAGTCTTT GATATATTAA ACGAAATAGC	15780
	TGAGGCACGC AGTTTAAAAA AAGCGGTGAT GACATTTGAT CCGCATCCGT CTGTCGTGTT	15840
40	GAATCCTAAA AGAAAACGAA CAACGTATTT AACGCCACTT TCAGATAAAA TCGAAAAAAT	15900
	TAGCCAACAT GATATTGATT ATTGTATAGT GGTAAATTTT TCATCTAGGT TTGCTAATGT	15960
	GAGCGTAGAA GATTTTGTG AAAATTATAT AATTAAAAAT AATGTAAAAG AAGTCATTGC	16020
45	TGGTTTTGAT TTTACTTTTG GTAAATTTGG AAAAGGTAAT ATGACTGTAC TTCAAGAATA	16080
	TGATGCGTTT AATACGACAA TTGTGAGTAA ACAAGAAATT GAAAATGAAA AAATTTCTAC	16140
50	AACTTCTATT CGTCAAGATT TAATCAATGG TGAGTTGCAA AAAGCGAATG ATGCTTTAGG	16200
	CTATATATAT TCTATTAAAG KCACTGTAGT GCAAGGTGAA AAAAGGGGAA GAACATTG	16260

55

TGCTGTTAGT ATTGAAATCG GCACTGAAAA TAAATTATAT CGAGGGGTAG CTAACATAGG 16380
 TGTAAGCCA ACATTTTCATG ATCCTAACAA AGCAGAAGTT GTCATCGAAG TGAATATCTT 16440
 5 TGACTTTGAG GATAATATTT ATGGTGAACG AGTGACCGTG AATTGGCATC ATTTCTTACG 16500
 TCCTGAGATT AAATTTGATG GTATCGACCC ATTAGTTAAA CAAATGAACG ATGATAAATC 16560
 10 GCGTGCTAAA TATTTATTAG CAGTTGATTT TGGTGATGAA GTAGCTTATA ATATCTAGAG 16620
 TTGCGTATAG tTATATAAAC AATCTATACC ACACCTTTTT CTTAGTAGGT CGAATCTCCA 16680
 ACGCCTAACT CGGATTAAGG AGTATTCAAA CATTTTAAGG AGGAAATTGA TTATGGCAAT 16740
 15 TTCACAAGAA CGTAAAAACG AAATCATTAA AGAATACCGT GTACACGAAA CTGATACTGG 16800
 TTCACCAGAA GTACAAATCG CTGTACTTAC TGCAGAAATC AACGCaGTAA ACGAACACTT 16860
 ACGTACACAC AAAAAAGACC ACCATTCACG TCGTGGATTA TTAAAAATGG TAGGTCGTCG 16920
 20 TAGcATTTaT TAAACTACTT ACGTaGTAAA GATATTCAAC GTTACCGTGA ATTAATTAAA 16980
 TCACTTGGTA TCCGTCGTTA ATCTTAATAT AACGTCTTTG AGGTTGGGGC ATATTTATGT 17040
 TCCAACCTTA ATTTATATTA AAAAGCTTTT TTCAAATAT TAACATTTAT TATATGTTAA 17100
 25 GCTAATATTG AGTGAATAAT AAGTTACAA TGAGATAAAG ATGATATAAG TACACCTAGA 17160
 GTAATAATCA AGATATTAAA AATAAAGTAT GTTTTTTTAA AAAATATAAC TTATATTTAT 17220
 ACTGATAAGG GTGGGACGAT AAGTCTATTT TGTAATAAAT AGATGGATAT CCCGCTCTCT 17280
 30 TTTTTTCCAA TTCAATATTT TATAACTAAT ATTAAATAC GATAATAAAT GATATGATAT 17340
 AACTATTAGA TTCAAGAGAG GAGATTTATA ATGTCTCAAG AAAAGAAAGT TTTTAAAACT 17400
 GAATGGGCAG GAAGATCTTT AACGATTGAA ACAGGGCAAT TAGCTAAACA AGCAAATGGC 17460
 35 GCTGTATTGG TTCGTTATGG AGATACAGTC GTGTTATCGA CGGCAACTGC ATCAAAAGAA 17520
 CCTCGTGATG GAGATTTCTT CCCATTAACA GTGAACTATG AAGAAAAAAT GTACGCTGCG 17580
 GGTAAAATTC CTGGTGGATT TAAAAAGAGA GAAGGACGTC CTGGTGACGA TGCAACATTA 17640
 40 ACTGCGCGAT TAATTGATAG ACCAATTAGA CCTTTATTCC CTAAAGGATA TAAGCATGAT 17700
 GTTCAAATTA TGAACATGGT ATTAAGTGCA GATCCTGATT GTTCACCACA AATGGCTGCA 17760
 45 ATGATTGGTT CATCTATGGC GCTTAGTGTG TCGGATATTC CATTCCAAGG GCCAATCGCC 17820
 GGTGTAAATG TGGTTATAT TGACGGTAAA TATATCATT ACCCAACAGT AGAAGAAAAA 17880
 GAAGTTTCTC GTTTAGACCT TGAAGTAGCT GGTCAATAAG ATGCGGTAAA CATGGTAGAG 17940
 50 GCAGGCGCTA GTGAGATTAC TGAACAAGAA ATGTTAGAGG CGATTTTCTT TGGTCATGAA 18000
 GAGATTCAAC GTTTAGTTGA TTTCCAACAA CAAATCGTCG ACCACATTCA ACCTGTTAAA 18060

	GAAGAAAAAG GACTTAAAGA AACAGTTTTA ACATTTGATA AACACAACG AGATGAAAAT	18180
	CTTGATAACT TAAAAGAAGA AATCGTCAAT GAATTTATCG ATGAAGAAGA TCCAGAGAAT	18240
5	GAACTTACTTA TTAAAGAAGT TTATGCAATT TTAAATGAAT TAGTGAAAGA AGAAGTTCGA	18300
	CGTTTAATTG CAGATGAAAA AATTAGACCA GACGGCCGTA AACCTGATGA AATCCGTCCA	18360
	TTAGATTCTG AAGTTGGTAT TTTACCTAGA ACGCATGGTT CAGGTCTATT TACACGTGGT	18420
10	CAGACTCAAG CACTTTTCAGT TTTAACATTA GGTGCTTTAG GCGATTATCA ATTAATTGAT	18480
	GGTTTAGGAC CTGAAGAAGA AAAAAGATTC ATGCATCATT ACAACTTCCC GAATTTTTCA	18540
15	GTAGGTGAAA CTGGTCCAGT ACGTGCGCCA GGTGCTCGTG AAATTGGACA TGGTGCCTTA	18600
	GGTGAAAGAG CATTAAAATA TATTATTCCT GATACTGCTG ATTTCCCATTA TACAATTCGT	18660
	ATTGTAAGTG AGGTACTTGA ATCAAATGGT TCATCATCTC AAGCGTCAAT TTGTGGATCA	18720
20	ACATTAGCAT TAATGGATGC GGGCGTACCG ATTAAAGCAC CAGTTGCTGG TATTGCTATG	18780
	GGCCTTGTTA CACGTGAAGA TAGCTATACG ATTTTAACTG ATATCCAAGG TATGGAAGAT	18840
	GCATTAGGTG ATATGGACTT TAAAGTCGCT GGTACTAAAG AAGGTATTAC AGCAATCCAA	18900
25	ATGGATATTA AAATTGACGG TTTAACGCGT GAAATTATCG AAGAGGCTCT AGAACAAGCG	18960
	AGACGTGGTC GTTTAGAAAT AATGAATCAT ATGTTACAAA CAATTGATCA ACCACGTACT	19020
30	GAATTAAGTG CTTACGCGCC AAAAGTTGTA ACTATGACAA TTAAACCAGA TAAGATTAGA	19080
	GATGTTATCG GACCTGGTGG TAAAAAATT AACGAAATTA TTGATGAAAC AGGTGTTAAA	19140
	TTAGATATTG AACAAGATGG TACTATCTTT ATTGGTGCTG TTGATCAAGC TATGATAAAT	19200
35	CGTGCTCGTG AAATCATTGA GGAAATTACA CGTGAAGCGG AAGTAGGTCA AACTTATCAA	19260
	GCCACTGTTA AACGTATTGA AAAATACGGT GCGTTTGTAG GCCTATTCCC AGGTAAAGAT	19320
	GCGTTGCTTC ACATTTTACA AATTTCAAAA AATAGAATTG AAAAAGTGGA AGATGTATTA	19380
40	AAAATCGGTG ACACAATTGA AGTTAAGATT ACTGAAATTG ATAAACAAGG TCGAGTAAAT	19440
	GCTTCACATA GAGCATTAGA AGAATAATAT TTAAAGTCAT ATGACGACAA TGTATCGTCA	19500
	TGTGATTTT TTATGCCACT TTTTACGAAG TGACCCGTTT TGAATTTGTT GTATTGAACA	19560
45	TTTTAAAACG CTTTATTATT TTGTGTGCAA CTGTTAATTA TCCTGTATGT ATAGTGATT	19620
	ATAGTGATCA TCAAGTGTTT TTTAACTTAT AATGAATAGT GAGTTTATAT ATGGACGGGT	19680
50	AACAAATTTA GGAGGTAAGA TTTTGAGTTT AATAAAGAAA AAGAATAAAG ATATTCGCAT	19740
	TATACCATTA GCGGGTGTG GCGAAATTGC TAAAAATATG TATATCGTTG AAGTAGACGA	19800
	TGAAATGTTT ATGTTAGATG CTGGACTTAT GTTCCAGAA GACGAAATGC TAGGTATTGA	19860
55		

	CCTTACACAC GGACATGAGC ACGCGATTGG TGCAGTGAGT TATGTTTTAG AACAAATTAGA	19980
	TGCACCAGTA TATGGATCTA AATTGACAAT AGCGTTAATT AAAGAAAATA TGAAAGCCCCG	20040
5	TAATATTGAT AAAAAAGTTC GCTACTATAC AGTTAATAAT GATTCAATTA TGAGATTCAA	20100
	AAACGTGAAT ATTAGTTTCT TTAATACGAC ACACAGTATT CCTGATAGTT TAGGTGTTTG	20160
10	TATTCACACT TCATATGGTG CCATTGTGTA TACAGGTGAA TTTAAGTTTG ACCAAAGTTT	20220
	ACATGGACAT TATGCACCAG ATATTAAACG TATGGCAGAG ATTGGTGAAG AAGGCGTATT	20280
	TGTCTTAATC AGTGATTCTA CTGAGGCAGA GAAACCTGGA TATAATACTC CGGAAAATGT	20340
15	GATTGAACAT CATATGTATG ATGCTTTTGC AAAAGTCCGA GGTGCTTGA TAGTTTCATG	20400
	TTATGCTTCG AACTTTATAC GTATTGAGCA AGTTTTAAAT ATTGCTAGCA AGCTAAATCG	20460
	TAAAGTGTCA TTTTtagGAA GATCACTTGA AAGTTCATT AATATTGCTC GTAAAATGGG	20520
20	GTATTTGAC ATTCTAAAG ATTTGCTAAT TCCTATAACA GAAGTTGATA ATTATCTAA	20580
	AAATGAAGTG ATAATTATAG CTACTGGTAT GCAAGGAGAA CCTGTAGAAG CCTTAAGTCA	20640
	AATGGCGCAA CATAAGCATA AAATTATGAA TATCGAAGAA GGCGATTCTG TATTTTTAGC	20700
25	AATTACGGCT TCTGCTAATA TGGAAGTTAT CATTGCGAAT AcATTAAATG AGCtTgTtAC	20760
	GnCTGGCGCA CATATTATTC CAAATAACAA AAAGATTCAT GCTTCAAGTC ATGGTTGCAT	20820
30	GGAAGAATTA AAAATGATGA TTAATATTAT GAAACCTGAA TACTTTATTC CTGTACAAGG	20880
	TGAATTTAAA ATGCAGATAG CACATGCGAA GCTAGCAGCT GAAGCAGGTG TTGCACCAGA	20940
	AAAGATTTTC CTGTGGAAA AAGGAGATGT CATTAAATTAC AACGGTAAAG ATATGATATT	21000
35	AAATGAAAAG GTAAATTCAG GAAATATTTT AATAGATGGC ATTGGTATTG GGGATGTAGG	21060
	AAATATCGTG TTGAGAGACC GTCATCTTTT AGCAGAAGAT GGTATCTTTA TTGCTGTTGT	21120
	AACGTTAGAT CCTAAAAATA GACGTATAGC TGCGGGACCT GAAATTCAAT CTCGTGGGTT	21180
40	TGTATATGTA CGTGAAAGTG AAGACTTATT ACGTGAAGCA GAAGAGAAAAG TACGTGAAAT	21240
	AGTAGAGGCT GGTTTACAAG AAAAACGCAT AGAATGGTCT GAAATTAAAC AAAATATGCG	21300
	TGATCAAATT AGTAAACTAT TATTCGAAAG TACAAAACGT CGTCCTATGA TTATTCCAGT	21360
45	AATTTCTGAA ATTTAATCAA AAAGTCATTA ACATAAAAGA GGTCAGAACA AGTCACTGAA	21420
	ATATAATGGT TGTGATGGAC AATTTACTTA TATTTTATGA TAGTCAATTG AAGGGGTAAC	21480
50	GATTAATCTG TTATCTTAAG TAAATTGATA CATAGATGAT ATTGTTCTAA CCTCTTTCAT	21540
	CGTCTGTTTG GACTACATAT TCTAAACATC AAATAGGAAA TTATATATAA TAACGTCGTT	21600
55	TTAACTAAGG CAACATAAGG AGGTGCGTCA ATTGGCACAA GCAAAAAAGA AATCGACAGC	21660

GATACGTTAT GTCATAGCTA TTTTAGTAGT TGTATTAATG GTGTTGGGTG TTTTCCAATT 21780
 AGGAATAATA GGTGCTCTAA TTGACAGCTT CTTTAATTAT TTATTGGGT ACAGTAGATA 21840
 5 TTTAACATAT ATTTTAGTAC TCTTAGCAAC TGGTTTTATT ACATACTCTA AACGTATTCC 21900
 TAmAaCTAGA CGAACGGCTG GTTCGATTGT ATTGCAAATT GCATTGCTAT TTGTATCACA 21960
 GTTAGTTTTT CATTTTAATA GTGGTATCAA AGCTGAAAGA GAACCTGTAC TTTCTTATGT 22020
 10 GTATCAGTCA TACCAACACA GTCATTTCCC AAATTTTGGT GGCGGTGTAT TAGGCTTTTA 22080
 TTTATTAGAG TTAAGCGTAC CTTTAATTTT ATTATTTGGT GTATGTATTA TTAATTTTT 22140
 15 ATTATTATGC TCAAGTGTTA TTTTATTAAC AAACCATCAA CATCGTGAAG TTGCAAAAGT 22200
 TGCACTGGAA AATATAAAG CTTGGTTTGG TTCATTTAAT GAA 22243

(2) INFORMATION FOR SEQ ID NO: 165:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5510 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

TTATTAATna TTAATATTTT TATTTTAAa AATAAAGCGA GGAGCTATCA ATGGAACAAA 60
 30 TTACTTCTGC ACAAATAAT AGAATTAAAC AAGCGAACAA GCTAAAaAG AAACGTGAGA 120
 GGGATAAAC TGGATTAGCT TTAATTGAAG GTGTGCATTT AATTGAAGAA GCTTATCAAA 180
 35 GTGGAATTGT AATTACACAA TTATTGCAA TTGAACCGGC AAGATTAGAT CAGCAAATTA 240
 WCGCATACGC GCAAGAAGTT TTTgAAATAA ACATGAAAGT TGCTGAATCT TTATCAGGTA 300
 CAGTgACACC ACAAGGTTT TTCGCAATCA TTGAGAAGCC GCATTATGAT ATTTCTAAAG 360
 40 CACAACAAGT ATTGCTCATC GATCGTGTTT AAGATCCTGG AAATTTAGGC ACATTAATTA 420
 GAACTGCCGA TGCTGCTGGA ATGGATGCTG TAATAATGGA GAAGGGTACG ACAGATCCTT 480
 ATCAAGATAA AGTGTTGCGA GCGAGTCAAG GTAGTGTTTT CCATTTGCCA GTTATGACAC 540
 45 AAGATCTCGA TACGTTTATT ACTCAATTTA ATGGTCCTGT TTATGGTACA GCACTTGAAA 600
 ACGCAGTgGc ATACAAAGAA GTTACTTCAA GTGATTCTTT TGCATTACTA TTAGGTAATG 660
 AGGGAGAAGG TGTTAATCCT GAATTATTAG CACATACTAC ACAAATTTA ATCATACCTA 720
 50 TTTATGGTAA AGCTGAAAGT TTAAATGTAG CGATTGCAGG TAGTATTTTA CTTTATCATT 780
 TGAAAGGTTG ACCGTGTTGA AAGTTTTCCG ATATAATTAT AATTAATTGT TTAACAGAAC 840

55

	ATAAATAATT	GTTTTAGGGA	GAATAATCGT	GACTGCAAGT	TATTCCAATT	ATTTAAAGTC	960
	TTTTCACCTT	TTTGGTTACT	TAAAGAGATT	TAAGTCGGAA	AGACAATCCG	TTATCAATAT	1020
5	TAAACAAGTG	TATGCTTAGG	CATAAATTTG	GGTGGTACCA	CGGAAATGAC	TTTCGTCCCT	1080
	TATTTTTTAA	GAGGATGAAA	GTCTTTTTTT	AGTTAAACAA	CAAATATGAT	AAATAGAAAA	1140
	TGAATAGTTC	GAATAGGGAG	GTCAGTGACA	TATGTCTGAA	CAACAAACAA	TGTCAGAGTT	1200
10	AAAACAACAA	GCGCTTGTAG	ATATTAATGA	AGCAAATGAT	GAACGTGCAC	TGCAAGAAGT	1260
	TAAAGTGAAA	TACTTAGGTA	AAAAGGGGTC	AGTTAGCGGA	CTAATGAAAT	TGATGAAGGA	1320
15	TTTGCCGAAT	GAAGATAAAC	CTGCGTTTGG	TCAAAAAGTG	AATGAATTGC	GTCAAACAAT	1380
	TCAAATGAA	TTAGATGAAA	GACAACAGAT	GTTAGTTAAA	GAAAAATTAA	ATAAGCCAAT	1440
	TGGCTGAAGA	AACAATTGAT	GTATCATTAC	CAGGTCGTCA	TATTGAAATC	GGTTCAAAGC	1500
20	ATCCATTAAC	ACGTACAATA	GAAGAAATTG	AAGACTTATT	CTTAGGTTTA	GGTTATGAAA	1560
	TTGTGAATGG	ATATGAAGTT	GAACAAGATC	ATTATAACTT	CGAAATGCTG	AATTTACCTA	1620
	AATCACACCC	TGCACGTGAT	ATGCAAGATA	GTTTCTATAT	TACGGATGAA	ATTTTATTAC	1680
25	GTACGCATAC	ATCACCAGTG	CAGGCACGTa	CGATGGAATC	ACGTCATGGT	CAAGGTCCAG	1740
	TTAAAATTAT	TTGCCCTGGT	AAAGTGTATC	GTCGTGACTC	TGATGATGCG	ACACATAGTC	1800
	ATCAATTTAC	ACAAATCGAA	GGATTAGTTG	TTGATAAAAA	CGTTAAAATG	AGTGATTTGA	1860
30	AAGGTACTTT	AGAATTGTTA	GCTAAGAAAT	TATTTGGTGC	TGATCGTGAA	ATTCGTTTAC	1920
	GTCCAAGTTA	CTTCCCATTG	ACTGAACCTT	CTGTAGAAGT	TGATGTGTCA	TGTTTTAAAT	1980
35	GTAAAGGAAA	AGGTTGTAAT	GTGTGTAAAC	ACACAGGATG	GATTGAAATT	TTAGGTGCTG	2040
	GAATGGTACA	TCCTAATGTA	TTAGAAATGG	CTGGTTTGA	TTCTTCAGAG	TACTCTGGAT	2100
	TTGCATTTGG	TATGGGACCA	GACCGTATTG	CAATGTTGAA	ATATGGTATA	GAAGATATTC	2160
40	GTCATTTCTA	TACTAATGAT	GTGAGATTTT	TAGATCAATT	TAAAGCGGTA	GAAGATAGAG	2220
	GTGACATGTA	ATGTTGATAT	CAAATGAATG	GTTGAAAGAA	TATGTAACAA	TCGATGATTC	2280
	TGTAAGTAAT	TTGGCAGAAC	GTATTACGCG	CACAGGTATT	GAAGTGGATG	ATTTAATTGA	2340
45	CTACACAAAA	GATATCAAAA	ATTTAGTTGT	CGGCTTCGTT	AAGTCAAAAG	AGAAACATCC	2400
	TGATGCTGAT	AAATTAAATG	TTTGCCAAGT	TGATATCGGA	GAAGACGAAC	CTGTACAAAT	2460
50	CGTTTGTGGT	GCACCGAACG	TTGaTGCAGG	ACAATATGTC	ATTGTTGCTA	AAGTAGGTGG	2520
	CAGATTGCCT	GGTGGTATTA	AAATTAAGCG	TGCCAAATTA	CGCGGTGAAC	GTTCAGAAGG	2580
55	TATGATTTGT	TCGTTACAAG	AAATTGGTAT	TTCAAGTAAC	TATATACCGA	AAAGTTTGA	2640

	ATATTTAGAT GATCAAGTAA TGGAATTTGA TTTAACGCGG AATCGTGCAG ATGCTTTAAG	2760
	TATGATAGGT ACTGCTTATG AAGTTGCAGC ATTATATAAT ACAAAAATGA CTAAGCCAGA	2820
5	GACAACATCA AATGAGCTTG ATTTATCTGC AAATGATGAA CTGACTGTGA CAATTGAAAA	2880
	TGAAGATAAA GTACCATATT ATAGTGCACG TGTGTTCAC GACGTGACAA TTGAACCTC	2940
	GCCAATTTGG ATGCAAGCAC GCTTAATAAA AGCGGGTATA CGTCCTATTA ATAATGTTGT	3000
10	TGACATTTCA AATTATGTGT TATTAGAATA CGGTCAACCA TTGCACATGT TTGATCAAGA	3060
	TGCGATTGGT TCACAACAAA TTGTGTTCG TCAAGCTAAT GAAGGCGAAA AAATGACAAC	3120
15	ATTAGATGAT ACAGAACGTG AATTATTAAC GAGCGATATT GTCATTACTA ATGGACAAAC	3180
	TCCAATTGCA TTAGCTGGTG TTATGGGTGG CGATTTTTCA GAAGTTAAAG AACAAACATC	3240
	AAATATAGTG ATTGAAGGTG CTATTTTTGA TCCAGTTTCA ATTCGTCATA CATCAAGACG	3300
20	TTTAAATTTA CGCAGTGAAT CATCTAGTCG TTTTGAAAAA GGAATAGCTA CTGAATTTGT	3360
	AGATGAAGCA GTCGACCGTG CATGTTATTT ATTACAAACT TATGCAAACG GAAAAGTGCT	3420
	AAAAGATAGA GTGTCTTCAG GAGAACTTGG TGCATTTATT ACACCAATCG ACATCACTGC	3480
25	TGATAAAATT AATCGCACTA TTGGATTTGA TTTGTCACAA AATGATATTG TTACTATTTT	3540
	TAATCAACTA GGGTTTGATA CAGAAATAAA TGATGATGTT ATTACAGTGC TAGTACCATC	3600
	ACGTCGTAAA GATATTACAA TTAAAGAAGA TTTAATTGAA GAAGTTGCAC GTATATATGG	3660
30	ATACGACGAT ATTCCATCAA CGTTACCTGT CTTGATAAA GTTACTAGTG GTCAGCTAAC	3720
	TGATCGCCAA TATAAACTA GAATGGTTAA AGAAGTGTTA GAAGGTGCTG GATTAGACCa	3780
35	AGCTATTACG TATTCGTTAG TTTCTAAAGA AGATGCTACT GCATTTTCGA TGCAACAGCG	3840
	TCAAACAATT GATTTATTGA TGCCAATGAG TGAAGCGCAT GCGTCATTAC GTCAAAGTTT	3900
	ATTACCACAT TTAATCGAAG CGGCATCATA TAATGTGGCA CGCAAAAATA AAGATGTAAA	3960
40	ATTATTTGAA ATCGGCAATG TCTTCTTTGC TAATGGAGAA GGTGAACTAC CAGATCAAGT	4020
	TGAATATTTA AGTGGTATTT TAACTGGAGA TTATGTAGTC AATCAATGGC AAGGTAAGAA	4080
	AGAAACGGTT GATTTCTATT TAGCAAAAGG TGTCGTGGAT CGAGTATCTG AAAAGTTAAA	4140
45	TCTTGAATTT AGTTATCGCC GTGCTGATAT TGaTGGATTA CATCCAGGTC GTACTGCTGA	4200
	AATCTTATTA GAGAATAAAG TTGTTGGTTT TATTGGTGAA TTACATCCAA TATTAGCAGC	4260
50	TGATAATGAT TTA AACGTA CGTATGTTTT TGAGTTGAAT TTTGATGCAT TAATGGCTGT	4320
	GTCGGTAGGT TACATTAATT ACCAGCCAAT TCCGAGATTC CCAGGCATGT CTCGTGACAT	4380
	TGCATTAGAA GTAGATCAAA ATATTCCAGC AGCTGATTTA TTATCAACGA TTCATGCACA	4440
55		

	AAAAGGTAAA AAATCAATTG CAATACGTTT AAATTATTTA GACACAGAAG AAACATTGAC	4560
	AGATGAGCGC GTTTCAAAAG TACAAGCGGA AATTGAAGCA GCATTAAATTG AACAAAGGTGC	4620
5	TGTTATTAGA TAATGATTTA AACCCCATGT ATAAGGATAT CTGAAGTAGA TTGATATCCC	4680
	TAACATGGGG TTTTATTTTT GGGTTCACCA ATTTGGTTCC AATGCATTTA AAAAGTCAAA	4740
	GAGGAACAGC GGAATACAGA TGATGcTTcG CACAACtGCA TAAAAGCCTC TAATGATTAA	4800
10	AAATCAAAGA GGCTTTAAAA TTTTTTGGGC TTTTTCACGA TTTTAAAAAT GCTTTTTTGA	4860
	AATGGTATCT AAACGTGAAA GACCGTATTT TTTTATAATT TTGGCGGCGA TTACATCGAC	4920
	TTTAGCACCG GCACCTTTAG GAATCGTCAT ATTAATATTT TTTGATATTT GATCCATATA	4980
15	TGTAACAAAT GCGTATCGAG AAATTATGCT TGCCACTGCA ATGGCTAATG ACTTCGATTC	5040
	TCCTTTTGTT TCAAATTTTG TTTCTTTGG AAGTGGTATA TCTGATAATG CGTAATGGCT	5100
20	ATACACTTCG CGTTTTGCGA ACTGATCAAT GACGATATAG TCTAATTGAG ACGAATCAAT	5160
	TTTTTCAAGT ACATTTTTGA TGGCTTCATT ATGAAGGGCA GCTTTCATTT TTAATTGAGT	5220
	CCAGCCTTTT GCTTGCTGAA TATTATATTT TTCATTGTGT AGTGTAAATA ATGAATGTGG	5280
25	TATGAAAGTA ACCAATTGCT CAGCAAGTTC TACAATTTTG GTATCGGTTA ATTTTTTTGA	5340
	ATCATCTACA CCCAAAGTTT TAAAAATAGG GACATGCTCT TTGGTAACGA AAGCAGCACA	5400
	CACAGTCAAC GGACCAAAGT AATCGCCACT TCCAGCCTCA TCACTACCAA TACAGTTAAA	5460
30	TTGrTCATAC ATTAaAGTTg TcCagAAAAG AATTAGCCAT ATTTnCCTTT	5510

(2) INFORMATION FOR SEQ ID NO: 166:

35	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 9623 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

	GnTTTAACTT ATAAATTTTA CGGGGGTAAT ATAATACTtA TTTACCTGTA ATATATGATA	60
45	ATTCTTCAGC GGCAGCTGCG TTGATAGTTC TATGAGAAAT GATACCTAAT CCTTTAACAT	120
	TGGATTCTGA AATAACGATA GAACCATCAC TGTTAACTTT TTCAACAAAT GCTACATGAC	180
50	CGTAATGTTG ATCTGCACCA AATTGTCCAG CCTCAAATAC AACAGCAGCA TGACGTTTTG	240
	GTGTATGACT TACTTGATAA TCACGGTATT GAGCTCGATT ATTCCAATTA TGTGCATCAC	300
55	CTAAATCACC TGAGATAGAT GTACCAAATT GTTTCATACG GTTATATACG TACCAAGTAC	360

	ATGAATCATC ATAATCCTTG ATAGAACGTT CATATTTATC TAAATCTGGC ATGCGTTCAT	480
	CGTCAAACG AGTTAATTGA TAGTGTTTAA TAATACTGTT TAATTTCTTA GCATAGTTTG	540
5	GATCTGTAGC ATATGTTTTA GATAAGTGTG ATGTTGCATC TTTATAAGAA TCGGCTTCCG	600
	ATTTCCATGT TGGTTTATAA ATTGTTTCGAT TGCCATCAAT ACCATTTTTA ATAAGGTCAG	660
	AGTAATCTTT TAGTGATTCT TTCGTGCTTG GATATTTTCG GAATCCAGCA TTAATACTAT	720
10	ACAATTGATT ACCATCAGCT TCTAATGTGT TAAAAGGAAC AGAATTCCTT TCaAAAGCAC	780
	CTTTGATACC GAATAAATTA TGGTTTGGTG ACWTAGCTAA AGCACTACGA CCTGAGTCAG	840
15	ATTCTAAGAT TGCTTGGGCA ATCATGACAG ACGCATAAAT ATCGTTATCT TGACCAATGC	900
	GATGTGCATC TTTAGCAATT GATTTGACAA ATTGACGTGT ATCTTTTGAG TCAACAACGT	960
	TAAATTGTCC GCTATCATCA TTGTTAGATA TACTAGGATC TGTTTCGAAT AATGATGTTG	1020
20	CACGTGTATC CTTTGTGATTA ACATCGTTAT TGAATGATTG AGCAGGTTTA GATTTATGTT	1080
	TCAATTCATC TTGTGTTGGT AACTGTGGAT TCTTTGTATT AGATTTTTCa TTTTGTCTT	1140
	TTTTAGATTG AGATGCATAA TCTTTTTGTG TTTTCTTTCG ATCTTCACTG TATTGATCCA	1200
25	AAATAGAGTC TAAAGCCGAA TCTGACATTG ATTGATTATC TTTTCGATGAA GATTTTTGAT	1260
	TTGCTTTATC GTCACCTGCT GGTGACTIAT TTGATTGATT AGGTTGTGTT GGCTTTGGCG	1320
	AATTTGGTTG CTTATTAGAT GTACTTGGTT TTGTATTGTT TGATTTAGGT GCTTTTTGAT	1380
30	TGTCTGCTTT ATCTTGTTTA GATGATTGCG TATCAGTGTC ATTTTGTATG CTATTGTCAC	1440
	TGTTTTTATT CGAATCATTT GTTGACTTTT CGCCATTACG AGGTTGTTCG TAATCAGAAA	1500
35	TATCCGAATT TAAATTGAAT AAGTTTTGGA TTAAAGTTGT TAATGAGTAA TTATCATCGT	1560
	ATTTATTTTT GGTTAGCAAT TGGTTTATAT TGGTTTGTGG TAAATTCTTA TAAATAAAAT	1620
	CAATGATATT GTTAGAGTCT GAAGTGCTGT CGTCTATAGT TTTAAATTTT TTGTCGTTAT	1680
40	TGTCTTGGTT ACTTGATTA TTTTGTCTG CTTTATCAAT ATCTTTACTT GTAGTATCCT	1740
	TAGAAGTTTC ATCGTCATTA GATTTTTTTG AATCATGAGA TGTGTCTTA GCTGTAGTAT	1800
	CTTTTTGAGG TGTATCAGCA TAAGCGgTAG GTGAaCTAA AGTAGGTAAT ACGAGCGTAG	1860
45	TTGATAGCAA ATAAATTAAA ATTTTATTTT TAGGCATATT TCGTATTCTC CCTTGAAAAA	1920
	TATAATAATT AAGTGTGATA ATAACTATG ATTTGTTATA ATTTATCGTA TGCTGAAAAT	1980
	AGTTGATAGG TATCAATCGA CTAAATATCT TCCAGTAAAT TGATTATACT AATTCACAAC	2040
50	GCAAAAATAA ATTAATTTAC AAAAAATATA TAAAAAATAT GAATAATTCC TACATAGGAG	2100
	TGTGACAATG AAGAACGCAT TTAAATTATT TAAAATGGAT CTGAAGAAAG TAGCTAAGAC	2160
55		

	TAACCTTATGG	GCAATGTGGG	ATCCATATGG	CAACACGGGA	CACATCAAGG	TCGCAGTCGT	2280
	TAATGAAGAT	AAAGGCGACA	CAATCAGAGG	GAAAAAAGTT	AATGTCGGTA	ATACGATGGT	2340
5	TAATACACTC	AAGAAAAATA	AAAGTTTTGA	TTGGCAGTTT	GTAAGTAGAG	AGAAAGCTGA	2400
	TCATGAGATA	AAAATGGGTA	AATATTTTGC	AGGTATTTAC	ATCCCATCTA	AGTTTACACA	2460
	TGAAATTACA	GGGACACTAC	GTAAGCAGCC	TCAAAAAGCA	GATGTAGAAT	TTAAGGTGAA	2520
10	TCAGAAGATT	AACGCTGTTG	CGTCTAAGCT	AACAGATACT	GGTTCGTCAG	TTGTCGTTGA	2580
	AAAAGCGAAT	GAACAATTTA	ATAAAACAGT	AACTCGAGCA	TTATTAGAAG	AAGCTAACAA	2640
15	AGCAGGTTTA	ACTATTGAAG	AAAATGTGCC	GACAAATTAAC	AAGATAAAAA	ATGCGGTATA	2700
	TTCAGCAGAT	AAAGCTTTAC	CTAAGATTAA	TGACTTTGCG	AATAAAATTG	TATATTTGAA	2760
	TAACCACCAA	GCGGATTTAG	ATAAATATGC	CAATGATTTT	AGAAAACTAG	GAAATTATAA	2820
20	AGGTGATATT	TTAGATGCTC	AGAAAAAATT	AAACGAAGTC	AATGGTGCTA	TTCCGCAACT	2880
	TAATGAAAAG	GCTAAGTTGA	TATTAGCTTT	AAATAATTAT	ATGCCGAAAA	TTGAAAAAGC	2940
	GTTAAATTTT	GCAGCTGATG	ACGTGCCAGC	GCAGTTCCCT	AAAATTAATC	AAGGACTTAA	3000
25	CATTGCGAGT	CAAGGTATTG	ATCAAGCTAA	TGGACAGTTA	AATGATGCCA	AAGGCTTCGT	3060
	CACACAAGTT	AGAAGTAGAG	TCGGTGATTA	TCAAGATGCA	ATTTCGACGG	CGCAAGATTT	3120
	AAATCGAAGA	AACCAGCAAC	AGATTCTCTA	AAATAGCGCG	GCGAACCAACG	AAACATCAAA	3180
30	TAGTGACCT	GCAGCTGGTA	ATGGTGTAGC	ATCAACGCCA	CCAAGTGCAC	CAAGTGGCGA	3240
	TACTGCACCA	AATAATAATG	TTACGCAAAA	TACCGCACCA	AATAGTAATA	ATGCGCCTGT	3300
35	ATCGACTACA	CCACAAAGTA	CAAGCGGGAA	AAAAGATGGT	CAAAGTTTTG	TAGATATAAC	3360
	AACAACACAA	GTCAGCACAG	CTAACGAGAA	CACACAAAC	ATTACAGATA	AAGATGTTAA	3420
	ATCAATGGAA	GCGGCATTAA	CGGGCTCTTT	ATTATCATT	TCAAATAATT	TAGATACCCA	3480
40	AGCGAAAGCC	GCACAAAAAG	ATAGTCAGGC	ATTACGTAAT	ATTCGTATG	GGATTTTAGC	3540
	ATCGGACAAG	CCTTCTGATT	TTAGAGAGTC	TTTAGATAAT	GTAAAGTCCG	GTTTAGAATA	3600
	CACAACGCAA	TATAATCAAC	AATTTATCGA	TACATTAAAA	GAGATTGAGA	AGAATGAAAA	3660
45	TGTTGATTTA	TCAAAAGAAA	TTGATAAGGT	AAAAGCAGCT	AATAATCGAA	TTAATGAATC	3720
	ATTAAGGTTA	GTAAATCAAT	TAAGCAATGC	ATTAAAGAAT	GGTAGTTCAG	GAACTGCTGA	3780
50	AGCTACTAAA	TTACTAGATC	AACTTTCAAA	ACTAGATTCA	TCATTATCAT	CATTTAGAGA	3840
	TTATGTTAAA	AAAGATCTTA	ACAGCTCTTT	AGTATCAATA	TCACAACGTA	TTATGGATGA	3900
	ATTGAACAAA	GGGCAAAC TG	CATTATCCAA	TGTTTCAGTCT	AAATTAAATA	CAATTGATCA	3960
55							

	AACAGTATTA CCAAGTATTG AACACAATA CATTAGTGCT GTTAAAAATG CTCAAGCAAA	4080
	CTTCTCGAAA GTGAAAAGTG ATGTAGCTAA AGCTGCTAAC TTTGTGCGCA ATGACTTACC	4140
5	ACAGTTAGAA CAGCGATTAA CTAATGCGAC AGCAAGTGTG AATAAAAAAT TACCAACGTT	4200
	ATTAAATGGT TATGATCAAG CGGTAGGATT ACTAAATAAA AATCAGCCAC AAGCGAAAAA	4260
10	GGCTTTATCA GATTTAGCTG ATTTTCTCA AAATAAATTG CCTGATGTTG AAAAAGATTT	4320
	GAAAAAGCG AATAAAATTT TCAAGAAATT AGACAAAGAT GATGCAGTCG ACAAATTAAT	4380
	CGACACACTT AAGAATGATT TGAAAAAGCA AGCGGGTATT ATTGCAAATC CTATTAATAA	4440
15	GAAGACTGTT GATGTTTTCC CAGTTAAGGA TTATGGTTCA GGTATGACAC CATTCTATAC	4500
	TGCACTGTCA GTATGGGTAG GTGCACTCTT GATGGTAAGT TTATTAACGG TTGATAATAA	4560
	ACATAAGAGT CTAGAGTCAG TCTTAACGAC AAGACAAGTG TTCTTAGGTA AGGCAGGATT	4620
20	CTTTATAATG CTTGGTATGT TGCAAGCACT CATTGTATCG GTTGGAGATT TGTTAATCCT	4680
	AAAAGCAGGA GTTGAGTCAC CTGTATTATT TGTACTTATA ACGATTTTCT GTTCGATTAT	4740
	TTTCAACTCA ATCGTATATA CGTGCGTATC ATTACTTGGT AACCAGGTA AAGCCATTGC	4800
25	AATCGTATTG CTTGTATTAC AAATTGCAGG TGGTGGGGGA ACATTCCCAA TTCAAACTAC	4860
	GCCACAATTT TTCCAAAACA TTTCGCCATA CTTACCATTT ACGTATGCAA TTGATTCATT	4920
30	ACGTGAAACA GTAGGCGGTA TTGTTCCGGA AATCCTAATT ACAAATTAA TTATATTAAC	4980
	GTTATTTGGT ATAGGATTCT TCGTTGTAGG TTTAATTTTA AAACCTGTAA CAGATCCATT	5040
	GATGAAGCGC GTATCTGAAA AAGTTGACCA AAGTAACGTT ACAGAATAAA AATTAAATCC	5100
35	ACACATTAGG GTTATAGCTC CTTAATGTGT GGATTTTAT GTTTTATAGAC AGAAGAGATA	5160
	GTAATTTCTG TCTTTTATGG GACGGTTGTT ATCATTGCTA TTATCCAGGA TGAATTACTA	5220
	TAGGACTAAT ATTACCGACA AAGTGAATAT CCTCGTCTTC CGTAGTTAAA ATAAAGCTAG	5280
40	AACCTTTTGT GATGTCATAG TGCTTATCGT TTAATGTTAA AGTACCAGTA CCATCGATAA	5340
	TTGTAATAA GCAATAAGCA TGTGGTTTAT TGAATTTTAA ATCTCCATGA ATATCCCAT	5400
45	TATATACTGC AAAATATTGA TTATCTACAA ATTGAGTTAC AGTGTGTGTG TCGATGTGAG	5460
	TTGTTATAGG AGTAGTATTT GGTTTCATGAT TGCCTAATTC AATCACATCT TTACTTTGCT	5520
	CTAAGTGCAA ATCACGCAAT TGACCATTTT GATCTCGTCT ATCATAGTCA TAAATACGGT	5580
50	ATGTCGTATC GGAGGATTGT TGTGTCTCTA AAATTAAAAT ACCCGAACCA ATGGCATGGA	5640
	CAGTGCCAGC AGGAACATAA TAAAAGTCAC CGGGCTTAAC AGGTATACGT TTGAAAAGAC	5700
55	TGTCAAATTC ATGATTATCA ATCATGTCTA TTAACGTCTG TTTATTATGT GCATGTACGC	5760

	GTTCGCCTTC GTGTTTTAAA GCGTAGTCAT CATCTGGGTG AACTTGAACA GATAATTTAT	5880
	CATTGGCATC TAATACTTTA GTTAGCAGAG GGAAACTATC TCGTGAATCA TTATCGAATA	5940
5	ATTACAGATG TTGTGACCAA AGTTGATCTA GGGTCATATC CTTGTATGGA CCATTGATAA	6000
	TTGTATTAGG ACCATTTGGA TGTGCAGAAA TTGCCCAGCA TTCACCAGTT GTTTCATTAG	6060
10	GGATATCATA GTTAAATGCT TTTAATGCAT GACCGCCCCA AATTCTGTCT TTAAAAACGG	6120
	GTTGTAAAAA TAATGCCATA GTTAAAACTC CTCTATATTT TCATTAATAA GTTATAAATT	6180
	TCTGTAGTAC TGTTTGCATT AATTAGTGAT TGGCGTGTCT CATCATTCAT TAACGCTTTA	6240
15	GATAAGCGCT GAAGTATTTT TAAATGTGTA TCCTGACTGT TGTTTGGTAC GGCAATTAAG	6300
	AATATCAATT GAGGTAGACT ACCATCTAGA CTGTCCCAT TAACACCATG ATTATTTTTTC	6360
	ATAACAGCTA CAATCGGTTG TTTTACAACA TCAGACTTTG CATGTGGAAT GGCCACGTTT	6420
20	ATGCCAATAG CTGTCGTAGm tCcATTTCAC GTTCTAGTAT TGCATTTTTT AAATGCGATG	6480
	TGTGCTCTAC ATAACGGCAA ATTTTAAGTT TATGAATCAA CATATCAATT GCTTCGTTTC	6540
	GAGACATGTC GTGATCAGTA ATTATCATAG TTTGTTGATC AAAAACATGA GAAGGTTTAT	6600
25	TGAGATGTGA ATGTTTCGCG GTGTTATCTA CATTGTCAAC CTCTGTATCA TGTGTGTAA	6660
	TATCTGTATC ATGAAGTTGC GTGTGTTGCG CTGGTGCATC TACTGCTATA ACTGGTGTAT	6720
	TGCGTTTTAA TAATAGTACA GTAGTCATTG TGACAAGACT ACCTACTATC ACTGCAAAGA	6780
30	TAAACCATAA TACATGATCA ATACCACCTA ATACAGCCAC GATTGGACCT CCATGTGCGA	6840
	CTCTATCGCC GACACCACCA ATGGCTGCAA TGAATGATGC AATCATTGCA CCAATGATGT	6900
35	TTGCAGGTAT AATGCGCAAT GGATCTTGGG CTGCGAAAGG AATAGCACCT TCAGTAATAC	6960
	CAAATAGTCC CATAGTGAAG GAAGCCTTAC CCATTTCTCT TTCGGAATGA TTGAATTTAT	7020
	ACTTTTGAAC AAACGTTGCT AAACCTAAAC CGATTGGTGG TGTACATACA GCAACTGCGA	7080
40	CCATACCCAT AACGGCGTAA TTACCTTCAG CAATAAGTGC TGAGCCAAAT AAAAATGCTA	7140
	CCTTGTTTAC TGGACCGCCC ATATCGAAGG CAATCATCGC ACCTATAATC ATCGCAAGTA	7200
	TAATAATATT AGCACCTTGC ATACTTTTTA ACCAGGTTGT TAATGCCTCA AAAATATTAG	7260
45	AAATTGGTGC ACCGATTAAA AATATAAATA TCAATCCTAC AACGACCGAT GAAATAATGG	7320
	GAATAATAAT GATAGGCATA ATTGGTGCCA TTGCTTTTGG AACTTTAATA TCTTTAATCC	7380
	ACTTTGCGAT ATAACCTGCT AAGAAACCAG CAACAATACC ACCTAAAAAT CCTGCGCCTG	7440
50	CATCACTGCC ATAAAACTA CCGTCAGCAG CGATAGCGCC GCCAATCATA CCAGGAACAA	7500
	GACCGGGcTT GTCAGCGATA CTAACAGCGA TATATCCAGC TAGTATTGGA ACCATAAATT	7560
55		

	ATCCTTTTGA TGTCGTTTCA CCGCCTAGAG TCAGCGCGAT GGCATAAGG AGTCCACCAA	7680
	CTACGATAAA AGGAACCATA AACGATACAC CGTTCATTAA ATGTTGATAC ACCATTTGAA	7740
5	TACCATTTTT AGACTTACCG CGATCTTTTCG AATGATAATT TGTTTCAGAT TGATAAATAG	7800
	GCGCATCTTG ATTAATGATA CGTTGAATTA GACCTCTCGG ATTATGAATC CCTTCGCGAA	7860
10	CATTTTCATT AATCAACCGT TTACCAACAA ATCGGGACAG ATCAACTTGT TTATCAGCTG	7920
	CAATTATGAC ACCGTCAGCT TCTTCGATGT CTTGCGTAGT TAAAACATTT TCAGCACCAA	7980
	CACCGCCCTG TGTCTCTACT TTAATATCCA CACCCATTTC TTTTGCTACC TGCTCAAGCT	8040
15	TTTCTTGAGC CATATATGTA TGTGCAATGC CATTTGGGCA TGAGGTAATA GCTACAATTT	8100
	TCATAAAATC ATCTCCTTTT CTATATTGTA AGCGTATTCT CGATACTAAA AAAAAGAATA	8160
	ATTACCGTTA CTAGTGGCAA TTATTCTTGT AAGTATTCAA ATAAGTGTG CTTTAACTA	8220
20	TGATCATCTA AACTACATAA ATGGTTCAGT GAATCATCAT CCAAGTAGC AATTAATTGC	8280
	ATCATTTGTT TTGTAAAAGC TTTGTCTTTA TCGGAAATCG CTAAGAAAAA GACAAGTTTG	8340
	ACATCGTGTT GTCGCCAAGG AAAACATCT TTTGTGCGAA AAATAAGCAC ATGTGATTGT	8400
25	AAAACTTTTT CAGGATCTCC ATGAGGAATC GCCATAAAAT TACCTATGTA TGTAGAAGAT	8460
	GATTTCTCAC GCTCTAAAGC TGATTCGATA TATCCTTCTA CAATCGCATG ATGTGCTTGT	8520
	AATATTTTTT GAGCTTCTTC AAAAATTTGC ACAGTATGCC GTGATTTTTG TTCAGTATTT	8580
30	ACGACAAGGA AATTGACAGT GTCCATATGA TGATGTGCTT GAACCGGATT TTGCTTTTGC	8640
	TTCAACAAGT GTCTGATTTT GTGACGATCA TCTTCAGAAA ATAATGGTGC AACCTTGATA	8700
35	GTCGTCAGGT GCTTAGGAAG TATGTTTAGC GTTTGTTTAG GAATATCATG GGTGCTTATT	8760
	AATAAATCTA CATTGTCAAA GTGATAGTGT GTTATATTTT CTAGTTTAAT CGTATTTATC	8820
	ACTGACAACT CTTGCGATAA GTTATTTIATT TTAGTTTCTA AAAAATTCGA CACACCTAGA	8880
40	CCATAATAAC AAGCAATGAC TACATTTAAT TGTGTTTTGG TACGACGCTC GATGGCAGCT	8940
	TGAAAATGAA TTGTTAAAAA TGCAATTTCA TCTTCGCTCA TCTCTATATC AGTATCAATT	9000
	GCTAATTTAT CAATCGCTTC AAAAAGTGTG TTAACACAA AGGGATAGAG TTTTTTAATC	9060
45	TCTATAACTA AAGGATTGTT TAAATAAATG TTTTGAGTGA TACGTAAATA TGCTTTACTA	9120
	AAATGATTAT ATAAATTTTG TTGTAAAATC GAATCTTCAT TGAAAGGTAC ATGAATACGT	9180
	TGCTGCATCA ATTCGATTAA GCGATCAATA TAACTTTGTA TAAATATACG TTCTATGCCA	9240
50	ATATCGAGTT TATTAAAATG ATAAGCAATA AAGAAAGAAA ACATATTGAT TACTTTTTTCG	9300
	TTCAAGTCAT AACCTAATCT TTCGTTGATT TGCTTAATGC AAGATTGAGA TATCAATTTT	9360

55

AGATGAATTA AAAGCTGTTG TATTTGAATA TCAGTTGTTT CAATACTATG TTGTTGAAGT 9480
 GTCTCTTGTA TAATATGCGA AATCATCCTT TGGTGTGAAT CAGGTAATTC aTTTAAAATT 9540
 5 AGGTCTTCAA CATGTACATG CCCTGATGAT AATTGATTTA AATGGATGAT GGCATTAGTG 9600
 ATATCATTAT CTGTTCCATC GAC 9623

(2) INFORMATION FOR SEQ ID NO: 167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1021 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

20 ACCGTGGAAA CACGTCTAGT CAATCAGAAA GCGATAAAAA TGTGACTAAA TCATCTCAAG 60
 AGGAAAATCA AGCAAAAGAA GAATTACAAA GCGTTTTTAA CAAAATTAAC AAACAATCAA 120
 GTAAGAATAA TTAAAAAATT TTGATATTGT CTATGTTTAT AGTTCACAAG CCATTCAACG 180
 25 TATTGTAAAC TAAGGATAGT GTATTTTTTT AATAGTAATT TGTGAGGAGG TGCCTATCTA 240
 TGGAAGAACA TTACTACGTA AGTATTGATA TTGGATCATC AAGCGTAAAA ACAATAGTAG 300
 GCGAGAAATT TCACAATGGT ATAAATGTGA TAGGTACAGG ACAAACCTAC ACGAGCGGTA 360
 30 TAAAAAATGG TTTAATTGAT GATTTTGATA TTGCGCGACA AGCAATCAAA GACACAATTA 420
 AAAAGGCATC AATCGCTTCG GGTGTTGATA TTAAAGAAGT TTTCTGAAA TTACCTATCA 480
 TTGGAACGGA AGTTTATGAT GAATCAAATG AAATCGACTT TTATGAGGAT ACAGAAATCA 540
 35 ACGGTTTACA TATCGAAAAA GTATTAGAAG GTATTAGAGA AAAAAATGAT GTGCAAGAAA 600
 CAGAAGTAAT TAATGTGTTT CCGATTCTGT TTATAGTCGA TAAAGAAAAT GAGGTTTCAG 660
 40 ACCCTAAAGA ATTAATTGCC AGACATTCAT TAAAGGTTGA AGCAGGCGTA ATTGCTATTC 720
 AAAAATCGAT TTTAATTAAT ATGATTAAAT GCGTAGAAGC ATGTGGTGTT GATGTATTAG 780
 ATGTTTACTC TGATGCATAT AACTATGGTT CAATCCTAAC AGCTACTGAA AAAGAGTTAG 840
 45 GTGCATGTGT CATTGATATT GGTGAAGACG TTACGCAAGT TGCTTTTTTAT GAACGCGGTG 900
 AATTAGTAGA TGCTGATTCT ATCGAAATGG CAGGGCGTGA TATTACaGAC GATaTTGCAC 960
 aAGGrTTaAA CACTTCTnAT GAAACTGCTG nAAAAAGTTA AACACCAATn TGGTCATGCA 1020
 50 T 1021

(2) INFORMATION FOR SEQ ID NO: 168:

(A) LENGTH: 7963 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

10	TAATCTATTA TAAAACTGT CCATACCCTT TGATTACCTT CTCTTCAGGT ACAGGCCACA	60
	CTTGAGGCCA TAAGCCATAT GCTTGCTGTG AATAAAATTG TGCCATTTGT AACAAATATAA	120
	TATATACAAA TAAACACCCA ATAATTGCTG TCACTAATGG ATATGATAAC CAAACCATTAA	180
15	ATAAACTGC AATAATTACT AACCTAAAGA TAATATTAAA TCGTCTCTC CCTCTTATAA	240
	AGCTTCTAAT AAATAAGAAT AAATACATCG CATTAGAGTT AAATTTACTA CCCTTTGGAA	300
	CTGGTAAAAG TATATCTAGA TAACTTCTTC TGACTGCAGA TTCTTTCAA TGTTTTACAT	360
20	CGGTGAACAT ATTAACAAAT TTATAATAAT TCATATGATG TCGATGTTTG ATTGCAATCA	420
	TTTTCTCCCA AGGATACAAA AAGCCTGTTT TATATTTTTT AACTAAAAT TCTATTACAA	480
	CAGGCAAAGC AACCATCACA AATGCGATGT ACCATTTTGG AGCTAATAGT AAGTAATATG	540
25	TTAGAGCAAA GGTGATGAAT GATATTAAAT TAACTTGCCA TGTTTTAAGT CCCGATTGAT	600
	ACCATTGCCA TCTTAAGCGT AAACCAACAT ATGGAAAAAT TAATGCACTG ACTCCAAAAC	660
	AAATATAAAA TGCCACATTA TGTGATTAA TATTGTAAAA CAACGGGAAC ATTACAATAA	720
30	CAATAATGAG TTGGATTAAT ATGCGCGCAA AGTAACTATA TAAATCGCA TGACGCATAA	780
	ATTGAGACAT GTGTTTTTCA AATGGTAATA AAAAGATTTT ATCCGCTTCT TTTAACAGTG	840
35	GTCsCmTTGG AAAAAATAGrT GTCAACGCAA CAATCACTGC TGCTATTaAT GAAAAATTGa	900
	TATTGCTTGG AATATGTTTT AACCATTcAC CATATCCaA AATAAATGCA CCCAGCAAAA	960
	TAAGTAAAAA GACCATGAAA TGACCATTAA ATATAAACTT ATTATAATAA TTTTtCTCTT	1020
40	TACGAAGGGC ATGTAATCTT TTATTAAATA ATGTGGTAgC TTGGTTACGC ATGTACATCT	1080
	CCACCTTGCG TCACATGAAT ATATATATCG TCTAATGTTT GATTATGTAA GCCAGTTTGT	1140
	TGTCTCAATG CTCTAAATC TCCAAATGCA ACGACTTCAC CTTCGTCTAG TATGaTAAAA	1200
45	CGATCACAGT AACGTTTCAGC TGTGCTAAA ATATGTGTAC TCATTAGAAC GGTTCTACCT	1260
	TCGTTTTTCT TTTCAACCAT TAAATCTAAC ATGGATTGAA TTCCTAATGG ATCTAGGCCA	1320
	AGGAATGGTT CGTCTATAAT ATACAATTCG GGATTAACTA TAAACGCACA AATAATCATG	1380
50	ACTTTTTGTT TCATCCCCTT AGAAAAATGA CTCGGAAAAA CTTTCAACTC ATTTTCTAAA	1440
	CGGAATGTCT TTAATAATGG CATTGCTCGA TTCATCGTTT CATCACGATC AATATCATAT	1500

55

	TCCGGAATAT AAGATAACTT TCTTCTATAA GCCTCTATGT CATCATTAAT GTTGATATCT	1620
	GAAATTGATA GAGATCCTTC CATAGGTGTA AGCAATCCTA GCATATGTTT AATCGTTGTA	1680
5	CTCTTACCAG CGCCATTAAG GCCAATAAGT CCAACAATTT CGCCTTTGTT TAATTCAAAA	1740
	TTTATATCTT TAATTACAGG GCGTTTTCCA TATCCACCTG TAAGCTGTTT TACTTTAACT	1800
	GTCATAAGGC ACCTCCATGA CTTATATTGT ACCAAAAATT ATAAAATGCT CATATTAAAT	1860
10	ACACATGTCC TAATATCGAA TTTTITAGCGA CAATGTTATA ATGAATGGTA ATACTAGTTG	1920
	AAAAGGAGTG TAGTCATCAT GTCAGAAACA ATTTTCGGCA AAATTTTAAC TGGAGAAATT	1980
	CCTAGCTTTA AAGTATATGA AGACGATTAT GTCTATGCCT TTTTAGATAT ATCACAAGTT	2040
15	ACTAAAGGAC ATACGTTATT AATTCCTAAA AAAGCTTCTG CTAATATCTT TGAAACTGAT	2100
	GAAGAAACAA TGAAACATAT CGGTGCAGCA TTACCTAAAG TAGCAAATGC TATTAAGCGT	2160
20	GCATTTAATC CTGATGGTTT AAACATTATT CAAAATAATG GTGAGTTTGC AGATCAATCT	2220
	GTATTTTATA TTCATTTCCA CTTAATTCCT CGATACGAAA ATGATATTGA TGGATTGGT	2280
	TATAAGTGGG AAACACATGA AGACATTTTA GATAACGATG CAAAACAACA AATTGCTGAA	2340
25	CAAATTCAG CACAATTTTA AATGTATGCT TAATCTAAGC TCGAACGGGT ATAATATGAT	2400
	TAATATTATA ACAATTGCGT TTGAAGTGAT AACATCAAGG TTAGCAATTT TAAACAAAAT	2460
	GAGTTATCAA GATAACAGAT GTTAAAAGTG AGGAGAATAT AAATGAAAGC ATCACGCATT	2520
30	CTATTCGGTA TCGGTGTTGG CGTAGCAGCT GGTTTTGTAG TTGCACTTCA AGGACGTGAC	2580
	GACAAAAGTG TCAAGAACAA CACGATCGAT CGTACTGCCC CTACTGTTTC AAAATCAGAA	2640
	CTACAACGTG AATTTGAAAC GATTAAACAA AGTTTAAATG ACATTTTAAA CTATGGTGT	2700
35	CAAATTAAAA ACGAAAGTGC GGAATTTGGT AGTTCAATTG GTGGTGAAAT TAAGTCATTA	2760
	CTTGGAACCT TCAATCTGA CATTAACTCT AATATTGAAC GTTTACAGTC ACACATCGAA	2820
	AATTTACAAA ATCGTGCGCA GGATATTGGA AACGAAATTT CTAAGTAGCA GGTACGTTT	2880
40	TCGATCAGAA CTATTTTAT TAGTAACAGC ATATTTATTT TTTAAAATTA AATGCCAAAT	2940
	AAACGAGATG ACATTAGAAA TTAGATATTT CTGTGTCATCT CTTTTTTAAA ACTCAAATGA	3000
45	ACTTATGTTT ACAAATTATA GGAAGACATT GTTTGTAGTG ATTTTCGCTT AAATCATATT	3060
	TATGAATTGA TTGAAAACAT TGCTTAGGAT TCATTGTGTT ATCCTTGCAC TTTGATTACG	3120
	CTTTACTTAA ATCATTATCG ACAAACAACA TACTTATATT TTCATTGAGC CGAACCTTAT	3180
50	ATACACATTA CATATACCTT ACTTGCACAA ATTATTAATC TGGTGTTTAT TATAATTACA	3240
	TATCACTATA TTTTITAGCAT TTGTATAACT TAGTTGGTCA AAAGATGCTT TTGCATATGC	3300

55

	TTTCATAAGT GATGCTTTAT TAGCAAGAAT ATGTGTTTCGC AGAAATTTGT TCTGCATTCT	3420
	ACTTCTACGC TAGTCAATCA GACAATTTTA CCAATCCCCA CTTTCGCGTT TCAAATCAAA	3480
5	CAATACGTCG CTCCTTTCTT CTTATATAAC AATTCTTCTA ACATGATATG TTACTATTGA	3540
	ATTACTGAAC CTGAGTTAGT TATAATCTAA CTTATATTGA AAAGAGATGA GGCCTAAGAT	3600
	ATGTTTTTAT GTAAAAGACA AATTGATATC AATGCACGAT TTGGTTTGCC TAGAATTGCA	3660
10	TTTATGAGTG CAGTTGCAAC CATCATTATG TTTTITAGTTA GTTATGAAGT AATGTATTTT	3720
	TTATCTAATA CGCCATTATC AGATAGACAT TTTCTCATCT TTTTATTACT TGTATTTATG	3780
	ACGTATCCAT TACATAAAAG TATACATTTA TTATTTTCT TACCATATAG AAAATCGTTT	3840
15	AAAGTTCATA AGTTAACTAA AAGAAAATGG CTTATATTCT ATAATACCTA CGTCAATCAA	3900
	CCTGTACACA AATTTTATTT TTGCATTAAC TTAATATTGC CGTTAATTAT CTTATCTGCA	3960
20	ATGTTTCGTTT ATCTAACAAT TTCATTCCCG CAATATGGAC ATTATTTTAT GTTCTTATTG	4020
	GCATTGAATT TCGGTATTTC CATTACAGAT TTATTATATT TAAAAATAAT TATATTTTCT	4080
	AATTATGGAC AATATATAGA AGAACATAGT ACAGGTATTA ATATTTTGAA AAAAATTAAA	4140
25	AATCCATATC ATTTATAACA AAATAATTAT AGCAAGGTGT TATTATTTGT TTTTAGGCTA	4200
	TGTAATAgcT tACAATCAAA TGTATATAGA CCTTGTTTTT TTATTTTCAT CAATTTCTAC	4260
	CCCTAAACCT AATGCTCTAG TCTGATGTCA TGGGTATTG ATTGGTGATA ATATAAAACT	4320
30	ATGTTATATT CACGATGATT AACTTACAAA GGAGTTTCAA CTATGAAGAT GATAAACAAA	4380
	TTAATCGTTC CGGTAACAGC TAGTGCTTTA TTATTAGGCG CTTGTGGCgC TAGTGCCACA	4440
	GACTCTAAAG AAAATACATT AATTTCTTCT AAAGCTGGAG ACGTAACAGT TGCAGATACA	4500
35	ATGAAAAAAA TCGGTAAAGA TCAAATTGCA AATGCATCAT TTAAGTAAAT GTTAAATAAA	4560
	ATTTTAGCTG ATAAATATAA AAATAAAGTT AATGATAAGA AGATTGACGA ACAAATTGAA	4620
	AAAATGCAAA AGCAATACGG CGGTAAAGAT AAATTTGAAA AGGCCCTTCA ACAGCAAGGT	4680
40	TTAACAGCCG ATAAATATAA AGAAAATTTA CGTACTGCTG CTTATCATAA AGAATTACTA	4740
	TCAGATAAAA TTAAATCTC TGATTCTGAA ATTAAAGAAG ACAGCArGAA AGCTTCACAC	4800
45	ATTTTAATTA AAGTTAAATC TAAGAAAAGC GACmAAGAAG GCTTAGATGA TAAAGAAGCG	4860
	AAACAAAAAG CTGAAGAAAT TCAAAAAGAA GTTTCAAAAG ATCCAAGTAA ATTTGGTGAA	4920
	ATCGCTAAAA AAGAATCAAT GGATACTGGT TCAGCTAAAA AAGATGGCGA ATTAGGTTAT	4980
50	GTTCTTAAAG GACAACTGA TAAAGATTTT GAAAAAGCAC TATTTAAGCT TAAAGATGGT	5040
	GAAGTATCAG AGGTTGTAA ATCAAGCTTT GGATATCATA TTATTAAAGC TGATAAACCA	5100

55

AAAAATCCAA AATTATTGAC TGATGCATAC AAAGATCTAT TAAAAGAATA CGATGTTGAC 5220
 TTTAAAGATC GTGATATTAA ATCAGTTGTC GAAGATAAAA TCTTAAACCC TGAAAAACTT 5280
 5 AAACAAGGTG GCGCACAAGG CGGACAATCC GGCATGAGCC AATAACACAA AACCGAGCGA 5340
 CCGTGGTTCA AAAATCATAC CACGGCCGCT CGGTTTTTTC GCATTAAAAA TCGGACAGAT 5400
 GAGCTCATGT TTCAGTATAC TCATCTGTCC GATATCTTTT AATTCTTAAT CGAGTGATTC 5460
 10 AGGATTGTAG AATCTACGAT TTTCAAGACC AAATATTTTA TCTGTAAACT GACCCTTGTC 5520
 AGTTTTTTTA TATGCCTTTT CAAACATATT CATTCTAGCA TCGATATTAT CGATATAGCA 5580
 TAAAATTTCT GCTTCTTTTA AGTATGGCAG TTTTGAGAAA CCATACTCTA ACTTACCATG 5640
 15 ATGAGATAAA ATCATATGTC TTAACAACAT GATTTCTTCT CCTTCAATGT TCAATTCACG 5700
 AGCTGCTTCA ACTACTTCAT CACTCGCAAT CGAGATGTGT CCTAATAAGT TACCTTCGAC 5760
 20 TGTATACGAC GTCGCAACAG GACCACTCAA TTCTCTAACT TTACCAATAT CATGCAAAAT 5820
 AATACCACTA TATAACAAAC TTTTGTTTAA CAATGGATAA ATGTCAAAA TTGATTTTGC 5880
 AATACGTAAC ATCGTTAATA CATGATAGCT TAAGCCACTC GCAAAGTTAT GaTGATGAGA 5940
 25 ACTAGCAGCT GGATATGTGT AAAATCGTTC TTGATATTTT TTCAATAAAT GACGTGTGAT 6000
 ACGTTGTAAA TTAGCATTTT CAATATCTAG CAAATAATGA GAAATCTCTT CTTGTATTTT 6060
 TGCCGGTGAT AAAGTGACAC CATCTACAAA TTGTTCTGTT TTTAATTGAT CTTCAGTTGT 6120
 30 CGCTAGTCTA ATTTGGTTGA CTTTCATCTG TTTATTTCCG CGATAGTTTA TGATGTCACC 6180
 TTTAACATGT ACAATTTCTT CAGGCTTGAT TGTGTCATA TCATTTTTTG TAGCCGTCCA 6240
 AAATTCGCT TCAATTCAC CACTTTTATC TTGCAAATGT AATGTCATAT AATCTTTACC 6300
 35 TTGTGCTGTT ACACCCTGTG TAGCTTTATG CACTAAGAAA AAGTGATCAA CTGAATCTCC 6360
 GGGATTTAGA TTCTCTATAT TTCTCATCGT TTCCCGCCTT CCTCTATTTT GTTTAATGTA 6420
 ATCACTTCTT TTGATGGAAC AATATTATCT TTTACACATG TAAAGTATAG TACTTGATAG 6480
 40 TGTTCGATA ATGATCGTAA ATAATTCAAC ATTTTTTCAG TACGTTTTTT ATCAAAATGA 6540
 ACAAATGCAT CATCAACAAT TAATGGGAAC GGATAATATG GTCTTAGTAC CTTAATTAA 6600
 45 CTGATACGTA AAGCTACATA AAGTAATTCT TTTGTAGATT GACTTAGTTC AACAGGATCA 6660
 TATAATTGAC CATTAACATG TTTAACCGTA ATTGAATCTT CATTATAGTT AATCATCGTA 6720
 TATCTGCCAT CTGTTAAATG CTTCAATATT TCTACCGCTT CATTAATAAC TTGAGGCAAA 6780
 50 CGTTTATCTT TAATTTGTTT AATGTGTTCA TCAACTAAAC TTTGTAAATA ACTTAAACTT 6840
 GCCCAATCTT TTGCGATATC ATTAAGTTGA TTTTAAAGAC TGTGATATTC ATGTCTTAA 6900

55

GCTTGCAATTT CAAGATATTG CTCATTATAT TCGTCAACTT GAGTAGCCAA TAAATGATCT 7020
 TCTTCTTCAA GTTGTGCAGT TGTTTTTTCA CTTAACTAG AACTTAATTC ATAAGAATAG 7080
 5 TTTTGGTTCT CAAGATATTT AGTTAAATCA TTAAACGAC TCAAATTACT AGTATAAGTT 7140
 TGGTAATCTT CATGATGTTG GTAAAAATCT TCTTCAGTAC CAACATTGAT AAAATCGAAT 7200
 AGTGCTGTAA TTTCTTTATT ATTTTCTTCT AATTGAGCAT TTAAATGATT TAATTCATTT 7260
 10 GTAACAAGTT TGGTATTTTC AGCATTAAATA CGCCATTTTT CATTCGTGTC TTCAGCTGAT 7320
 TTCAACCATT GTtGCACATC GTGGAATAAA GATAATTTGT TGAAATAAAC AAATTGTGAT 7380
 TTTGTAACAG CTTCAGCATG ATTGTAGAAT GTATCTAATT CTTGAACCAA TGCTGGCGT 7440
 15 TGTTGATTTA AATCACTGAT ATGTTGATCT AATGCTTTAA TATTCGCCAT TGTAGAAATA 7500
 CTATCAACAA TTAAATCATT TGAAATTTTA GATGATAAGT ATAATTCATC CTTAACGTTT 7560
 TCAACTGTGG ATTGTAATTC ATCATGACGC CCTTTCGCAT CATTTAAACG ACCTTCAATA 7620
 20 TACTGACGTT TCTCTTCTAA AATATCTTTA TTTTCAAAG CTTGTTGCCA GTGATCACGA 7680
 ATGCGATATT GCTCATCAAG ATCAAAATCT AAGTCATAAT TTTCATCTAA AATGGCTAGT 7740
 25 TGTGCTTTAA TTTCTTCGAT TTCATCTGTG ATGGCCTCGC TATAATCTAC TTCTTTTGAT 7800
 TTAGACATGA TGATACCGAT AACAAATACT AAAGTTAATA CTGCGAAAT AATACCAAAC 7860
 AACATGTTGT TTGAAATAAA TGAGAAGGCA GTTAAACCAA TACCTACTAA TGTTAAAAGT 7920
 30 ATAAACGTTG TTCGKAACAA TTTTGTGACGT TTTTGtTTT CTT 7963

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3958 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

ATATTGTCTT TACAATAGTT TGCTATGGAG GTAATTAACC AATAGGAGGA ATTTATAATG 60
 45 GCAGTAATTT CAATGAAACA ATTACTAGAA GCGGGTGTTC mCttCGGTCA CCAAACACGT 120
 CGTTGGAACC CAAAAATGAA AAAATATATC TTTACTGAGA GAAATGGTAT TTATATCATC 180
 GACTTACAAA AAACAGTGAA AAAAGTAGAC GAGGCATACA ACTTCTTGAA ACAAGTTTCA 240
 50 GAAGaTGGTG GACAAGTCTT ATTCGTAGGA nCTAAAAAAC AAGCACAAGA ATCAGTTAAA 300
 TCTGAAGCAG AACGTGCTGG TCAATTCTAC ATTAACCAA GATGGTTAGG TGGATTATTA 360

GAAGATGGTT TATTCGAACT ATTACCTAAA AAAGAAGTAG TAGAACTTAA AAAAGAATAC 480

GACCGTTTAA TCAAATTCCT AGGCGGAATT CGTGATATGA AATCAATGCC TCAAGCATT 540

5 TTCGTAGTTG ACCCACGTAA AGAGCGTAAT GCAATTGCTG AAGCTCGTAA ATTAAATATT 600

CCTATCGTAG GTATCGTTGA CACTAACTGT GATCCTGACG AAATTGACTA CGTTATCCCA 660

GCAAACGACG ATGCTATCCG TCGCGTTAAA TTATTAAGTG CTAAAATGGC AGATGCAATC 720

10 TTAGAAGGTC AACAAGGCGT TTCTAATGAA GAAGTAGCTG CAGAACAAAA CATCGATTTA 780

GATGAAAAAG AAAAATCAGA AGAAACAGAA GCAACTGAAG AATAATCAAC TGTGGAATCT 840

GACTTAGATA TAGTTTAAAT GGGTGATAAG ATATTAATGC TTATCACCTT TTTTAAAAAG 900

15 AAAATCGAGG CAAATTACAA ATATTCAATT AGAGTATTGG CAATCTTGCC TATAATAATG 960

CTAAAATCAT AATATATAAA ATGATAACTT ATTGGAGGAA TAATGAATGG CAACTATTTT 1020

20 AGCAAACTT GTTAAAGAAT TACGTGAAAA AACTGGCGCG GGTATGATGG ATTGTAAAAA 1080

AGCGCTAACT GAACTGATG GTGACATCGA TAAAGCGATT GACTACCTAC GTGAAAAAGG 1140

TATTGCTAAA GCAGCTAAAA AAGCAGACCG TATTGCGGCT GAAGGTTTAG TACATGTAGA 1200

25 AACTAAAGGT AACGACGCAT TATCGTTGAA ATCAACTCTG AAACAGACTT TGTGCTCGT 1260

AACGAAGGTT TCCAAGAGTT AGTTAAAGAA ATCGCTAATC AAGTATTAGA TACAAAAGCT 1320

GAACTGTTG AAGCTTTAAT GGAAACAACT TTACCAAATG GTAAATCAGT TGATGAAAGA 1380

30 ATTAAAGAAG CAATTTCAAC AATCGGTGAA AAATTAAGTG TTCGTCGTTT TGCTATCAGA 1440

ACTAAAAGTG ATAACGATGC TTTCGGCGCT TACTTACACA TGGGTGGACG CATTGGTGTA 1500

TTAACAGTTG TTGAAGGTTT AACTGACGAA GAAGCAGCAA GAGACGTTGC TATGCATATC 1560

35 GCTGCAATCA ACCCTAAATA TGTTTCTTCT GAACAAGTTA GCGAAGAAGA AATCAACCAC 1620

GAAAGAGAAG TTTTAAAACA ACAAGCATT AATGAAGGTA AACCAGAAAA CATCGTTGAA 1680

AAAATGGTGG AAGGACGTTT ACGTAAATAC TTACAAGAAA TTTGTGCTGT AGATCAAGmT 1740

40 TCGTTAAAAA CCCTGATGTA ACAGTTGAAG CTTTCTTAAA AACAAAAGGT GGAAAACTTG 1800

TTGACTTCGT ACGCTATGAA GTAGGCGAAG GSTATGAAAA ACGCGAAGAA AACTTTGCGG 1860

45 ATGAAGTTAA AGGACAAATG AAATAATCTG TCATAAGTA AAACAAGGAA GAAGACACCT 1920

TTAATGTTGC TTTATTAAAA TGTAATCAT TCTAATAAAA CGACAACTGT GTCTTCTTTA 1980

CTGTATATG TTACATATAT TCACGATAGA GAGGATAAGA AAATGGCTCA AATTTCTAAA 2040

50 TATAAACGTG TAGTTTTGAA ACTAAGTGGT GAAGCGTTAG CTGGAGAAAA AGGATTTGGC 2100

ATAAATCCAG TAATTATTAA AAGTGTGCT GAGCAAGTGG CTGAAGTTGC TAAATGGAC 2160

55

	TTAGGTATGG	ACCGTGGAAC	TGCTGATTAC	ATGGGTATGC	TTGCAACTGT	AATGAATGCC	2280
	TTAGCATTAC	AAGATAGTTT	AGAACAATTG	GATTGTGATA	CACGAGTATT	AACATCTATT	2340
5	GAAATGAAGC	AAGTGGCTGA	ACCTTATATT	CGTCGTCGTG	CAATTAGACA	CTTAGAAAAG	2400
	AAACGCGTAG	TTATTTTTGC	TGCAGGTATT	GGAAACCCAT	ACTTCTCTAC	AGATACTACA	2460
	GCGGCATTAC	GTGCTGCAGA	AGTTGAAGCA	GATGTTATTT	TAATGGGCAA	AAATAATGTA	2520
10	GATGGTGTAT	ATTCTGCAGA	TCCTAAAGTA	AACAAAGATG	CGGTAAAATA	TGAACATTTA	2580
	ACGCATATTC	AAATGCTTCA	AGAAGGTTTA	CAAGTAATGG	ATTCAACAGC	ATCCTCATTC	2640
	TGTATGGATA	ATAACATTCC	GTTAACTGTT	TTCTCTATTA	TGGAAGAAGG	AAATATTAAA	2700
15	CGTGCTGTTA	TGGGTGAAAA	GATAGGTACG	TTAATTACAA	AATAAATTTA	GAGGTGTAAA	2760
	ATAATGAGTG	ACATTATTAA	TGAAACTAAA	TCAAGAATGC	AAAAATCAAT	CGAAAGCTTA	2820
20	TCACGTGAAT	TAGCTAACAT	CAGTGCAGGA	AGAGCTAATT	CAAATTTATT	AAACGGCGTA	2880
	ACAGTTGATT	ACTATGGTGC	ACCAACACCT	GTACAACAAT	TAGCAAGCAT	CAATGTTCCA	2940
	GAAGCACGTT	TACTTGTTAT	TTCTCCATAC	GACAAAACCT	CTGTAGCTGA	CATCGAAAAA	3000
25	GCGATAATAG	CAGCTAACTT	AGGTGTTAAC	CCAACAAGTG	ATGGTGAAGT	GATACGTATT	3060
	GCTGTACCTG	CCTTAACAGA	AGAACGTAGA	AAAGAGCGCG	TTAAAGATGT	TAAGAAAATT	3120
	GGTGAAGAAG	CTAAAGTATC	TGTTCGAAAT	ATTTCGTCGTG	ATATGAATGA	TCAGTTGAAA	3180
30	AAAGATGAAA	AAAATGGCGA	CATTACTGAA	GATGAGTTGA	GAAGTGGCAC	TGAAGATGTT	3240
	CAGAAAGCAA	CAGACAATTC	AATAAAAGAA	ATTGATCAAA	TGATTGCTGA	TAAAGAAAAA	3300
	GATATTATGT	CAGTATAAAA	CTAATATACA	ATGACATATT	AAAATGCCAG	TATTAAACGA	3360
35	TAATGTAACA	TTTAAATGG	GCATGTTTAA	TTAAATCAAA	GATGCATGTG	ATAATTTAAA	3420
	TTCAGAAATGA	GCATAAAAAT	GGTGTTTTAA	CAAGTTAATT	AAACATATAC	TTTATAAATA	3480
	ATAGGCATTA	GGTATATTGC	TATAATAAAG	TTATGTAATT	TTTAACCTCA	GTATGTATGT	3540
40	CACATTTCTG	GTGTAAACTG	TACCGAGTCA	GACTTTGGTA	CAGTTTTTTT	ATTTGCTTAT	3600
	TCAATGCATT	AAATGAGTAT	GATAAAATGA	TAATGATTGT	TTAGTAACTT	ATACTATATG	3660
45	ACAGAGATGA	TCAGGCTCGG	AGGAAAGACC	ATGTTTAAAA	AGCTAATAAA	TAAAAAGAAC	3720
	ACTATAAATA	ATTATAATGA	AGAATTAGAC	TCGTCTAATA	TACCTGAACA	TATCGCTATT	3780
	ATTATGGATG	GTAATGGGCG	ATGGGCTAAG	AAGCGAAAAA	TGCCTAGAAT	TAAAGGTCAT	3840
50	TACGAAGLAT	GCAAACAATA	AAAAAAATTA	CTAGGGTAGC	TAGTGATATT	GGTGTTAAGT	3900
	ACTTAACTTT	ATACGCCTTT	TCCACTGAAA	ATTGGTCAAG	ACCTGAAAGT	GAAGTAAA	3958
55							

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5333 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

10	ATTAAAACAA CTTAATATAC CTATTATGG TGGTCCTTTA GCATTAGGTT TAATCCGTAA	60
	TAAACTTGAA GAACATCATT TATTACGTAC TGCTAAACTA AATGAAATCA ATGAGGACAG	120
	TGTGATTAAA TCTAAGCACT TTACGATTTT TTTCTACTTA ACTACACATA GTATTCTCTGA	180
15	AACTTATGGC GTCATCGTAG ATACACCTGA AGGAAAAGTA GTTCATACCG GTGACTTTAA	240
	ATTTGATTTT ACACCTGTAG GCAAACCAGC AAACATTGCT AAAATGGCTC AATTAGGCGA	300
20	AGAAGGCGTT CTATGTTTAC TTTCAGACTC AACAAATTCA CTTGTGCCTG ATTTTACTTT	360
	AAGCGAACGT GAAGTTGGTC AAAACGTAGA TAAGATCTTC CGTAATTGTA AAGGTCGTAT	420
	TATATTGCT ACCTTCGCTT CTAATATTTA CCGAGTTCAA CAAGCAGTTG AAGCTGCTAT	480
25	CAAAAATAAC CGTAAATTG TTACGTTCCG TCGTTCGATG GAAAACAATA TTAAAATAGG	540
	TATGGAACCT GGTATATTA AAGCACCACC TGAAACATTT ATTGAACCTA ATAAAATTAA	600
	TACCGTACCG AAGCATGAGT TATTGATACT ATGTACTGGT TCACAAGGTG AACCAATGGC	660
30	AGCATTATCT AGAATTGCTA ATGCTACTCA TAAGCAAATT AAAATTATAC CTGAAGATAC	720
	CGTTGTATTT AGTTCATCAC CTATCCCAGG TAATACAAA AGTATTAACA GAACTATTAA	780
	TTCCTTGAT AAAGCTGGTG CAGATGTTAT CCATAGCAAG ATTTCTAACA TCCATACTTC	840
35	AGGGCATGGT TCTCAAGGTG ATCAACAATT AATGCTTCGA TTAATCAAGC CGAAATATTT	900
	CTTAECTATT CATGGTGAAT ACCGTATGTT AAAAGCACAT GGTGAGACTG GTGTTGAATG	960
40	CGGCGTTGAA GAAGATAATG TCTTCATCTT TGATATTGGA GATGTCTTAG CTTTAACACA	1020
	CGATTCAGCA CGTAAAGCTG GTCGCATTCC ATCTGGTAAT GTACTTGTTG ATGGTAGTGG	1080
	TATCGGTGAT ATCGGTAATG TTGTAATAAG AGACCGTAAG CTATTATCTG AAGAAGGTTT	1140
45	AGTTATCGTT GTTGTTAGTA TTGATTTTAA TACAAATAAA TTACTTTCTG GTCCAGACAT	1200
	TATTTCTCGA GGATTTGTAT ATATGAGGGA ATCAGGTCAA TTAATTTATG ATGCACAACG	1260
	CAAAATCAAA ACTGATGTTA TTAGTAAGTT AAATCAAAAT AAAGATATTC AATGGCATCA	1320
50	GATTAAATCT TCTATCATTG AAACATTACA ACCTTATTTA TTTGAAAAA CAGCTAGAAA	1380
	ACCAATGATT TTACCAGTCA TTATGAAGGT AAACGAACAA AAAGAATCAA ACAATAAATA	1440

	GCTTTTCTT TATATATGAT GAGCTTGAGA CATAAATCAA TGTTCATGTC TCTACAAAGT	1560
	TATATTGGCA GTAGTTGACT GAACGAAAAT GCGCTTGTA CAAGCTTTTT TCAATTCTAG	1620
5	TCAGGGGCCC CAACATAGAG AATTTGAAA AGAAATTCTA CAGGCAATGC GAGTTGGGGT	1680
	GTGGGCCCCA ACAAAGAGAA ATTGGATTCC CAATTTCTAC AGACAATGTA AGTTGGGGTG	1740
	GGACGACGAA ATAAATTTTG AGAAAATATC ATTTCTGTCC CACTCCCGAT TATCTCGTCG	1800
10	CAATATTTTT TTCAAAGCGA TTAAATCAT TATCATGTCC AATCATGATT AAAATATCAC	1860
	CTATTTCTAA ATTAATATTT GGATTTGGTG AAATGATGAA CTCTTGCCT CGTTTAATTG	1920
	CAATAATGTT AATTCCATAT TGTGCTCTTA TATCTAAATC AATGATAGAC TGCCCCGCCA	1980
15	TCTTTTCAGT TGCTTTCAAT TCTACAATAG AATGCTCGTC TGCCAACTCA AGATAATCAA	2040
	GTACACTTGC ACTCGCAACA TTATGCGCAA TACGTCTACC CATATCACGC TCAGGGTGCA	2100
20	CAACCGTATC TGCTCCAATT TTATTTAAAA TCTTTGCATG ATAATCATT TGTGCTTTAG	2160
	CAGTTACTTT TTTTACACCT AACTCTTTTA AAATTAAAGT CGTCAACGTA CTGATTGAA	2220
	TATTTTCACC AATTGCCACA ATGACATGAT CAAAGTTACG GATACCTAAA CTTTTCATAA	2280
25	CTGCTTCATC TGTAGTGTCT GCAACAACCG CATGAGTAGC GATATCACTA TATTCATTCA	2340
	CTCTATTTTC ATCATGGTCG ATGGCCATTA CATCCATGTC TAATGCATT CACTCACGAA	2400
	CGATACTACC TCCAAAACGA CCTAGACCGA TGAATACATA TTCTTTACCC ATACTCGCCC	2460
30	TCCATTAAAT GATTTTCATC AATTCATTGA AAATATAAAT TAAAATTAT TATAAATGAG	2520
	TACCCCAACT AAATTATCTA AATGCAGTAA TGCAAGTAAA TGAAAGTTGG GGTATCGTCT	2580
	CAACTTATGA TTTCTTCTCT TCAACATATT CTTTGTGCGA AACAAATAAT CTTAATAATA	2640
35	ATATTAACGA TGGAAGTAAT AAAAGTAAAC CTAAATAAAA GACAATCACT AATGTCCAGC	2700
	CCATTTCTGG ATTAACATAT GCATCTGTAA TTTTACAAA CGGATATAAA AGGTATGGCA	2760
	ATTTACTAAT TCCATAGCCA AAGAACGCGA ACATCATTTG TAAAATAACA AATACAAAAG	2820
40	CCTAACCATG TTTTCTCTTA AAGAATGTTA ACAATGAAGC TAATGCAAAG AATAAGAAAC	2880
	TTATACCAA CATCCACCAA TAGTCAAAAA CAGCTGAATA AAAATGTTCA GAATTTTGAA	2940
45	TGCGTAATGA TAGAAATACG AATAAACAAA TGATAATCAT CGGCGGCCCT AAAAATATGT	3000
	GCCATTGTCT TGTAAATTA TATGCTGGTT CGTCATTTGC TTTTGTAGCA TAATATGTCA	3060
	AAAATCCTGA TGAAATATAT AAACTGAAA TAATTGCCAA GAATACTACA GACCAAGCAA	3120
50	ATGGGCTTAA TAATAACTGC ACCCAATCTA GATCGATAAC ATTGTTTCGA ACATTAATAT	3180
	AGCCACCTTC TGTAATAGTT AAAGCAGTAG ATAATGAAGC TGGAATTAAT AATCCACTTA	3240

55

	AAC TGT TTT CT CAACGATATC ATAATCAGTG CTATTGAACC TGGTATTAAC AATACCGTGC	3360
	CTAAATATTT GATTGACTCT GGAAAGAAAC CTACGAATCC TACGAAGAAG AAAACAAAGA	3420
5	ATACATTTCGT AACTTCCCAA ACTGGGTTTA AATAACGTGA AATTAAAGTGA TTAATTTTCT	3480
	TTTCATCACC AGTTAACTTT GAATGCAATG CGAAGAAACC TGCCCCAAAA TCTATAGAAG	3540
	CAATAATGAT ATAGCAAAAT AAAACAACCC ATAACACTGT TATACCTATA AATGCATAAA	3600
10	TCATTTTTCT ATTTCTCCTC CTTGCTTCTT GGCTAAACGA TTTACATCTT CATACGCCGG	3660
	TTTATTTTTA AACATACGAA TTAATACGTA TGCACATGTA TACATTAAAA TGATGTACAA	3720
	TATGCCAAAT AAAATTGTAA CGAAGGTTAT TCCGCCTGCT TGTGTTGCTG CTTCTGCCAC	3780
15	GCGCATATAA CCACGAACAA TCCAAGGCTG TCTACCCATC TCTGTTAAGA ACCATCCAAA	3840
	TTCTATAGCT AGCATTGAAG CTGGGCCTGT TAATAATATT CCATAAAGCA TCCATTTATG	3900
20	AGTAGAAAAC TTTCTAAGCT TTTTAAACAT TAAAGTTAAG ACATAAACAC CTGAAATGAC	3960
	AAAACATAAA ATTCCCATCG TTACCATTAA ATCAAAGAAA TAATGGACGA TCATAGGCGG	4020
	ATGTAAACTT TTTGGAAAAT CATTTAACCC TTGTACTTTA GTTTTGACAC TATTATCTGC	4080
25	TAAGAACTC AATAGTCCAG GTAATTCAAT CGCACCTTTA ACTTGCTGAG TCTTTTCATC	4140
	TAACACACCA AATAATAATA ATTTGGCATG GGAAGATGTA TCGAAATGCC ATTCATAAGC	4200
	TGCTAATTTT TCAGGTTGGA ATTTATGCAA AAATTTTGCA GATAAATCCC CTGCCAACAT	4260
30	AGAAAGTAAT GTTGAAAAGA ATCCAATAT CATAGACATT TTCAAAGCTT TCTTATGGTA	4320
	GACAGTATCT TTAGGTTGAC GATTACGCAA TAATTTAAAA GCTGCTATTG ATGCAATAAC	4380
	AAATGCCATC GTCATACCGG CTGIAGTAAT TACGTGAAAT GATCGAACTA TAAACGATGG	4440
35	GTTAAACATC GCTTCTATAG GTTGAACATT GACCATCTTT CCATTCTTCA ACTCAAAACC	4500
	TGCAGGCGTA TTCATAAATG AATTCACCTGA AGTAATGAAG AATGCTGAGA AAGAGCCACC	4560
	AATAATTACT GGTATACTAA TTAAGAAATG TGTCCATTTA TTTTAAAC GATCCCAAGT	4620
40	ATATAAATAT ATACTTAAGA AAATAGCTTC AAAGAAGAAC GCAAATGTTT CCATAAATAA	4680
	TGGAAGTGCA ATAACGTGTC CACCCATTTT CATAAATGTA GGCCAAATCA ATGATAATTG	4740
45	AAGTCCTATA ATTGTACCTG TAACAACTCC CACTGCTACA GTAATTGTAT AAGCTTTAGC	4800
	CCATCTTTTG GCCATAGCTA TATATTGAAG ATCATTTTGG CGAATACCTA AAAATTCTGC	4860
	AATTGCGAAC ATTAAAGGCA TACCAACACC AATCGTTGCA AAAATGATAT GAACTGCTAA	4920
50	AGTCATAGCT GTCAAAAACC GACTGATTTT AACTGTATCC ATTAAAAAC ATCACCTTTT	4980
	TCTTTTTTTG ATGACAACAC AATGAACCTA ATTATAATTG CTATAATGTG TATTTTTAAA	5040

55

GAATTTCAAT GTATAATTGT GTATATTACA TTAGAATAAA GCACGAAGGA GCATGATACA 5160
 TGTCAGAAAT AATCGTTTAT ACGCAGAATG ATTGTCCACC TTGTACATTT GTAAAAAATT 5220
 5 ATCTAAATGA GCATCACATT GATTTTGAAG AGAGAAATAT CAACAATCAA CAATATCGAA 5280
 ACGAAATGAT AGATTTTGAT GCTTTTTCOA CTCGGTTTAT TTTGTTGAAT GGC 5333

(2) INFORMATION FOR SEQ ID NO: 171:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11126 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

ATACGTGACC CTTTATCCGA AAATTTCTTT TCATATTCTG TTAAAATATT ACTGCCATCG 60
 TCTTCTTGAT GTAAATTTAG ATTTATTTTT GTAAAATACA TTCCAAATTG AGACATACTT 120
 TCTAAACTGT AGGCAAATAG TCCTCTGTTA TCAGTTTTTA AATGTAAATC TCCTTCATCA 180
 25 TTTAAGATTT GTTGATACAA CGCTAAAAAC GTATGATACG TTAAACGTCG TTTTGATGA 240
 CGATTTTTTG GCCATGGATC TGAAAAGTTC AAATAAATAC GCGAAACTTC GCCGTCTTTA 300
 AAATATTCAT TTAATTCAAT GCGTCATTA CAAATAATCT TTAAATTGT TAAACCCATC 360
 30 TCTTTAACTT TATCCAATAC TTTATAAAGC ATACTTTTCT CACGTTCCAT TGAATATAG 420
 TTAATATGAG GATTTTGAGC AGCTAATGTT GTAATAAACT GCCCCATACC CGAACCAATT 480
 TCAATGTGTA TCGGTTGCGT TTTATCAAAC CATTCACTCA TTTTCCCTGc ATGTTGACCG 540
 35 TCCATGTCAA CCAATTCAGG ATGATCTTTT AAATAATCTT CAGCCCATGG TTTGTATCGA 600
 ACTCTCATAT TTTATTCTCC TCTTAAATAA ACATGTTACT ATTCATAACT TCATTTAGGA 660
 ATTTAAGCCA AGTGTTTATA TCCTTATATC TTTTTTGCTC TTCATACCAT TGAACAAGAC 720
 40 CTATAGATTG AATTACCGTA TACCATTTC ACGTTTTATT TAAATTCAAG CTCTCTTGAA 780
 CACCATATGT TTCAAGCCAT TCAGACCATT GTTGTGTGG AACATAGTTG TAAAGCAGCA 840
 TTCCGATATC AATTGCCGGG TCTGCAATCA TTGCACCTC CCAATCAACT AAAAATAGTT 900
 45 CATCTCGATC GGATAATAAC CAATTATTAT GATTCACATC ACCATGTACA ACAGTGAAAA 960
 AACGCGAATC TAAACTCGGT ATATGCTCTT CTAATAGGT TAATGATTTT CTCACAATAT 1020
 50 GATGTGTTAA AACTTCTCTT GATAAAGAGG CATTAAATTT ATTAAGCATA ATCTCAGGAG 1080
 TAATAGGTTT CATTTCATTA CGCTTTAACA TACTTAATAA AGGTCTAGAA TTGTGTATCT 1140

	TTTTCCAATG TTGTGCTGTA ACAACCTCGC CTGTTTCTAT GCGTTTCGTC CATACTAATT	1260
	TGGGCACAAT ACCTTCTGCT GATAATGCCG CAATAAATGG ATTTGAATTT CGTTTTAAAA	1320
5	ACAACCTTTTG TCCATCTTGT TCAGCCATAT ATGCTTCACC AGATGCACCA CCTGCTGAAT	1380
	CAAGTGTCCTA CCCTAATTGA TAAACTGCT CCAACTCGTC CACCTCACTT TCAATTAGAA	1440
	AATGGCTCTA GAAATAGGTT TTTCAAGAGC CATATATTCT AATTATATAAC ACCATACTGG	1500
10	TACAAATATT ATGTCCAGAT AATTATTGTA AATCCTCAAC CAATGCCTAC ATTACACGAC	1560
	TAAATTTAAA TCGTAATGTC TGTCAATTGAC ACCATACATT CTATAGTCAC TTACTTGACA	1620
15	TATAATGTTA CCGTGTCTAA AACTACATGT TTTTGAATCT CTGTAGGCGA TAAACTcTAG	1680
	TTTTCAAAAT AATTGCTATC CCATTTTCAT GGTTAGCATA AATTATATGAA CTGTAACATT	1740
	TACGTACTTA GTAAAATATG ATGCACATCA TATTTGTAC TCATAGAAAA TTTTATAAT	1800
20	TTTATCATTa TATTTCAACT GAAAATGAGA AACAAATGG CACTTTTTTAC TAATATGTGT	1860
	TTTCTAAACA ACACCTTTTAA GCTTCGTTTT AAATTATAAC ATAATTCATT TACGAAAGTT	1920
	GATAAATTTA AGTAATTTAA TCTAAAAATA TGATGAAAGA ATTTTAAATA CTGTGTGACT	1980
25	CTATATACTT TTCAAATCCT TCTTGTAGTT GACGTGTAAT TGGGCCAACT TTACCATCAT	2040
	TAACTGGTTC ACCATCTAAT TTAATAACAG GTGTAACCTC AGCTGAAGTA CTTGAAACAA	2100
	TAACTTCATC TCGTTTTTTC AAGAAATCTA CAGTAAACGT TTCTTCTTTA AATGGGATGT	2160
30	TATAGTCTTC GGCAATTTTT TTAATTACAA TTCGTGTAAT ACCATTAGA ATATAGTTGT	2220
	TAATCGGATG TGTATAAATC ACACCGTCTT TAATTGCATA AGCATTACTT GAAGATCCTT	2280
	CAGTTACAGT TcCACCTCGA TGTTGAATTG CTTCAACTGC ATTATATTTT ACAGCATATT	2340
35	CTTTTGCTAA TACATTcTCC TAATAAGTTC AAGCTTTTAA TGTCGCAACG TAACCATCGG	2400
	ATATcTTCAA CGGTAACACC ATTCACACCA TTTTCTAAAT GATCATAAGG ACGATCATAA	2460
	CTCTTTGTAT AAGCAACAAT TGCTGGTTCT ACTTCAGGTG TCGGGAAGCT ATGATTCCTT	2520
40	TCAGCTACAC CACGCGTTGC TyGAATATAA ATTGCCCCAG TTTCAATTTG ATTCATATCA	2580
	ACTAATTTAC GAGATAGTTC AATTAATTCT TCTACAGAAT AATTTAAATC TAAACCAATC	2640
45	TCATTGGCAC TACGTWAAA W TCTTTCATAA TGTTCTGTTA CTGTAAATAA CTTACCATTa	2700
	TATACTCGAA TGTATTCATA AATACCATCG CCAAATACGT ATCCTCTGTC GTTGTATGAA	2760
	ACCTTTGCTT CACTTGGA CT TACAACTCA CCATTTAAAA AAATTTTTTC CATATATTAT	2820
50	TCCTCCACGC ATAATGAATA AATTGCTTCT AAGTAAATAC TAGTTGCGTT AAATAACTGT	2880
	TTTTTAGTGA TATATTCATT TTTCTGATGC ATTAAATCTT CAGAATCACT AAACATTGCG	2940

55

	TCAGTCATAT	CATTGTTTG	ATTCTATAT	GCAGTAACTA	ACTTTTGAC	AAAAGGATCA	3060
	TTTTTATCAA	CATAATGTGG	TGGTTGGACT	TTACCTAATT	TCACCTCAAA	GCCATATTGT	3120
5	TGAATCTCAT	TTGCAAAACG	ATCCATAGCT	TTTTCAAATT	CAAATCCTTC	TGGGTAGCGT	3180
	AAGTTGATAC	CGAAAAGACC	TGCGTTTTCA	TTATCATATG	TAATAACACC	AATGTTAGTT	3240
	GTCACGTCAC	CCATGACATC	TGTATGGAAT	TTCATTCCCA	TCTTTTCACC	AAAATCTGAA	3300
10	TTAAATAAGT	AGCGATTACT	AAATGCTACA	AACGCTTG TG	CATTATTATC	AAGATTTAAT	3360
	GATGCTAAGA	ATTTTAGTAA	GTAAGACCC	GCATTCACAC	CGATAGATGG	ATCCATACCA	3420
	TGAACCGCTT	TACCTTCAAC	TGTTAAAACT	AGAATGCCAC	TATCAACAGT	ACTATCACCT	3480
15	TGTAAATGAT	TTTGTCTTAA	AAAGTACTCA	AAGTCTTGAA	TAACATCTGT	CATATTTTCT	3540
	TTAACAAGCA	CTCTTGCTTC	TGCATGATCA	GGTACCATGT	TGTAACGTTT	ACCAGATTTA	3600
20	AAAGTTATTA	ATTCATAATC	AGGTTTCATCT	TGATCTTCAG	TAAGTTTATT	TTGAACTAAA	3660
	TCAAATGTTG	TAATGCCTTT	TTCAACATGA	ATACATGGAA	ATTCTGCATC	TGGTGCAAAA	3720
	CCTAATGTTG	GCATTTCCTC	TGTTTTAAAA	TAGCGATCCG	TACATTTCCA	ATCAGATTCT	3780
25	TCATCCGTAC	CAATAATCAT	ATGAATACGT	TTCTTCCAAT	CCACATTCAT	ATCTTCTAAT	3840
	ATCTTAATTG	CATAATAAGC	AGCAATTGTT	GGACCTTTGT	CATCAAGTGT	ACCTCTAGCT	3900
	ATGATAGCAT	CTTCTGTTAC	AACCGGCTCG	AACGGATTAC	TATCCCATCC	ATCACCAGCA	3960
30	GGAACAACGT	CAACATGACA	TAAGATACCT	AATACGTCAT	TTCTTTTACC	TGCCTCAATT	4020
	CTTCCTGCAA	TATGATCCAC	ATCATGTGTT	GTAAATCCAT	CTCTATGTGC	AATTTTCATC	4080
	ATGTAGTCTA	ATGCCTTACG	AGGACCTGGA	CCAACTGGTG	CGTCTTCTGA	TGCTTTTGCA	4140
35	TCATCTCTCA	CACTTTCAAT	TGCTAATAAT	CCTTTTAAGT	CATTAATGAT	TTGATCTTCG	4200
	TATTGTTGAA	CTTTTTCTTT	CCACATTCGA	AATCGACTTC	CTTTTTTCTA	TAAGTTAAAT	4260
	TCTATTTTAC	ATGAAAAGAT	ATAAAAACTA	CAATAAGATG	TCAGAAAATA	ATAAAAAGGA	4320
40	ACAAAACGAT	GCTATTGATA	TGACACAAAT	CATAAATAGC	TGCTTTGTTC	CTTTTTTAAT	4380
	TTATATATTT	AAAATACACA	TATTCAAGAG	CTCGAGATAT	AAGTCAATGT	ACTAGGCACA	4440
45	CAATTTAATA	TTGACAGTAA	TTAACCGAAC	GAAAATGCGC	CCCGGGGCCC	CAACATAGAG	4500
	AATTTGAAA	AGAAATTCTA	CAGACAATGC	AAGTTGGCGG	GGCCCCAACA	TAGAAGCTGG	4560
	CCAATAGTTA	GCTTTCAATA	ATGTGCAAGT	TGGGGTAAGG	GCCCCAACAC	AGAAGCTGGC	4620
50	CAATAGTCAG	CTTTCATAAA	TGTGCAAGTT	GGGGTAAGGG	CCCCAACACA	GAGAATTTTCG	4680
	AAAAGAAATT	CTACAGACAA	TGCAAGTTGG	CGGGGCCCCA	ACACAGAAGC	TGGCCAATAG	4740

55

	TAAAGAAATA CGTTTTCTTT AGATATTAGT ATTTCTTATG AATGAGTTTC ACGCATGTAT	4860
	TCTTCTTTCT ATATGCATAT TAGCTATGAC TAACGATAAA GAACCTGAAA CACTAATAAA	4920
5	TGTCCTATAG TTTACAATAT TATATTGGCA GTAGTTGACT GAATGAAAAT ACGCTTGTA	4980
	CAAGCTTTTT TCAATTCTAG TCAACCTTGC CGGGGTGGGA CGACGAAATA AATTTTGCTA	5040
	AAATATGATT TCTGTCCCAC TCCCTTATCA TTTCTGTCCT ACTCACATCT TATTCTTTAT	5100
10	CAGATAATGC ATTTTTATTG TTTTTTAAAT CTTCTTCAGT GACGATACGT AAATTATTAT	5160
	TTGGTGTGCG CCACCTTCAT CATCAAATTT ACCTTTTTCA ATACTTTCGT CAGTCTTATT	5220
	GTCATATTCG GTAAATTTTG ATTTTCTTTC TTTGAAAAAT GCTTTTGGAT TATTTTTTAA	5280
15	TCTATTAGCA TATTCTTTCG GATTTGTTTT TACTTCTTTA ATTGTTTCAT TAGCAATTGT	5340
	TCCTAATTGC GTCGCTTTAT CCTTAGCATT ATCTTTATAG CTTTGAGGAT CTTGTTTATA	5400
20	TTTATTATAT TCcTGcTTTC AGCTTGTCAC GACTATCTTT ACGTGTAACA AGTACAGCTG	5460
	CTACAGCGCC ACCTATACCT AAAATCGCTT TAAATAAATT ACCTTTTGCC ATATCAATCG	5520
	TCTCCCTTTT ATTTATAATT TAATTTGTCA AAATCATTTT CAGTTAATAA ACGATATTCT	5580
25	CCTGAATCTA AATTGCTGTC CAATTCTAAA TCAGCAATTT TGATACGTCT TAAATGTAAT	5640
	ACCTCATTTT GAATGCTATG AAACATTTCGT TTAACCTGAT GATATTTTCC TTCATAAATT	5700
	GTTACGTGTG ACGTTTGATT ATCAATATAA GTTAATATTG CAGGCTTAAC CTTGCCATCA	5760
30	GTCAGTGTtA CACCCTCTTT AAAAGCTTGA ATGTCGTCTT CAGTGATAGG ATTTGCTGAA	5820
	ATAACTTCAT ATTTTTTAGA AACATGTTTG TTTGGACTCA TTAATTCATG ATTAATAATCA	5880
	CCATCATTCG TTATCAATAA AAGCCCTTCT GTATCTTTAT CAAGACGACC AACCGGAAAA	5940
35	ATATTTAGAT GTTGGTATTC AGGTATTAAA TCAATAACGG TTTTGAATG ATGATCTTCA	6000
	GTTGCTGATA TATAACCTTT TGGCTTATTT AACATAATAT AGACATTTTC AATGTATTCT	6060
40	ATTAATTCTC CACGAACGT TATCTTATCG TTTTCTGGTT CTATATGTGT TTTTGGTGAT	6120
	TTAATTACTT GTTCGTTGAC ATTTACAAGG CCTTTTTTAA GTAACGTTT GACCTCATT	6180
	CGTGTAACGA CGCCCATATT TGCTAAAAAT TTATCTATTC TCATCGTAAA AACCTAACTC	6240
45	TACGTCTTAA TTTTTCAGGA ATTTACCTA AGAATTCGTC CGCAAGACGC GTTTTAATTG	6300
	TGATTGTACC GTAAATTAGA ATACCTACTG TAACACCTAA AATAATAATG ATTAAGTAAC	6360
	CAAGTTTAGT AGGTTCTAAG AATAGATTTG CAAGGAAAAA TACTAATTCT ACACCTAGCA	6420
50	TCATAATAAA TGAATACAAG AATATTTTTG CAAAATGAAT CCAACTATAG CTGAATTTAA	6480
	ACTTCGCATA TTTTTTAAGA ATATAGAAAT TACATCCAAT TGCAAATAAT AATGCGATAC	6540

55

	ACTTGATAAC TACAGAAGCT AAAATAACAT AAAGTGTAA TTTCTGTTA TCTATACCTT	6660
	GTAACATTGA TGCCGTTACA CTTAATAGTG AAATTAGTAT TGCTACAGGC GCATAATAGA	6720
5	ATAATAAGCG ACTACCATCA TGGTTAGGGT CATGACCTAA AACAATTGGA TCGTAACCAT	6780
	AGAAAAGTGT GAATAATGGT TGTGCCAAGG CCATAATTCC AATACTAGCT GGAACAGTTA	6840
	TAAACATTAA TACACCAATA GATGTTCTAA TTTGATGATG CATTTCATGT AAGCGACCTT	6900
10	CTGCAAATGT TTTTGTAATA TAAGGAATTA AACTCACTGC AAAACCAGCA CTTAATGATG	6960
	TCGGAATCAT TACAATTTTA TTAGTTGACA TATTTAGCAT ATTAAAGAAT ATATCTTGTA	7020
	ACTGTGAAGG TATACCAACT AAAGATAAAG CACCGTTATG TGTAAATTGA TCTACTAAGT	7080
15	TAAATAATGG ATAATTCAAA CTTACAATAA CGAACGGTAT ACTATAAGCA ATAATTTCTT	7140
	TATACATCTT GCCATATGAC ACATCTATAT CTGTGTAATC AGATTCGACC ATACGATCAA	7200
20	TATTATGCTT ACGCTTTCTC CAGTAATACC AGAGTGTGaa TATACCAATA ATCGCACCAA	7260
	CTGCTGCTGC AAAAGTAGCA ATACCATTGG CTAATAAAAT AGAGCCATCA AAGACATTTA	7320
	GTACTAAATA ACTTCCGATT AATATGAAAA TCACGCGTGC AATTGCTCA GTTACTTCTG	7380
25	ACACTGCTGT TGGCCCCATA GATTTATAAC CTTGGAATAT CCCTCTCCAT GTCGCTAATA	7440
	CAGGAATAAA GATAACAACC ATACTAATGA TTCTTATAAT CCAAGTAATA TCATCGACTG	7500
	ACCAACCGTT TTTATCATGA ATGTTTCTAG CTAATGTAA TTCAGAAATA TAAGGTGCTA	7560
30	AGAAATACAG TACCAAGAAA CCTAAACAC CGGTAATACT CATTACAATA AAAGTCGATT	7620
	TATAAAATTT CTGACTTACT TTATATGCCC CAATAGCATT ATATTTGCA ACATATTTG	7680
	AAGCTGCTAA TGGTACACCT GCTGTCGCAA CTGCAATTGC AATATTATAT GGTGCATAAG	7740
35	CGTATGTGAA CGGCGCCATA TTTTCTTGTC CACCAATTAA ATAGTTGAAT GGAATGATaa	7800
	AAAGTACGCC CAATACCTTG GTAATTAATA TACTAATGGT AATTAAAAAG GTTCCACGCA	7860
	CCATTTCTTT ACTTTCACTC ATTACGAATC TCCCTATCTC ATGTTTATTA AAGTTTTGTA	7920
40	AACTAAAAGC TGTTTCTCTG TAAAATCATT TTTCATTATT ATGAATATAT CACAAAACCT	7980
	TATTTTATTG TCGTATATTC AATGAATTAT CATAACAAAA TTATCAACAC ATTGTCATTG	8040
45	AATACTAGAT TTTGATTAGA ATATTACGAA ATTTATATA AACATTATAC TACTATTTGA	8100
	GATGAACATC GCATAACAGT AGAAAAATCA TTCTTATCAT ACACATACAT CTTCATTTTT	8160
	TATGAAGTTC ACATTATAAA TATATTCAAC ATAATTGTCA TCTCATAACA CAAGAGATAT	8220
50	AGCAAAGTTT AAAAAAGTAC TATAAATAG CAATTGAATG TCCAGTAACA AATTGGAGG	8280
	AAGCGTATAT GTATCAAACA ATTATTATCG GAGGCGGACC TAGCGGCTTA ATGGCGGCAG	8340

55

EP 0 786 519 A2

	GTAAACTCAA AATATCTGGT GGCGGTAGAT GTAACGTAAC TAATCGATTA CCATATGCTG	8460
	AAATTATTAA GAACATTCCT GGaAATGGGA AATTTTTATA TAGTCCCTTT TCAATTTTTG	8520
5	ATAATGAATC CATCATAGAT TTTTTTGAGT CTAGGGGTGT TAAATTAAAA GAAGAAGATC	8580
	ACGGGCGTAT GTTTCCAGTT TCCAACAAAG CACAAGACGT GGTGATACA TTAGTGACAA	8640
	CTATCGAACG CCAACATGTA ACGATTAAAG AAGAAGAAGC TGTTAGTAGA ATCGAAGTTA	8700
10	ATACAGACCA AACTTTCAC TACATACTC AAAATAATAG TTATGAAAGC CATTGCTAG	8760
	TGATTGCTAC AGGTGGTACA AGTGTCCCTC AAAGTGGTTC AACTGGTGAT GGTATAAGT	8820
	TCGCACAAGA TTTAGGTCAT ACCATTACTG AGTTATTCCT GACCGAAGTT CCAATTACAT	8880
15	CAGCTGAACC TTTCATCAAA TCCAATCGTC TAAAAGGTTT AAGTTTAAAA GATGTTGAAT	8940
	TGTCAGTACT TAAGAAAAAT GGTAAAAAAC GCATCAGTCA TCAATGGAT ATGTTATTTA	9000
20	CTCATTTTGG TATCAGTGGT CCAGCTGCAT TAAGATGTAG TCAGTTTGTT TATAAAGAAC	9060
	AAAAAATCA AAAGACACAG CACATTCTA TGGCAATCGA TGCAATTCCT GAATTAAACC	9120
	ATGAACAATT AAAACAACAC ATCACATCAT TATTATCGGA CACACCAGAT AAAATCATT	9180
25	AAAACAGTTT GCATGGTCTA ATTGAAGAGC GCTACTTACT GTTCATGCTG GAACAAGCAG	9240
	GAATCGATGA AAATACCACA TCACATCACT TATCAATCA ACAATTGAAC GACTTAGTAA	9300
	ATATGTTTAA AGGGTTTGTA TTAAAGGTGA ACGGGACATT ACCTATAGAT AAGGCATTTG	9360
30	TCACAGGTGG TGGTGTGTCA CTTAAAGAAA TTCAACCTAA AACAATGATG TCTAAATTAG	9420
	TTCCGGGATT ATTTTATGT GGTGAAGTAT TAGATATACA TGGTTATACT GGTGGTTATA	9480
	ATATTACAAG TGCACTCGTA ACAGGACATG TCCTGGATT ATATGCCGGA CATTACTCAC	9540
35	ATGCATCAAT GGAATAATAG TATAAAATTT GGTTCGATTC TCTTAGTAG ATCAACTTTT	9600
	TCATTCAAAT AAAATGACC TTAATATAAC TGAGTCACTA AAAAGTGTG TTATATTAAG	9660
	GTCAATTCGT TAATTATGAT TCTTTTCGT TTTTAGTACG TCTTCTAGCT AACAAAGCCG	9720
40	CACCTGTAAT CAGTGCAAT TCTTCAATG GTAAATCCAT TCCTTCAGAA CCTGTATTTG	9780
	GAAGTCTTT TTCAACTTTG CGCGATTCAT GTGTCTCTTC TTTTAAATA GGCGTACAAA	9840
45	CTTTGGAGC TGGCTGAATT TCTTTGGTG ATACTTCGT CGCTTCAGCT GGTAATTTAA	9900
	TTGCTAAAT TTCATCAACA ATGAATTGCG TGTGTGTTT GATGTCATTT AATGTCGCAT	9960
	CTTCATCAAT CATTCTATTG CCATCTGCAA CATATTGATC AATTAATACT TTTACTTTAG	10020
50	CTAATTGTTT TGGTGTGCG ATCGCTTTGA ATTTGCGATA TGTTTGTTGA GCAATGTTAT	10080
	CAATTCGCAG TAAGCTATTT TCTTTTTCAG TAATTACTGC TTCTATATCG CTTAATGCAA	10140

55

CATCCATTG TAATTTTAAA GCAGTTATAG CTTTTAATGC ATCAGCCTTA TTACGATTAC 10260
 TTACTTTTCG ATAATTTTGC ACTAAAGCAG TGACGCGTGC AAGATCATCA TTAATCGTTT 10320
 5 TTTCAGCATC TGGCTTTTTA ATAGGATGTA CATCTAAATC ATGTATTGTT TGTAGATTTA 10380
 ATGATGCTGT TTTATCAACT TGTGCATTGC TACGATCTTG ATCAATTTGT CCAATAGCAG 10440
 TGTCAAATAT ATTTTGTAAAC TGTGCTAATA TACTATTTCT TTCTTCTACC GTTGCTTGAA 10500
 10 TATTCGCTTC AATTGCTTGT TTTTATCGT TGAATAATGT TGTCAATTGT TCTCGAGCAG 10560
 ACGCCTTTCT GTTAATAACA GGTTCGATTT CACGAATTTT GTTTTCTCA TCATGCAATA 10620
 AATATGCCAC ATCTGCATTA GTCAGTGCAC TAGCAATTTG TTGTTTAGCT TTAATTAAC 10680
 15 CTTTTTCAAC TTGTGCTATT GCAATATTTT GTTCTTCATC TGTCGCTTCG TTATTTGCTT 10740
 TAATTAAATT AATTTTATTT GTAGCGATAT TTTGAATTTG TTGTAATGCT GTTGCTTTAA 10800
 CTGTTGTGCG TGGTTTAATT TTTGAAATAA TATTTGAGC ATTATACTA TCTTGATTAA 10860
 20 CTTGGGCAGT CTTATCTGCA TGATTGATCT GATCAATAGC CTGATTAAGT GCTTGTTCTA 10920
 CTAAATGTTT AGCAGCTAGT CTTTCTTCTT CAGTTGATAA ATCGCTTTGA TCGATTAGTG 10980
 25 CATTTTGAGC TTCGGCTTTT ACACCAACAG ATTGACGCGC TGCTGGTTTA ACTTGAAC 11040
 TAGGTAAAAT CACTTTGATG TTGTCGTTGC CATCAGTCnC AGTnCGATCC ACTTCTGCAT 11100
 TCGTTTTGTT TTGTGCAATG TCATT 11126

(2) INFORMATION FOR SEQ ID NO: 172:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3660 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

TTGCCCCGCA CGGCGGTGTG nTTCCTAGAA ATAATGAATA TAAAGaGAAA TATATAACAA 60
 CGATTTTGAA TTATGAACCT GGTGATATCG TTACAATCAA ACGTGTGAGA GATAAGACCG 120
 45 ATTTGCTAAT ATATTGTCT AGTAAAGATA TTTCTATTGG TAATGAAGTG GAAATTGTAT 180
 CGAAAGATGA AATGAATAAA GTAATTATCA TTAAACGTAA TGATAATGTA ATTATTGTCA 240
 GTTACGAAAA TGCAATGAAC ATGTTTGCTG AAAAAATAAA TAAAGAAGCC ATAAAGATAT 300
 50 CCATGATTGA ACTGATAAAG ACATATGGAT AATTGCTTTA GGCTTCTTTT TTATTAGTTA 360
 ATTTATCAAG TGAGTATATT TGAGTAAAAT ATTCACTGCA TAAAGATTGA AGATAATCCA 420

	CTGTGGACTC GGACGCTGGA AAGTCAATTT AGCAATCGTC CAACTAGATT GTAGAACTTC	540
	GCCTAATAAT ACACCTAAAA TATATTGATA ACTCATTGTG ACAAGTAGTT GAATTTCTAC	600
5	TATATTTTCA TCTTTTAATA TAAAATACAA CATGATAGAA ATTAAAGTTA TAACAACAAT	660
	GGGTGAGCCT TTTCTAGATG TTAAATTAATA AAAATAAATA AATATCAATA AATAGGTAAA	720
	TATAAGAAAA CTAGGTATCT GATAATGGCT CGACGCTAAA CCTATCAATA ACATAATAGG	780
10	TGGCATAAAA TAACCACCAA TCGTTGTAAG CCATTGGCCT GCTAGATGTC TAGATTGTGT	840
	AATTGCGAAT CCTTGTGTGA ATGTCTGTTG TCGCTCTCGT GGACTTGTTA CAATGACTAA	900
	ATCTTTTGCA CGGCCACCAG CGAGTTTATT AAACAGTACA TGACCAAATT CATGTGTTAA	960
15	AACAGGGATA TAGTTTAAAA TGACATCTAA ATAGTTCAAA ACAGGCTTAT GTCTATATTG	1020
	ATGAATAGCA ATATAACAAG CTGCAACAAT AACGATAATG TATATATTAA GTTGAATTGT	1080
20	CGTATTAAAA AAGTTTGATA AATAATTCAT TGTTAACCTC ATATAAGATA TTAATTTAAA	1140
	GTTTGCTTAT CACTTATTAT AAATGATATT GGCATCAATA GCGTTAGACT TTAGACTTAC	1200
	CTTAGTTAAA CTAATTTTAA TTTTGA AAAA GGTGAATATG TGTAAAATA AAGCAAAATC	1260
25	ATTTGATAT AAATAGGATG AATATAAATA CTGTTAATAT TGATTACACT AACATAATAA	1320
	TGAAATAAGA TAGGAGATTC CTGTTATGAC TGTGAAGAA AGATCCAATA CAGCCAAAGT	1380
	TGACATTTTA GGGGTCGATT TTGATAATAC AACATGTTG CAAATGGTTG AAAATATTAA	1440
30	AACCTTTTTT GCAAATCAAT CAACGAATAA TCTTTTATA GTAACAGCCA ACCCTGAAAT	1500
	AGTGAATTAC GCGACGACAC ATCAAGCGTA TTTAGAGTTA ATAAATCAAG CGAGCTATAT	1560
	TGTTGCTGAT GGGACAGGAG TAGTCAAAGC TTCGCATCGT TTAAAGCAAC CTCTAGCGCA	1620
35	TCGTATACCT GGTATTGAGT TGATGGATGA ATGTTTGAAA ATTGCTCATG TAAATCATCA	1680
	AAAAGTATTT TTGCTAGGGG CAACTAATGA AGTTGTAGAA GCGGCACAAT ATGCATTGCA	1740
40	ACAAAGATAT CCAAACATAT CGTTTGACACA TCATCACGGT TATATTGATT TAGAAGATGA	1800
	GACAGTAGTG AACGhAnTTA AACTGTTTAA ACCTGATTAC ATATTTGTAG GTATGGGATT	1860
	CCCTAAACAA GAAGAATGGA TTATGACACA TGAAAACCAA TTTGAATCTA CAGTGATGAT	1920
45	GGGCGTAGGT GGTTCCTTG AAGTATTTGC TGGGGCTAAA AAGAGAGCGC CTTATATCTT	1980
	TAGAAAATTA AACATTGAAT GGATATATAG AGCATTAAATA GATTGGAAAC GTATTGGTAG	2040
	ATTAAAGAGT ATTCCAATAT TTATGTATAA AATAGCCAAA GCaAAAAGAA AAATAAAAAA	2100
50	GGCGAAATAA TCATGATGAC AAAAATAAAA CCGAGGAAAT CCTTAAATGG AGATTCTCGG	2160
	TTTTTTCGGT TTATTTAATA ACGAAGCGGG ACTCATCGAG TTTGTTTCTA AATTCTTTTT	2220

55

CATCAAGTTC ACCGTAATCT TTAACTTTC CGCCTTCAAT CCAAGCAATC TTAGTACAAA 2340
 ATTGTCTCAC TTGTCTAAG TTATGACTAA CGAAAAAGAT GGTTTGTGTT TGCTCTTTAA 2400
 5 ACTCGTAAAT TTTATCTAAA CATTITTTGTG CAAAAGTTTG GTCACCTACA GATAAAGCTT 2460
 CGTCAATGAC TAAGATATCT GGATTAAGTG TGATATTAAT TGAAAAACCA AGTTTTCGAC 2520
 GCATACCACT TGAATACTTT TTAAGTGGTT GATAAATAAA CTCACCAAGT TCACTAAATT 2580
 10 CAATAATCTT AGGTGTCATC GCTTTAATTT CTITTCGCTT AAAGCCCATC CATAACATTT 2640
 TAAATTCGAT ATTTTCAATC CCTGTAAGTT GTCCACTCAA GCCAGCACTA ATTGCGATAA 2700
 CGCTGACTTC ACCATTACGA TCCACTTTGC CAACAGTAGG CGACAAAGAA CCGCCAATGA 2760
 15 TATTGCTCAA CGTTGATTG CCGGAACCAT TGATGCCAAC AAGCCCTATG ACGTCACCTT 2820
 CATATGCTTT TAACTAATG TCATCTAAAG CGAAAAATGT TTTGTTTTTA TGTTCGGGAA 2880
 20 TGAGCGCATC TTTCATACGT TCTTTATTTG TACGATAAAT ACGATATTCT TTTGTTACAT 2940
 TTTTAATGTT TACCGAAACG TTCATTGTGA GACCTTCCTT ATTCACATTT ATCTAGATTA 3000
 TAATATACTA CTCAACAGTT GTTAAATTTT AAAACCTGTT GTAAAGTGTA TAGAAGATTT 3060
 25 TGTTATTATC AGAGTGGGTG TTTTGACACA AAATGTTAAT CATCAATGAT AACAAATGATA 3120
 TTTAAAACT AAACTTATTT CAACTTACAT GATTGTATAC TATAATGTAT TTGTAATAAA 3180
 CTAATATTTT AAAGAACTAG ACAATAATTT TGATAGCATC CATGTATAGT GATAGTATTT 3240
 30 ACAACAATTA TTATAATACT ATTTAGTTAA GTAGAGAAAT AGTTAAACAT TTGAAAGTGT 3300
 GGTTTAATGG AATGTCAGCA ATAGGAACAG TTTTAAAGA ACATGTAAAG AACTTTTATT 3360
 TAATTCAAAG ACTGGCTCAG TTTCAAGTTA AAATTATCAA TCATAGTAAC TATTTAGGTG 3420
 35 TGGCTTGGGA ATTAATTAAC CCTGTTATGC AAATTATGGT TTAGTGATG GTTTTGGAT 3480
 TAGGAATAAG AAGTAATGCA CCAATTCATG GTGTACCTTT TGTTTATTGG TTATTGGTTG 3540
 GTATCAGTAT GTGGTCTTC ATCAACCAAG GTATTTTAGA AGGTACTAAA GCAATTACAC 3600
 40 AAAAGTTTAA TCAAGTATCG AAAATGAACT TCCCGTTATC GATATACCGA CATATATTGT 3660

(2) INFORMATION FOR SEQ ID NO: 173:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 13868 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

55

	ATTAATCACT TGTGTGTAG AGTCTTGTC GTTTTGGTTA TGATTGTTAG CCATGATATA	120
	CCTCCCTTAC AACACTCGTG GACCAGAAGT TTTCTGATCT CTCACATTAA CTTCTAACTT	180
5	ACGTACTGGC ATTTCTGTGA AATATTCTAC ATTCTTTTAA ATATCCGAAC GAATTGCTTC	240
	AGTTAAAGAT TGAACCTGAA CATTATTTGG TACGAAAAAG TCAGTTTTAA TGTGATATA	300
	AGATTTATTT TTTTGTAT ATAGTTTCGC AACTACATT GGTGTCTTA CTTGATCATA	360
10	TTTTGCAACC GTATCGAATG CCGTCTTTTC AACAGCTTTA CGAGATACGT AAACATGACC	420
	ATCATCGAAG TCTTTGTATA ATCCAGGTTT TCGATGCGTA GGTGTGAAGA TACTAAATAC	480
	TAATATAAGA CCTATTAATA TCAATAGTGC AGCAAGTGAA ATAAGTAATG GTTGAACCA	540
15	TTCAAATTGA AGGAAGTAGT CTTGATATTC AGTTATACGT CCATCTTGGA TATACATGAA	600
	TAACAGGAAC CCCACGATTA CTACTATTAA TAAGCCAAGG ATAAAGTTTT TAAGTCGTTT	660
20	CACCCCTAAC GACACCTCCT TAGTTAAAGT TAATTTAAAA ACATATTAAA TATGTACCCA	720
	TCAGTTTTTT TCTTAAACAT AATAAATTAA TAACTTTAAA TTTATTTTTA ATATATAAGA	780
	TGAAGTACCA TTTAGTAATA TATCCCTAG TTTTGTAAA TAAAACCTCA TTATTAATTA	840
25	ATTyTCGTCA ATATGTTTTG AAGAACGATA TTCTAAAATA TCTGGGTCAC GATGTTAAT	900
	TAAAACCTTA TTACTATTC TCGGTTTCTC CTCACTCAA GATTTTATAA GCGACCATAT	960
	CATCGCTATA ATGACCACGG AAAATGGTAA CGCAGCAATG ATTAATAAAT TTTGAATTGC	1020
30	TTGAGTACCA CCTGTGTAAA TCATGATGAT TGCAAATAAT GCCATAATGA TACCCCAACT	1080
	CACTTTGACA AATGACTTCG GATTAATATC ACCACTTGAA CTCAACATAC CTAAAACATA	1140
	AGTTGCCGAA TCCGCTGATG TAACAAAGAA AATCATAATA ACAAGTAAAG TAATTAAGCT	1200
35	TAATACAAAA CCTAGCGGAT AATGTTGTAG CGTCGCAAAA GTTGCTGTTT CTGTCGCAGC	1260
	TTTAGCAATA TCGGCAATAT GATTATCTTG TAAGTAAATT GCTGACGCGC CGAATACCGC	1320
	AAAGAATATA AAGCAAATA ACGCCGGGAC AAAAAGTACA CCTAGAATAA ATTCTTTAAT	1380
40	CGTACGTCCT TTTGACACAC GTGCAATAAA TATACCTACA AATGGTGCCC AAGATATCCA	1440
	CCATGCCCAG TAAAAGATTG TCCAATTTG TAACCATTGG AATTTTGTAC CACCTGTCGG	1500
45	AATGCGTAAA CTCATACTAA AGAAATTTGC AATATAATTA CCTAGACCAT TCGTAAATGT	1560
	ATTTAAAATG TATAGCGTTG GCCCAACAAT AAAAAGACCA ATAAGTACTA CAAAAGCAAG	1620
	TAACATGTTG ATATTACTCA ACGTTTTGAT ACCTTTATCG ATACCTGACC ATGCTGACCA	1680
50	AGTAAATAAT ATGGTTGCAA TGACAATCAA GATTACTTGC ATCGTGAAGT TACTCGGTAC	1740
	ATTAATAAAA AAATGTAAAC CTTCGTTTAT TTGCAATGCA CCGAAACCTA ATGTTGCAGC	1800

55

	CATTGCCTTT TCACCTAATA AAGGCGTCAA TGTCAGCGCTG ACTAAGCCAG GATATCCTTT	1920
	ATGAAAGCTA AAATATGCAA AACTAGCGC GACAATACCA TAGACTGCCC ATGCATGAAT	1980
5	CCCCCAATGG AAAAATGAAA ACTGCATTGC ATCATTAAAT GCAGATTGCG TGCCAGCTTT	2040
	ATGAATAGGC GTTAATTGTA AGGCATGACT GATTGGTTCT GCCGTTGTCC AGAACACAAG	2100
	TCCTATTCCC ATACCAGCAC TAAATAACAT AGCAAACCA GACGGCAATG AGAATTCAGG	2160
10	ATCTTCGCCT TCTTCACCTA ATGTAATGTT TGCCTATCTC GAAAAATAAA TATACACACA	2220
	GACAAATAAA ATAATAAAA CGAGCAATAA ATAATACCA GAAAAATGTA GCGCAATAAA	2280
15	TGTAGTAATG TTTTGCCTGA GTTTTTCTAA CTGTTTCGGA AATATTGCTC CAAAAGCAAC	2340
	AAATATCGTA CATATCACTA AAGATACCCA AAACACTAGA CTTACTGATT TATTTTTCAT	2400
	AAATACAAAC CCTTTCTGTG TAATGGTAAG TTCATACCCA TAACTGCAAC ATTTTAATCA	2460
20	TTTGTAATTT TATATAGACA CAATTAATAA TGCCTCATCT TTTAAAAATG ATATATAAAA	2520
	CACACTCAAA TTATTTATCA TTGAGCAACA AAGTATTTTA TTGTATTTAA GTAATGCCTT	2580
	TCTAGTGCAT TATTGATTTG ATACCTGCAA AGTTGCCATA TTTCCGTTTA GAATCAATAG	2640
25	TCGCTAGACA CAAAAAATAA GTCGCCTATA CAGTATTTTC TGCATAAGGC GACTTTACTT	2700
	ACTAATCTAT ATATTAATTA CTAATTTTCC AATCATTGAT TGTTTTTCCA ACAATTGATG	2760
	TGCTTGATAT AAGTTTTTCTG GTGATAAACC TTCAAAACT TGTGTCGTTG TTGGTTGGTA	2820
30	ATGCCCTGAT TCTATATTTT TCCTAATATC TTCTAAATAC TCATGTTGTT TAATCATATC	2880
	AGGCGTTTGA TGAATTGGAC GCGCAACAT AAATTCATGT GTAAATGTTA TACTTTTTAA	2940
35	TTTTAATGCA TTTAAATCTT GATCTTCATT AAAAGCTACG ATAGTCGTAA TATGCCCTAA	3000
	TGGTTTTATC AGTTCAATCA TAGTATTGTA ATACAAGTCT GTATTATAGG TGCAAAATAT	3060
	ATAATCTACT AATGGAATTT CTTTAAATTG ACGCACTAAA TCCTCTTTAT GATTCAATAC	3120
40	GATATCTGCG CCCATCTTTT CACACCACTC TGTGTTTCT TGTCGTGATG CTGTTGTAAT	3180
	GACAGTTAAA CCATACCGTT TAGCAATTTG AGTGGCTATA CTGCCTACAC CACCGGCACC	3240
	ATTAATGATT AAGACAGACT TCCCTTCGTT TTCAGCAGGA TTCGTAGAAA TTTTAAATGT	3300
45	ATCAAAAAAC GTTTCATATG CCGTAATACC AGTTAGCGGT AGACTAACCG CTTCAATTAGC	3360
	ACTTATGTTG TGTGGTGCTT TTGCAACTAT AGCTTCTGAC ACCAATTGAT ATGTCGCATT	3420
	TGATCCTTGT CTATTTGGCG ATCCAGCATA AAATACAACG TCACCCGGAC TAAATAATGT	3480
50	AACGTCTGGT CCGATAGCTT CAACAGTACC AATAGCATCA AACCCAAGTA CACGAGGTGC	3540
	TTGAGTGACT TCCATTTGTC GTTGCTTTGT ATCTACAGGA TTTACACTAA TGCTATTTAC	3600

55

	ATTTCCTTCT	TCCAATTTAA	AGGGCTTCTC	AAATCCTATC	ATTTTCATAT	CGTTTCACCT	3720
	CATTTATGAA	CTTATTTCTT	ATTATACAAA	ATAGAAGCCA	TGTGTGCTTA	TATCGCAGCA	3780
5	TCATGACTCC	TTTTTCATTT	GAATATATAA	ATAATTACAG	ACGACTTTTCG	TATTAAATTT	3840
	TAGACTTATT	TCTACCATGT	TGCTGAACAA	ATTTACTTTA	GATAAAAAAT	TATTAAATTT	3900
	TGGTCAATT	ACAAAGTTAG	TTTGTTAAAA	CGTgATACTT	TATTATTCCG	TTACTTTAAT	3960
10	AACTTGTTTA	CCAAAGTTAT	CGCCAGTaa	TAAATTTTAA	AATGCATGTG	GCGCATTTTC	4020
	AAAACCATCT	TCAATGGTTA	CTTGTGACTG	AATTTTACCT	TCTTGAACCC	ATGTTGCAAG	4080
	CTGTTCACTA	GCTTCTTTAA	AAGCATTAGC	GAATTCACCT	ACCAAGAAGC	CTCTCATCAT	4140
15	TACTTGCTTC	TTAATAAGCG	TACCTTGAAT	ACGTGGTCCG	ATATCGGCTT	CAGGATGATT	4200
	ATATGACGAA	ATTGCGCCAC	ATACTGGTAC	ACGTGCAAAA	CGATTTAAAT	GCTTAAATAC	4260
20	TTCATCGCCA	ACTGTTCCAC	CAACATTTTC	AAAATAAACA	TCAATACCAT	CTGGTACTGC	4320
	TTGTGCTAAC	GCTTCTGCAA	AATCCTCTTT	CTTATAATCA	ATACCAGCGT	CAAAGCCCAG	4380
	TGTCTCTGTT	AAATAATTTA	CTTTTTTGTC	GCCACCCGCA	ATACCTACTA	CACGGCAACC	4440
25	TTTAATCTTA	GCAATTTGAC	CTACAACCTGA	ACCTACAGCA	CCAGATGCAG	CTGAAACCAC	4500
	AACAGTATCA	CCGGCTTTAG	GTTGTCCAAT	ATCAAGCAGA	CCATGATATG	CTGTTTGTCC	4560
	TGGCATTCCCT	AAAACACTTA	AATATAAATC	AAGTGGTACA	TCTGTCGTTG	GAACTTTAGT	4620
30	AATTTGATCC	GCTTGGACAT	GATTAATGAT	TCGCCAAGGC	AACATACCTA	CAACGACATC	4680
	TCCTTTTTTA	TAATCTGCGA	GTGTGCAATC	AATTACTTTT	GCAACGACAT	GGCTAACAAT	4740
	CGGTTTACCA	ATTTCAAAG	GCTGTACATA	CGAATCTGCC	TTAGTCATAC	GTCCTCTCAT	4800
35	ATATGGATCC	ACTGAAATAT	ACAGCGTTTG	TACAAGTACA	CCATCGCTCT	CAAGTTaGG	4860
	CGTGfCAATC	TCTTCaATTT	TGAATGTATC	CTCTTGAGGC	ATGCCkTCTG	GTATTTTGTT	4920
40	AAAAAGAATT	TGTTTATTTT	GCATCATTAA	TCACCTTTCT	TTATTTGAAA	CTTTTACTTA	4980
	TTTGTTACTT	AAGCGTTAAG	TTTGAATTGT	GTCtTCGTGA	TGTCTGTATG	CAAATACATT	5040
	CTTAGTTGTT	ATATTTTGAC	TTAAGCACTG	ATTCATTTCAT	GTAACCTCAA	CCACATTATA	5100
45	TTTGCTATAA	TCATAAATTT	AAAATGTTAC	GACTTAGACA	TTTTATGGAA	ACTCTCAAAC	5160
	AATAGATAAT	TTTTGAAAAG	CTCTAATATT	ACAAGCTTTT	TTGCCCCAGA	AAACTAGCA	5220
	GTTGCTTTAT	TTTTTCCATA	AGAAGTCGAT	TAACTCATTA	GCAACATTTT	CATTCTCATG	5280
50	AAGCTGACTA	TGTTGTGCAG	GCTCACCTTC	ATATTTAGAT	TCTCGATAAC	TTTTCGGACT	5340
	ATTTCCCAGT	AAATATTTTA	ATGATTTCGA	AGAACTATTA	GACACTCTGC	CGTCTGAATG	5400

55

	ATCTTTTAGC	ACGCGTAATT	GCTGATAAGG	TTGATTCATT	CGACTTGGTT	TACCATCTTT	5520
	ATCAACTGTA	ATTTTCATTGA	CATCTTCATT	CATATTTAAA	ACACCATTAA	ATGTCCCTGC	5580
5	AATATTCACT	TGTTTGTTTA	ACTGTGGCAG	TGACTTGTCTG	TTACCATATG	TCATCATATA	5640
	TTGTGCAAAT	GTAAAGTTCC	CCATTGAGTG	ACCGACAAAG	TTGAATTTAT	CGAAATTGTA	5700
	TTCAGATTGT	AACCTAGTCA	GTACATTTTT	AAACCACGCA	GCATTCTTAT	CCAAATAGCC	5760
10	TTGTCTGTTA	TTTTCAAGTT	CAATTTTCAC	AATAGGATTG	ACTGCATCTT	TTCTTAGTTT	5820
	CCCTTTAAAT	GTCACGTCAC	CATCCTTTGA	AACGTAAGCA	GTGATGATAT	CTTTAGTTAC	5880
	CCCTCTTTTT	TCTGCTTGCT	TCACCATAAA	CTTTTCAGAA	TTGGCACTAC	CACCAAATCC	5940
15	ATGTAAGAAC	AATGTTGGAA	TTGGCTTTTT	AACAAATTGC	TGTTGTTGTA	TTTTAAATGT	6000
	TTGTGCCTGT	CGTTGACTAA	ACACCACCAT	AATAATAGAG	CCTATAATAA	TAGCGACCGC	6060
20	TAACAATGTC	GTAATAATTA	CAAAAATTTT	CTTCACACTT	TTAACTCCCA	TTCATGTCTT	6120
	TTATATAAGT	ATAAAGGATG	TGATTAAAAA	TGTCCTTTAG	TTGATTTTGA	ATACATCATT	6180
	AACTTTAAAG	ATGACTTTGG	AAAGTTGTCC	GTTAACGTTT	GTTAATTGAT	TGCTTCTTTA	6240
25	GCTTTCATG	GTGTGTCACC	CATTGATTAA	TATATAAATA	TGTATATGCA	TGTTTAATTT	6300
	ATCTCTTTCT	ATAAATAAAG	ACCTACCAGC	ACTCGACTGA	TAGGCCTTTT	AATATCTATA	6360
	ATTATTTAAT	TTCTTTTGTT	TCGGCTAACT	CTTTGTACCA	ATAAGCACTT	TTCTTAGGAT	6420
30	AACGTTCTTG	AGTCTCAAAA	TCGACATAGA	ATAAACCATA	TCGTTTTTCA	TAACCATTTG	6480
	ACCAAGAGAA	CACATCCATT	AATGACCAAA	TAAAGTAACC	TTTAACATTT	GCACCATCTA	6540
	TAATAGCATC	TGCAATAACG	TTCAAATGTT	GTCTTACATA	ATCAATACGT	GCATCATCAT	6600
35	GAACTGTTTT	TTCAGATTCA	ATAAATTCAT	CTTTATATCC	TAAACCATTT	TCAGTGATAT	6660
	AAATCTTATG	aTAGTTAGGA	TAATCTTTAA	CAACACGCAT	GaTTTGATCA	TATAAACCTT	6720
	GAGGATAGAT	CATCCAGTCC	CAGTCTGTGC	GAGGTACGTC	GACATCAAAT	TCACGTTGTC	6780
40	CGACACCTTT	AAGTTGGTAT	TTAGAACCGC	CTTTATCACC	TGTCGCATTA	TGCGTGATTT	6840
	CAGATTCTCC	ATCGTAACCT	CTCATCCAAT	CACTCATGTA	GTAATTGATA	CCTAAGAAGT	6900
45	CGTTTAAATC	TTTGGCTGCA	TCTAAAATGG	CATAATCTTC	ATCTGTAATG	TTTAATTTAC	6960
	CGCCATTAAAC	AGATAAGATA	TGTTGCACAC	CTTCCATCGT	TTACAGAGAA	TACTTACCTA	7020
	AATATGTTGC	ATCTAAGATG	AATTTATTAT	GGATGATATC	TTCTAATTCT	GCTGCACGAA	7080
50	CATCTTCAGG	ATTTGATGGA	TCGAACGGAT	ATTTTGTTGG	CAATGCGTGT	ACAACACCAA	7140
	TTTCTCCTTT	GTATCCGCCA	TCTTTAAATA	ATTTTACTGC	TCTAGCATGA	GCCACCATCA	7200

55

	CTACTAAATA TTGACCATCA CCAATAGGTC CAATTTTCATT GAATGTAGTC CAATATTTTA	7320
	CTTCTGGGAA TTCTTTAAAA CAATATTCAG CATAATCTAC AAAGTAGTCA ATCGTTTTAC	7380
5	GATTTAGAAA ATCGCCATCT TTGTGTAAaCA CTTCTGGTGT ATCAAAATGA TGCAATGTTA	7440
	CAAATGGTTC AACATGACGT TTATGACACT CTGCAAATAA CTTATGGTAA TACTCAACAC	7500
	CTTTAGGGTT AACTTCGCCA TATCCATTTG GGAAGATACG AGACCATGCA ATTGAAATTC	7560
10	GGATACCATT AACACCGAAT TTTTCACCTA ATTCTAAATC CACTGGATAT CTGTTATAAA	7620
	AATCACTCGC TGGTTCTGCA GTGTACCAAT AGTTTTCTTC TAAATACGTA TCCCATGCTA	7680
	CGCGACCTTT ACCATCCGTA TTTGTGCGAC CTTCTGCTTG ATATGCTGCT GTTGCTCCAC	7740
15	CAAAAATAAA ATCTTCAGGT AATGTTTTAG TCATATGAAA AACTCCTATT CTTAATTTTC	7800
	AAATTGTTGT TGAACGAAAT CAAGGGCTGC TTGGCCATCT CGTGTCAATT TGATATATTC	7860
20	AGCACCTTGA GTCTTCGCTA ATTTAATACC TAATCTATCT GTATCTTGCT TAATATCTTC	7920
	ATAGTTAGAC GCAACTTGTG GCGCTAAAAT GATTAAATTGG TACTCTTTCA TAATGTCCAT	7980
	ATGTGCGCCA TATCCGCCAG CTGCGCTTT CACTGGCACA TGATATTCTT CAGCTGCTTT	8040
25	ATTAAGTGCA TTGGCTAATA ATCCACTTGT ACCACCACCG GCACAAAGTA CTAAGACATT	8100
	TGTTTGTTCT GTGATATTTG AAGCTTTAGC TGCATCGTCT GATACACCAC TTGCCGCTAA	8160
	AATTGAATCA GCTTTTTTCG TATCAAAGTT TGCTGCAACT TTTCTTTTA AATCTGAATT	8220
30	ACTTTCTTTA CGTCCTTCTT CTTTCATCAAG AATTTCACTA TCATAAACTT TTAGGAATGG	8280
	GTAGTAAATA ATAATATCTA CAACAATCAA AGTAATAGCT AGTACGAATG ACCATAAAACC	8340
	AAAACCTGTA CCCATGATAA TGCCCAATGG ACCTGGTGTG GTCCAAGGTA AATTCACACT	8400
35	AAAACCTGTA CCCATGATAA TGCCCAATGG ACCTGGTGTG GTCCAAGGTA AATTCACACT	8460
	TGCTAATACA AATGGAATAA AGAACACAGG ATTCAATACT AGTGGTGCAC CAAATAAAAT	8520
40	TGGTTCGTTT ACACCAAAGA ATGTTGGTAC AACTGATGCA CGTCCAATCG CTTTGTTCG	8580
	TTTAGATTTC GTCATCCACA TAAACATGAA CGGGACGACC AATGTTGCAC CCGTACCTCC	8640
	AAATGTAACG ATAAACATTT GTGTACCTGA TGTAATAATT TTATCTGCGT GTTCTCCAGC	8700
45	TTGAAGCAAC TTGAAGTTCG CTTTCGATATT CGCATATGTA ATGGCTGCAA TTGCTGGCTC	8760
	TACAATTGAC GGACCATGAA TACCTACAAA CCAGAATAAT GCAAAGGCAC CAAAGATAAT	8820
	TGTGACACCA ATCCATCCAT CTGCTGCTGT AAATAATGGT TCGAATAATT TTAAAATACC	8880
50	TTCCGCTACA TTTGATTAA AGCTGTTGCG AATGACTAAA TCTAATGCAT AAAGAATGAT	8940
	GATTACCGCT GAAAATGGAA TTAAGTCCTT AAATACTTGT GAAATATTCG GCGGTACTTC	9000

55

AAATGCTGAT AAGAATGCTG TTAATAAACC TTTAGTTCCG ATAAATGCAC TTAAGAATCC 9120
 ACCATCTTTG GCTGGATCAG AAGCTAAGAA CAAGAATCCA CACATCGCTG CTAGCATTGT 9180
 5 AGAAATAAAG TTAATTTGAT TTGTACTTTC TAGCTTACGG TTAAATGAAT CTGTTAAAGA 9240
 TTTGCTGTC GTTCCTGCTA CTAAAAATGC TACAAGCCCC ATCGTATAGT TATATGGTTT 9300
 CATTAAATG GCTTCCATGC CTTTATCCCA TTTAAACCA AATATATTTG GCACATATGC 9360
 10 AATTAATAGA AAGATACTTG AGAATAAGAT GACAGGCATT GCAGAAATAA ATCCATCACG 9420
 GATGGCTCTT AAATATATGT TACGTGATAA TTTCTCGAAA AATGGCTTCC CTTTTTCAAT 9480
 TTGTGCGATC AATTTTTGCA TCATTGTCAT CACCCTCTTT TATAAAATTC TAATAAATGC 9540
 15 TTCATTAAAT CTTTCAGTAA AATTGTTGTC ATTAAATGAT CTTGACCATG CATCATCGTT 9600
 ACACTGTATG CAATATCATC ACCTTGCGCT TCTTTAGCCA ACAGGCTTGT TTGTGCTCTA 9660
 20 TGCGCTTCCG CAATGCAATT GTTTCCTTCT TCAATCAGTG CATCTGCTTT TGCAAAATCT 9720
 CCAGCTTGAG CTGCTGTAA TGCTTCTAAA AACTTAGAAC GTGCATCCCC TGCAAAATGCA 9780
 ACAATTTCAA AACCTAATAA TTGGACTTCT TCTCTATTCA TAGCATTAAAT CCCCTTTTAA 9840
 25 ACTTATTTTC TTTGTTTCCA AGATGTCGCA GTATCTTTTA ATACTTTATT TAAGTCATCA 9900
 ATATTTTGA AACCAAGTGT ACGTAACCAT TCACGAGCAG CATCTTCACC TTGTTCAATG 9960
 TATACTTGAA CAGCACCAGA CCATGTAGCA CGGCCACAAA GTACCCCGTT GAATTTAGCA 10020
 30 CCAGCTTCGT GAGCAAATTT TAAAGTTTCT TGGATAAATT CCGCAGAAAC ACCAGCACTT 10080
 AAGTAAATGT ATGGTAAATG AGTTGCTGCA TCTTGATCTT TAAAGTGTG TGCCGCTTCC 10140
 TCTTTTGTAT AAACCACTTC ACCTTCAGCA AAGCCTTCTA CATATTTTCT GTTTACTGGT 10200
 35 ACTTCAACTT TCAATACATC AACGTTAAAG CGTGGTTCTG AGAATAATTT CATGCTTCG 10260
 TTAACCTTTC TAGGCTTAAC TTTTGCGAAT TcAACAcTAC CGTTATcAGG AATGTTGTCA 10320
 TCGTATGTTA ATACTTCTAA AAAGAATGGA ATATCTTCTG CAACACATTc TGAACCGATT 10380
 40 CTTTCAATGT ATGCTTTCTT TTGAATGTTA ATTTCTTCAG CATCATCAAC ATCATAGTAA 10440
 AGTAAGAATT TAACAGCATT TGCGCCTTGT TCTTTTAAAC GTTTTGCAGA CCACTCTACT 10500
 45 AAACAGTCAG GTAAACGACC TTTAGCGTTT ACGTCATATC CAGTTTTTTC ATAAGCAAGT 10560
 AATAATCCAC AATCTTTGTT ACGTGCATCT GAAGCTGGTA AACCATATTc AGGATCTAAT 10620
 AAAATTGAAG ATGCATATTG TGTTAATTCT TCCGCAACTA ATACTTTTAA TTGTTCAATT 10680
 50 TGAGCTACAG TTGGTTCTTC AGTTTGATGT TTTGCCATCA TCGTTTTTAA AGCACCACGT 10740
 TGGTCAAATG CTAATGCAGA AATGATACCT TCGTTGTTAC TTAATTGTTc AATTGATGCG 10800

55

TCATCATAAT TATTTAAATT GACATAACCT GTTGTGCTT CTTGTGCATT CAGCATGCCT 10920
 AAAGTATTGG CTTTTTTTAG TAAATCGTGG TCGTTTTCAT GATTAAGAAT TGCTGAAGTA 10980
 5 ATCCAGCAA CTGTAGAATC ACCTGAACCA ACCGGATTTA ATACACTTAT TGTGGAATA 11040
 TTCACTCTAT AGAATGTATG ATTGTGCTTA GCGAATGCAC CTTGTGCACC TAAAGACACA 11100
 ATAATCCACT CAATCCCTTC GAATAAGGGT TGTGACACTG CCTGTTTCAA ACTTTCTAAA 11160
 10 CTTTCATCAA GTGGCTGGTT AAGCAATTGA TATAGTTCAG AAATGTTTGG TTTAATGACT 11220
 GTAGGTTTGT ATGGATTTTC CAAAAGTtG CACCCGAGCA ATCTAATATC 11280
 ACAGGCACAC CTTTGTTTTG GCATCGTTCA ATGATTTGTG CATAATAATC TTGATTTAAT 11340
 15 CCTTTAGGTA AGCTACCTGA AATAGCAACT GCTTCAACTT TTTCTAATAA TTGTTCAAAA 11400
 TGTTAATAA ATCCTGCAGC CTCTTGATTA TCAATCTCCG GTCCCTGCTC TAAAATTTCT 11460
 GTTGTGCGC CTTCATGTAA AATTGCAATG CAGTTTCGTG TTTACCCTT AATGTTATAA 11520
 20 AATGCATGCT TGATGTCGGC ATGATCTAAT TTTTtagCAA TAAATTGACC TAATTCACCG 11580
 CCAATAAAAC CACTCGCAAG GACTGGCTCA CCTACTGCG CAAGTACTCT TGTACATTT 11640
 25 AAACCTTTAC CACCAGCTGT TTTACTTACT TCTTGAACAC GATTAACATC ATCTAATTTT 11700
 AATGCTGTTA ATGGGTATGA AATATCAACG GATGGATTTA ATGTTAAAGT TAAAATCATA 11760
 TGTGTCGTCC CTTAATCGTG GTATTGCGCT CTGTCCCAT TTTCTAAGAA TTCATCAAAG 11820
 30 AAATGTGGAT CAGCTTGATC TGCATTGCTT GTTTCTAAAT GTTTAATTTT AGCGATTAAT 11880
 TTTTGTCTT CTTAGTTGG TTTATATTCA GCATTAATAA ATGCATCGAT AATATCGCAC 11940
 ATTAATAACT CACCTATAAT ACGTCCACCG AAGCCAATAA CGTTCGCATT TAATCTTCT 12000
 35 TTAGCGTATA ACGCTGATGT CATATCACGT ACTAGTGTG AACGAACGCC AGGTACTTTA 12060
 TTTACAGCAT TGTTAATACC AACACCTGTT CCACAAATAC AAACACCTAA GTCTGCATTA 12120
 CCGCTAACAA CTTGTTGCGC AACTTTTTTA CCAAAAATTG GATAATGTGT TCTGTGAAA 12180
 40 TCGTATGTTT CTACGTCAAT GACTTCATGT CCTTTTGATT TTAATAATTC AGATACACGC 12240
 ATTTTGTAT CTGTAACAAT ATGGTCGCAT CCTAATGCAA TCTTCATAGT AATTTTCTCT 12300
 CCTTAGCACA TTTTATTAAG CATATCTACG CGGATTGGT GTCTACCACC ATCGTATTTA 12360
 45 CCTTCAACAA AACCTTTAAC GACATTTTTC GCTAATGTGT CTCCAACAAT TTCAGATCCC 12420
 ATAGTGATCA TTCTTGAATT GTTATGGCCT CTAGTCATAT ATCCAGAGCG TTCATCTGAT 12480
 50 ACTTCAGCAG CAATCATGCC TTTGATTTTT GTAGCAACCA TAAAGCTACC TGCACCAAAT 12540
 GCATCGATAA CAATACCTAA GTTACCTTCT TGACTTTGAA CATCTTTTGC TACAGCCAAA 12600

55

TCTAATAAGT ATGATTTGAT GACTTCTTTT AATCGTTTGC CAGCTTCATC TGAACCAATA 12720
 ATAATCGCCA TAATAAGACT CCTTTTTACT TTAATTTTGA AATACCTTTC TTAAATGTG 12780
 5 ACATATTTAT TTGTAGGTTA TGAAAATCTT GAGAAAAGGC TTTCAATTG ATTACGTTTA 12840
 AATTATAAAC ATAAACAAAC AATAAATCAA CATAATATGT TTATAATATG TTTGTTTATG 12900
 ACGTATTTTC AAACAATAAG TGAACATTCA TATTGTGGTG TTGTTTAAAT TAGGTATTTCG 12960
 10 TCTGAAATTG TAGTAAACT TTGTCGAGGT TCCCGTTGac ATAAATTTGC ATAAAAAaA 13020
 GCCCATAAAT GAATGCAAT TCACATTCAC TTATGAGCAT ATAGATACAT ATTTTAACAA 13080
 TGCAGTTATA CTTTTAATTT AGTCGACTAC TTCAATATAT GTTTTAATCG TTTCTACTTT 13140
 15 TTCTTCATCT TCATAGTCCA TGACCACTGC AGTCAATTTCG TTTAACTGAC AAAATGATGT 13200
 AAAATCTTCT TTGCCAACTT TCGTATGATC GATTAAACAAG TATTTTTCaA TTGAATTACT 13260
 20 TAGTGCCAGT TGTTGCGTAT AGGCTTCATC TAATGTAGAT GTCATCACAG CACCTTTATT 13320
 TACTGCGTTA CTACTAAAGA ACATCTTGCT AAATCTTAGT TTTTCCAACA TGGCGTTTCGc 13380
 cATTTACCTT ACAAATGCTT CTGTAATATG GCGCATTTCa CCACCAATTA AATAGACACG 13440
 25 AAAATGTGCT GTTTGTTTTT CTAACAAAAT TTTATACACC GGCAACAAT TCGTAATAAT 13500
 TGTGAGCGTA TGATGATTGA CTTCTTCTGC TAATAGTTCC ACTGTTGTTc CTGGTCCGAA 13560
 AAACAAAGTA TCCCCATCTT CAATTAATGA TGCAGCTTTT TTAGCTATAA ATCGTTTTTC 13620
 30 TGCAATTTGA CGGGTATGTT TTTCTTTATG CGATATTTCT TTATACTGAA ATGTTGAATT 13680
 ACTGCGTGCA CCACCATGAA TCTTCGTTAA AATCCCTTTA TTTTCCAATT CAATTAAATC 13740
 TCTTCGAACT GTCATATCAG ACACATTTAA ACCTTCGACG ATTTCATTCG TTCTTATCGT 13800
 35 GCCCTTTTTT TTAAGTAGTT TAGCAATTC GTCCAAACGT TCATGTTTAT TCAATGTAAA 13860
 ATTGCTC 13868

(2) INFORMATION FOR SEQ ID NO: 174:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4549 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

50 TTAAGTCAAC TTTGTCTATA CGGTTTGGAT CtTCTaCCCA ATGTCTTATA AAAGACAATC 60
 CCGCACCTGA AACATAACTC ATGAAATAAG AAAATGGTAT ACCATTAATT TGATCATTTT 120

	AATCTTTACC	CATACGAAAC	ATCAATTGAT	AAAATGCGAT	GTCTTTTTCT	ATCATTCTTA	240
	TTAAACGGT	CATAATTTGA	TGTATGTTAT	CCGTGGATAA	CTTAACTGCT	CCATTTAACT	300
5	TCTCATCATG	AATGAAGTCT	CTTATTTCT	CCAACGCTG	GTCTCTAAT	TTTTCAAGCA	360
	AATCATACTT	ATCATAATAA	TGCGTATAAA	ATGTACTACG	GTTAACATCA	GCTAAATCTG	420
10	CAATTTGTTG	CACAGTAATC	TCTTCTAATT	GGTGTGATG	TAAAAGTTCA	ATAAATGCAT	480
	TTCTCATTGC	AACTTGTGAT	TTTCTAATAC	GTCGATCTAT	AGTCATTTAT	ATCAAGTCCT	540
	CCCCAATGAT	TATAAACGTT	ATGTTCATT	TCCCACAAAT	CTCCAACATT	GATGATTGGC	600
15	ACACAATGTT	TACCTGTTTA	ATATAGGTGA	TACAAACAAA	CAGAAAAAGG	TGATAACAAT	660
	GAACCAACAT	TTACTAGGAA	ATCCAAAATT	AACTGTAACT	CATGTCAATG	AAGTTAAAGC	720
	CGGTATTAAC	CACATCGTTG	TCGACAGTGT	TCAATATGGA	AATCAAGAAA	TGATTATGGA	780
20	AAAAGATGTC	ACTGTGGAAA	TGCGCGATGG	CGAAAAATTA	TATATTAATA	TTTTCAGACC	840
	AAATAAAGAT	GGCAAATTCC	CTGTAGTTAT	GTCTGCAGAT	ACTTACGGTA	AAGATAATAA	900
	GCCTAAAATC	ACAAATATGG	GTGCCCTTTG	GCCAACATTA	GGTACCATT	CGACATCTAG	960
25	TTTTACACCT	GAAGAATCAC	CAGACCCAGG	ATTTTGGGTG	CCAAATGATT	ATGTTGTAGT	1020
	TAAAGTTGCA	TTACGCGGTA	GTGACAAATC	CAAAGGCGTC	TTATCTCCAT	GGTCAAAAAG	1080
	AGAAGCGGAA	GATTATTACG	ATGTATTGAA	TGGGCAGCAA	ATCAGTCATG	GAGTAATGGA	1140
30	AATATCGGGA	CAAATGGTGT	TTCTTATCTT	GCGGTGACTC	AATGGTGGGT	CGCATCATT	1200
	AATCCACCAC	ATTTAAAGC	AATGATTCTT	TGGGAAGGCT	TAAATGATAT	GTATAGAGAA	1260
	GTAGCCTTTC	ACGGAGGTAT	ACCAGATACT	GGCTTTTATC	GTTTCTGGAC	TCAAGGTATT	1320
35	TTTGCGAGAT	GGACAGATAA	TCCAAATATC	GAAGATTGGA	TTCAAGCACA	ACAAGAACAT	1380
	CCTCTGTTTG	ATGATTTTTG	GAAACAGCGT	CAAGTGCCAT	TATCACAAAT	TAAAACACCT	1440
40	CTACTAACAT	GTGCTAGTTG	GTCTACACAA	GGTTTGACAA	ACCGTGGCTC	TTTTGAAGGA	1500
	TTTAAACAAG	CTGCATCTGA	AGAAAAATGG	CTATATGTGC	ATGGACGTAA	AGAGTGGGAA	1560
	AGTTACTACG	CTAGAGAAAA	TCTCGAACGC	CAAAAATCAT	TCTTTGATTT	TTACCTTAA	1620
45	GAAGAAAATA	ACGATTGGAA	AGATACGCCT	CATGTCATTT	ATGAAGTTAG	AGATCAATTT	1680
	TATAAAGGCG	AATTCAAATC	AGCGTCACGT	GTCCCTTTAC	CTAACGCAGA	ATATACACCA	1740
	TTGTATTTGA	ATGCTGAAAA	TCACACATTG	AATCATGCAA	AGATTAGTAG	CGCGCATGTC	1800
50	GCACAATATG	ACTCTGAAGA	TAAACAACAA	GATGTAAGTT	TTAAATATAC	GTTTGACAAA	1860
	GATACTGAGT	TAGTTGGAAA	CATGAACTTA	AACTATGGG	TAAGCACTAA	AGACTCAGAT	1920

55

	CCTGATTTTA ATCATATTGA AAATGGTCAA GTAGCTACTG GTTGGTTACG CGTATCACAT	2040
	CGTGAATTAG ATCAAGAAAA ATCCTCAATC GCGCAACCTT GGCATAAACA TGAAACAGAA	2100
5	TTAAAGTTGT CACAAGATGA GATTGTACCT GTTGAAATCG AATTGTTACC TTCAGGCACG	2160
	CTATTTAAAC AAGGCGAAAC ATTGGAAGTT GTTGTAAGG GTAGTGAAAT TGTAATTGGT	2220
	AATAGTACTC CTGGCATGAA AACACGTTAT GAACATGAAG AAACCGTAAA TAAAGGCATG	2280
10	CACATGATTT AACTGGTGG TAAATATGAT TCACAATTAA TCATTCCCTAT CGTTAATTGA	2340
	TATGCAGCAA TTACGGTCGC TTTTGATTAA AAGTGACATA GTGATAGGAC TGTATAACAA	2400
	GAGAAAGCCA CACGCTTGA ATCTTAAACC AAGGTGTGGC CCTTTTTATT ATTGATGGCT	2460
15	ATTTAATTTT ATAACACTAT CGTATTTTCT TTTTCATGAA TCATTTCAAT AATGACATTA	2520
	TCTTCATTCA TTACTGCTAC TTTAGGTGCA TGGTTTTTAA TTTCTTCTTC ATTCAACTGT	2580
20	GCATAAGTCA TGATTATGAC TACATCGCCT ACTTCAACAA GTCTTGACGC TGCACCGTTT	2640
	AAACAAATTT TACCACTACC TCTTTCACCA GCTATTACGT ATGTTTCAAA ACGTGCACCA	2700
	TTATTATTAT TCACGATGGC TACTTTTTCA TTTGGCAAGA TGTCTACCGC TTCCAATATA	2760
25	TCTGAATCAA TCGTAATGCT ACCTACATAA TTAAATTTG ACTCAGTCAC TCTTGCTCTA	2820
	TGAATTTTAG CATTCATCAT TGTTCCTATC ACTTTATTCA GCTCCAATTA TTATATTATC	2880
	TATTAAACGC GCTTTTGAAA ATTAAACAGC TAACGAGATA AATATGCGTC CAGTTATTTTC	2940
30	GTGTGTCTCT ACTAATTGAG GATAACTATA AACAGCAACT TCTTCAATGC GTTCACCTAT	3000
	ATGTGATTCA AGATATTGAG TAACCCTGTC TATAATTACT TTACTTTGAC GTTCACCGTC	3060
	TTGATACAAC GCTTGTGCTA ATAGCAAACT TTTACTTAA TGTACCGCTT CTTGTCGTTT	3120
35	TTGCTCCGTT AAATAAACAT TTCTTGAAC TTTGCGCAA CCATCTGCTT CTCGAACGAT	3180
	ATCAATACCA ATAATTTCAA CGGCATGATT GAAGTCTTTT ACCMTTTGCT CGaCAATAGC	3240
	CAATTGCTGG GCATCTTTTT TACCAAAATA AGCATAATCC GGCATAACAA TATTAAATAG	3300
40	CTTATTAACT ACTGTTACCA CCCCATCAAA ATGCCCTGGT CCGtTCGCTC CTTCTAACAC	3360
	ATCAGCTAAT GGGCCTACTT TGACATCAAT ACCTAATTCA CCTGGATACA TATCTTCTAC	3420
45	TGCAGGATGA AAAACAATGT CCGCTCCTAC TTCTGATACT AATTCTAAAT CTTTATCAAT	3480
	TTGTCTCGGA TAAGCATCGA AATCTTCGTT TGGACCAAAT TGTAATGGAT TAACAAATAC	3540
	ACTCACAATT GTAATATCAT TTGTACTAAC TGATTGCGGT ACCATCGTTA AATGTCCATC	3600
50	ATGTAAGGCA CCCATTGTTG GGATAAAACC AATCGTTGTG CCTGAGCGTT TGGCTGCTTT	3660
	AACAATGTGT TGCATCTCTT TTACCGTAGT AATCAGCTTA GTCATTGTTA TTAACCTCAT	3720

55

GATCGTATTG TTTTAAACCA TCCACACCAA CACTAAAATC AGCAAATTGC TTCACAAATT 3840
 TCGCTTTATG TTCAACACCA TAATTTAACA TATCGTGATA AACCAATACT TGACCATCTG 3900
 5 TACCTTTTCC TGCACCAATA CCAATGACTG GAATTGTTAA GTGCTTGCTA ATTTCTTCTG 3960
 CTAAATCATT TGGAATTGCT TCAAGTACTA ACGCAACTGC ACCAGCTTGT TCTACATTTT 4020
 TCGCGTCTAA AATAAGTTGc TCCGCTGCTT CTTCGTTGTC ACCTTGTAAT TTATACCCCA 4080
 10 TAACGCCAAC ACTTTGAGGT GTTAATCCTA AATGTGCAAC AACAGGAATA CCAATTGCCG 4140
 TTGCTTTTTT AATAAATGGT GTAATATGCG CTCCTTCTGC TTTAATTGCA TTTGCATTCCG 4200
 TCTCCTGATA AAGCTTTAGA GCATGATTTA AGTCTTGTGT CATAGAGATG CCTACTGCAC 4260
 15 CAATCGGCAT ATCAACAAC ACAAATGTAT TTGGTGCGCC TCTTCTTACT GCACGACCGT 4320
 GATGAATCAT ATCTGCTAAC GTCACCTGTA CGGTACTTTC ATAACCTAAT ACAGTCATAC 4380
 20 CAAGTGAATC CCCAACAAGA ATCATATCAA TACCCGCTGC TTCCACTTGT TTAGCACTTG 4440
 GAAAATCATA AGCTGTTACC ATAGAAATTT TAGTTTGCTT TTGTTTCATA TCTATTAAAT 4500
 GACTTACTGT TTTCAATGTT ATTCAACCTC TTTTGCAGT ATnATTAGA 4549

25 (2) INFORMATION FOR SEQ ID NO: 175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8339 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

35 TTATCTTTTG TTGTTTCCTT AGACAAACGA CTAACCACAT TATAATGGAC TAATTTATTA 60
 ATTFIATTTA ATTCCATTAA GTTATCCGTA AACTAAGTG AAGATGCGGA GTTCACTCTC 120
 40 GTTGTTACTC TTCGTTTTAA TAAAGCACCT CGTAATAATA CAATCATTCT TCTTATTAAT 180
 GATGCTTGTC TATATACCTG TGTTCTTTCA GCATAACGCA TATAGTTTTC AAGTACACTA 240
 TTCGTTATTT GTCCTTCATC TACTAAAGAC TCTAATGTTT TkGTTTCTAC ATTAAGCA 300
 45 ATTTTTTGTA GACGTTCTAA TTCTTTAGAG TTTTCATCAT CTTTCTCTAC AGTTTTTAAA 360
 AATGCTAATT TATCATGATA TTCTTTAATC ACGTTACCAT ATTTAAACT TGTTCGAAA 420
 GTAGATTTTT GATTTAGATA ATCAATAACT TGTTCTAATA TATAAATTCT AGCAACTTTA 480
 50 AACGACATAT TGCCAATTAC TGTTTTAGGT GCAGGTTTCG TTAATAATGG CAATAATACT 540
 TCGCCAACTA CCAAACTAAT AATAACCATA CCAGATGCAA TAAATAATAA GTCGTTTCTA 600

55

	ATTGTTCCAT GCACACCACA TAACGTCATA ATTAAAGCGT ATAAACTTCG CTTTGGTGGT	720
	TTCTCAGTCG TTGGATTATC ATCATCATTT TTAGTCATCA TTTTTTGGAA TGGACTGATG	780
5	GCTAAATAAA AATAAGGATA TAAGACATAA ACCCAAACAA ATCTAAATAG ATAGACAGCT	840
	AAAGCAACAA CAATAGTGAT GCCTATTAAA AAGATTAAAT TGTGCGGTTT TGTTTTGATA	900
	ATTTTAATAA TAACCTCAGG TACTAAAAAT CCTAATATTG AAAAAACAAA GCCATTTAAA	960
10	ACATAACCTA GTATATTCCA TGTATGATTG TAACTCATTT GCAGTTGTGT ACGTACTTGC	1020
	ATAATTCTGT CACGTTTCGA ACCATGTACA AGTCCTGCAA CTACTGCTGC AATGATTCTT	1080
15	GATGCGTGaa ACAATTCAGC AATTAAATAC GTAACAAATG GTGTTAACAA TTGAATAATT	1140
	GTAAACATAT TAATGTTTTT ATATCCTCGA CGCATCAATG TTAATCGGAA CCTTACTAAT	1200
	GCCATACCTA TAAGTAAACC AACCCTGCG CCACCAATTG ATGCAATTAA AAACAACCTGA	1260
20	ACAGCATCAA CAAGTGAAAA AGCACCTGTA ACTAATACTC CAACAGCTAT TTTAAATGAA	1320
	ATAATACCAG CAGCATCATT CAATAATGAC TCACCTTCAA GAATTGTCAT TGCTCCTTTT	1380
	GGCAAGACCT TTCCTTTAGT GATTGCTTGC ACTGCTACTG CATCAGTAGG ACAAAGAATG	1440
25	GCAGCAATTG CAAATGCTGC TCCAATAGGT AAATCTGGCC AAATCCAATG AATAAATAAA	1500
	CCTACACCTA TCACAGTAGT AATGACTAAT CCTAATGCCA TCATCATCAC TGGCTTAATA	1560
	TATTTCTTAA AATGGACTCT AGAAACATTA ACACCTTCTA CAAATAACAA AGGCGCAATC	1620
30	ATTGTTACCA TAAACAATTC AGAATCAAAA TTAAATTGAA CAGGGATTGG GGTAAATAAT	1680
	AGTAACATGC CCAAGAAAAT TTGTATAAAT GCTAGGGGTA CTTTAGGTAT GAAAGTATGG	1740
	ACAAACGAAC TTAGTATCAC AACAGCTATA AATATAAGAA TTGTTTCAA TATTTCCAAA	1800
35	CTTTCACCTC TCTAAAAAGT ATTGTTTAAAT TGAAAATTAA GTATCACATC TCGTTGTAAT	1860
	TATACTTTAG AGGATAAATT GAGTTAGCGA CCACAAAAGC ACTTTAATAT AGATATATGT	1920
40	CTACGATTGC AGTACTTAAA TTTGCAATTA TTTAATTTTA TTTTATCACT AATTGTTTGT	1980
	ATAAATAAAC AACTTGCTTT CACATAACAA CATTAACCTA TAATACAAAA AATGAGCACC	2040
	TTAAATCGA CTAACCAATT TCaAAGTACT CTTTTAATGA TTAATTTTGA AAACAGATTT	2100
45	TCaAAAGCAT TGTTATGCTT AACAAATTTAG CCAACACTTC AATCGTTTTG ATACCATTTT	2160
	TTACGATGCT CTTCTCGTTT TTCAGCACGT AATTGTAATG CTTCTGTAGA GTTTTGTTC	2220
	TTTGAACCTA ATAATATTGA TGCATGTGTG TGAGCATCAT TTTTTCGATA CATATAAGCG	2280
50	CCGTTGCGAT AAGCAGCGCG AGCGACTAAG TGCATGCCGA CTGGTGAAGT TAAATTAATA	2340
	AAAACAAGTG ACAGTAATAA ACGCACACTG AAAATCCTG TATTCACAAT AAAATAAATT	2400

55

	CTTAAGAAAA CATCTTGGA TTTACAGATA CCTATTGCAC TAATAAGAGC AATAAAACTA	2520
	CCTAACAACA ACATCACAGC AGCAATAAGA CTAAAGATTT CTTTTGTTAT TTCCATTAAA	2580
5	CACATGCCCC CCACCAATAA AGCGTGATAT TGAAACAGAA CTTACAAAAG ATATAATGGC	2640
	AATGAGCATG ATTGAATCTA AGAAAGAAAC GGTGCCCATTA AGTACACTTA ACACACCCAC	2700
10	AATTGACATT ACGACAGCAC TTGTTGTATC AAATGTAACG ACACGATCTG CTGTTGTAGG	2760
	TCCCTTGATT AATCTAAATA AACAGATGAT TAATGCAATT CCAAAAATAA TGAGTGAAC	2820
	AATAATCATA ATATGTGTTA TTGTTTGTAT CATCGCGACA CCTCCAATAT TAAGTCTTCA	2880
15	TAATGCTTAA TACTTCTTAA CAACTATCT TTTTCTTTTT CTGACACGTC GATACTATGA	2940
	ATAAAAACT TTTTAGAGTC TTGAGAAATT CGTATTACTG TAGACCCTGG AGTTATAATA	3000
	ATTAAATTG TTAATAATGT TATTGACCAA TCACTTGTTA GTCTTGTTTC ATATGAAAGT	3060
20	AATCCAGGGT TCATATCTTT TGTTTTAAAA AGAATATAAT TAATCGTGCT AATGCTAGAT	3120
	GTTATTAATT GATATAAATA AACACCTAAA AATTTAATAG CTACCCATAT TTTTCTAACA	3180
	TAAAAATCAT CGCTGAAAAA CCTGTGTAAT ATATAAATGA CAATTAAACC AATTAGATAT	3240
25	CCAGAAAAGA AAGTCGAGAA TTTAAATGA TCTTCATCTT GAAATAATAC CCATAAGAAT	3300
	GCAATGATAA TATTTAAAC TATTTGATTC ATTTAGTCCT CTCCTTTCAA ATGCGGATTT	3360
	ACAAGTTTTT GATATAATTG ATCACTCGTG TTCAACTCAG TTGCATCACT TGTAACATTT	3420
30	AACACAACAG GTGCAGCAAT TCCGATTGCG ATAACCACAA CTACTAAAAT ACTTAAAATT	3480
	CTTTTTCGAT ATAGCGGGAT TTTCTTAAAA TTAACCTCCT CCCCATCTTT ATCTCCAAAA	3540
	TACATATAAA AAAGTATCCT AAATAAATG TACATTGCAA TTAGACTAGT AATAATCATT	3600
35	AACGCTAGTC CAATATAATT GCCATTTTGC AATGCACCTT GGAAAATAAG TACTTTCCCC	3660
	GGAAAGCCAC TAAATGGAGG CACGCCGCCA ATAGCAAAAA TCATTATAAT AAACGCAACT	3720
40	CCAAATAAAG GTTCTTTTTT AGCTAAGCCA TTCAAATATT GATATTGTCG ATAGCCTGTA	3780
	ATGTAACTA AACTACCAAT AATAAAAAAT AGCAATGTTT TTACAACAAT GTCATTTACC	3840
	AAATAAATA TTGCACCATT AATACCTGCA AACGTGTTTG TTCCTAAACC TAAAATGATA	3900
45	AATCCTATTG AGATTATGAC TTGGTAAGCT GCAATCTTTT TAATATCTTT ATAAGCAATG	3960
	ACACCTATAG CGCCGATGAC CATAGTTATA GCAGCCATAG TTGCTAGCAA TGGATGTATG	4020
	AGATCATTAT GTTGATCAAA TAGTAAAGTG AAGAATCGAA TTAATGCATA GGCCCTACT	4080
50	TTGGTCATTA ACGCTGCAAA TAATGCTGCA AGCTCAGTAT TTAACACAGC GTAGGCTTTG	4140
	GGTAGCCACA TAAAAGGAC CAGCGCTGCT TTCGCACTAA ATGCGACTAA GAAGATTAAT	4200

55

	AAGTTTAAATG TACCTACTGT TTTATAAAGT AAACCTATAC CTAATAAGAA TAGCCATGAA	4320
	CCAATAATAT TCAAGACAAC ATAAATAATT GCAGCACGTA ATTGTTCTAC AGATTGTCCA	4380
5	AGTGTAATGA GTACAAATGA CGCTAGTAAC ATAATTTCAC ACATGACGTA TAAATTAAAT	4440
	AAATCTGATG TTAGAAAAGA GCCTATCACG CCAACACTTA AAAATAATAT GAACGATGGC	4500
	AAGTGATAAC GATTTGCTTT ATGTTGCGCA CGCCCAAATC CGTATGCCAT AATTAAAGTA	4560
10	ATCACAAACG AAGCGGTTGT AACCATAATT AAACCTAAAG AATCTCCTAA AAACCTGTATA	4620
	CCAAAGGGCG CTGACCATCC TCCAAAGTCT AGCGTAATTG GACGGTGACG CTGAACATAA	4680
	ATTAATAGCA TTAATGAAAT AATTGTGGTG ATAGTCATTG TACCTAAGTA TAAATATTTA	4740
15	GAAATACGAT CATTATTTTT TAAAAATACA AGGATTAAGG CACAAAGGAA TGGTAATAAC	4800
	ATTGGTAAAA TCAATAAGTT ACTTAGCATC ATCTTCCCCC CTTAGGCCTT CAATTTTCATC	4860
20	TTCTTTTGTT ACTTTATAAG TTCTATAAAC AAGTACAAGT AAAACGCAG TCATCCCCAA	4920
	CCCTATAACT ATTGCAGTTA GTACAATAGC TTGTAACAAG GGATCAACAA ACAATTGGTT	4980
	TCCACCAGTT ATTAGTGGTT CTGATCTACT AGAACCATAC GTTCCCATAC TCATAATAAT	5040
25	GAGATTACCA GCATGAGTAT ATATTGAAAT TCCGATTACA ATACGAATTA AATTGATTGA	5100
	TAAAATCATA TATGTTCCCTA TAAACACTAA AAATCCTATA ACTAGTAATA ATATTAAATT	5160
	CATGATCGAC CTCCGCTAAG CGACAACATC ACTGTGACAA TAACACCAAC AACTGAGAAT	5220
30	AAAATACCTA ATTCAAAAAG TGTATTGTA CTTACATGAA TTTGTCCTAA AATTGGAAGT	5280
	ATCCAAGTTG TTTCATATTG AGACAAAAAT GGTMTTCCAA AAAACATAGG TATTATCGCA	5340
	GTAATAGATG ATACCAATGC TCCAATAATC ATTAAATTC TAAATCAAT CGGTAAACTT	5400
35	TCTAAAACCT CTTCAACATT AAAAGCCAGA AACATTAAAA TAAACGCTGA ACTAAATATT	5460
	AAAACACCAA TAAACCCACC ACCAGGATTA TTATGACCTG CGAAGAAGAC ATAGAATCCG	5520
40	AAAGTCAATA AAATAAATAC AACAAAGTTTC GTGACCGTTC TTAACACGAC ATCATTCTCT	5580
	TTCATCTTGT CCCCTCCGAT CTTGATAATT TAATAATGtg TAAATACCTA GCCCAGTAAT	5640
	AATTAACACT AATCCTTCAA ATAATGTATC TAATGCTCTA AAGTCACCAA GTATCGCATT	5700
45	TACAATATTT TTACCACCTG TTAGTTTGTC AGCTTTTAAA TAAAAGTCTG ATATTGATGA	5760
	TAAACCATCT GTTTGTGTG TAATAAAAAAT TAATGATACA ACAATAAGTG CCATCAAGAG	5820
	TGATACAGAA ATTTTAATTA TTTCTCTTTT TTTGTTAGCG TTAGATCTTG GCACGTTTGG	5880
50	TAATCTTGAA AAACCTGACAA TAAATAGTAT CGTCGTTATT GTTCAACTA CTAGCTGAGT	5940
	CAATGCTAGA TCAGGGGCTT TCATTGCTAT AAAGAATAAG GTCACAACAA ATCCGATGAC	6000

55

	GACAGTTACG ATTGCTAATA TAATTTCTAA TGCCCCAAAT TCAGAAACAT GTAACGTATG	6120
	TACTTTAGGA AGTCCaATTC GAATATAACC ATATCCAATG ATAATCATAA ATATGCCTAA	6180
5	GGTCATAATA ATGTACTGGT TTAAACGATC TTGCATAACA CGTTTAAATC GCTTCGTAGC	6240
	AAACTTTTCA AAATGTCGAT ATACCATCTC ATAGCTTTTT GAAACTGAAA TCTGTCTAAT	6300
	TTTACCTGTG AACACTTTTT TCCAATCTAC TTTGATTGCT AGTACACTAC CCAATAAAAT	6360
10	AATGATGATG GTTAAAAGAA GCGGTATGTT AAATCCATGC CATTGCGAAA CATGTGGTGC	6420
	CAATTGATCA ATTTGATGAT TACCACCTGA TACAGCTCTT AATGCAAGAA CGATAATCCC	6480
	CTTCCCAAAT ATATnTGGTA CAAAAAAGAT TACAGGTACT AGCACCATTA aTATAAGAGA	6540
15	TGGTAAACTA aACAACCATG GTTCGTGGAT ATTTTTTTTA GTAAAAACCT TAGAATCATA	6600
	TTTTGtCCAA AATACTTCTT TTACCATGTA TAGTGCATAT GTGAATGTAA AAACACTCGC	6660
20	AATAACACCA ACAAACACGA TAGCTATCAT TGAATCAAA CTAAATTGGG ATAATTGTCC	6720
	AGTTTGTGTT AATGCATCTA AAAACATTTT TTTACTTAAA AATCCATTTA AAAATGGTAC	6780
	TCCAGCCATA GATAGAGCCG CTATCGTCAT GACTAGATTC ATTTTAGGAA ATAGTTGACG	6840
25	CATTCCACTT AAAATTCGTA TATCCCTTGA ACCTGCTTCA TGATCTAAAA TACCTACTCC	6900
	CATGAAAAGC GCACATTTAAGATGGCATG ATTcATTAGa TGAAATAGcG CACCArATAA	6960
	TACmAATACA TAAATaGATG CTATTGCGTC TTGTTGGTGT TGAGCATATC CGCCACCTAT	7020
30	ACCCACCATA GCCATAATCA TCCCAAGTTG ACTGATTGTA GAGTACGCTA GGATACCTTT	7080
	TAAATCCCAT TGTTTTAAAG CTGTAATTGA ACCAAATAAC ATTGTTATTA AACCAACAA	7140
	CGTAACGATA TATACGTACA TATTGCTaNG ACCTAATAAT GGTGTAAATC GAAGTAATAG	7200
35	AAnGATACCA GCTTTTACCA TCGTGGCTGA ATGTAAATAA GCACTTACAG GTGTAGGTGC	7260
	AGCCATTGCT CTAGGTAGCC AGTATGAAAT GGAraTTGTG CTGATTTTGT AAATGCACCT	7320
	AATAAAAACA TAAAAATCAT AGGGATAAAC AATCCATGAT TCTTAATATG ATCTGCTTGT	7380
40	CCTAATATCT CTGTGATGTT ATTCGTTCTT GTCATGATAT ACAGCATAAT AAAACCAACT	7440
	AATAACGCCA ATCCACCAAA TACTGTAATC ATAAATGATT GAATCGCACC AAATTGACTG	7500
45	TCACCATTGT TATACCAATA TGAnATCAAT AAAAATGATG ATmCACTCGT TAATTCCCAA	7560
	AAaATGTACA TCMaTATCGT ATTGTCTGAT AATACaATAC CAATCATACT GAACATAAAT	7620
	AACGTTAAAT AAAAATAAAA CCTTGGTAAA TTGTCTTTTC GAGAGGATAA ATATTGAGTT	7680
50	GCATAGAAGA ATACTGCAAT TCCAATAAGT GAAATAATAA GAGAAAACAT TAACTTAAA	7740
	CCATCTAAAC GTAAATCTAA ATTAATATCT AATGTCTTAA TCCATGGAAT AGAGGTAGAA	7800

55

GGTGCAACCA ACGCTATGTA CCCGGCATAT TTAGCCAATG CTCTACGTTT AGACATTAGA 7920
 AGTATCATCG CCATAATCAC AAGTATAGCA ATTAATAAAT AAACCAAACCT CATTATTAGC 7980
 5 CTCCTTTGTT TCTATAATTG TAATGAAATA TAAATACTAT GTTCACACTC ATTTTCTAAA 8040
 CCGATAAAAT TTAGTGTTTC AATAGCAGAT TGATGCCCTA AATACTTTTG AATGACTGGT 8100
 ATAAGTATAC CTTTTTGATA AGCATGATAT GCAAATGTCT TACGCAATGT CGTTAGTCCT 8160
 10 ACATTATCTA TACCAGCTTC AATTGATGCT TGGTGAATT A TTCGATATGC TTGCTGTCTA 8220
 GATAATACTT GATTTGTTCG TAGTGATTGA AAAAGAACGT CTTCAATCGA AAGACTCCTG 8280
 TCCTCTATAT ATTGAAGTAG TTCTTTCGAT AATGTTTCTG GTAACCTAAT TTTAATCAA 8339
 15

(2) INFORMATION FOR SEQ ID NO: 176:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 588 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

CCCGATTTT TTACGTAATC TAATACATAC GGCAAAATCA ACTTTAATCA AAAAAGACTC 60
 ATACACAATG CCTTTAAAGC ACATGTATGA GTCCTTTTTA GTAGTTTATA TCAAAAATA 120
 30 GTTTAATGTA TAAATTAGTT TTTGTTTACA GATGCGTCGT AGATTGATTC TACAGCATCA 180
 CCTAAAGCTT TATCGAATTC TTCTTTAGAT TGATCAGCTC TTAAATCACT AGCTAATGCA 240
 CGTGAGAAAC TTGCGATAAG TTCAGCGTTA TCTTTAAGTA ATTCATTTGC TTTTCTCTG 300
 35 CTGTAACCAC CTGATAATAC AACGACACGA ACAACATTAG GATGTTTCAGC TAACTCTTTG 360
 TATAAGTTTG GTTCAGTAGG AATTGTTAAT TTCAACATTA CTAATTGATC AGCATTTAAG 420
 CTATCTAAAC CTTTTTTAAG TTCAGCTTTT AATACTTTTT CAATTTTCAGC TTTGCTTTTT 480
 40 GCATTAAATAT TAACTTCTGG TTCGATAATT GGAACATAAC CTTTAGCAAT AATTGTTTA 540
 GCAACTTCAA ATTGTTGTTC AACAACTCTT TTGATACCTT GTCATTT 588

(2) INFORMATION FOR SEQ ID NO: 177:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2841 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	ATAGAGTnCT GGrACTTACT ATGACATATG GCGCTAGAAT GGCTGAGCCA GGTGAATTTA	60
	CAAAACGTGC CTTTTTAAAT GGTCTGATTG ATTTATCTCA AGCTGAAGCA GTTATGGACT	120
5	TTATTCGCTC GAAGACAGAT AGAGCTTCTA AAGTTGCGAT GAATCAAATT GAAGGTCGTC	180
	TAAGTGACTT AATCAAAAAA CAACGTCAAT CTATATTAGA GATACTCGCT CAAGTGGAAG	240
	TGAATATTGA TTATCCTGAA TACGATGATG TTGAAGATGC GACTACTGAA TTTCTTTTAG	300
10	AGCAGTCTAA AGAAATCAAA CAGGAAATTA ATCGTTTATT AGATACCGGT GCGCAGGGTA	360
	AAATTATGCG TGAAGGTTTA TCTACAGTTA TTGTTGGTAA ACCAAACGTA GGTAAATCAT	420
	CGATGTTAAA TAATTTAATA CAAGATAATA AAGCGATTGT AACTGAGGTA GCAGGTACTA	480
15	CTAGAGATGT CTTAGAAGAG TACGTCAATG TTCGTGGCGT GCCATTAAGA TTAGTTGATA	540
	CTGCTGGTAT ACGTGAGACA GAAGATATAG TAGAGAAGAT TGGTGTGAA CGCTCTAGAA	600
	AGGCTCTTAG CCAAGCAGAC TTAATTTTAT TTGTATTAAA CAATAACGAA GCATTGACwC	660
20	AAGAAGATTA CACATTATAT GAAGTGGTTA AAAATGAAGA TGTAATCGTA ATTGTTAATA	720
	AAATGGATTT AGAGCAAAAC ATAGATATTA ATGAAGTTAA AGATATGATA GGTGATACGC	780
25	CATTAATTCA AACTTCAATG TTAAAACAAG AAGGTATTGA TGaATTAGAA ATACAAATTC	840
	gAGATTTGTT CTTTGGTGGa GAAGTACAAA ATCAAGATAT GACTTATGTT TCTAATTCAA	900
	GACATATTTT ATTATTAAAA CAAGCAAGAC AAACGATACA AGATGCGATT GATGCAGCAG	960
30	AATCTGGTGT GCCTATGGaT ATGGTACAAA TTGATTTAAC TAGAACTTGG GAAATATTAG	1020
	GAGAAATTAT TGGTGAGACT GCAAgTGATG AACTCATCGA TCAGTTATTC AGTCAATTCT	1080
	GCTTAGGTAA ATAGTAATTG AAATAGACGG AATACCGTCT TAAGAAGGCT AGTAAGATAT	1140
35	CAAATAAGGA GGTTTATATT GTGGTTCAAG AATATGATGT AATCGTTATA GGTGCGGGAC	1200
	ATGCAGGTGT AGAAGCAGGT TTAGCATCTG CAAGACGTGG TGCTAAAACA TTAATGCTAA	1260
	CAATAAATTT AGATAATATT GCATTTATGC CATGTAACCC ATCTGTAGGT GGACCAGCTA	1320
40	AAGGTATCGT TGTTCTGTA AATTGATGCTT TAGGTGGACA AATGGCAAAA ACAATCGATA	1380
	AAACACACAT TCAATGAGA ATGTTAAATA CAGGTAAAGG ACCTGCTGTA AGAGCACTAA	1440
	GAGCGCAAgc AGaTAAAGTA CTTTATCAAC AAGAAATGAA ACGCGTGATT GAAGATGAAG	1500
45	AAAATTTGCA TATAATGCAA GGTATGGTAG ACGAACTTAT TATAGAAGAT AATGAAGTTA	1560
	AAGGTGTACG TACAAATATT GGTACAGAGT ATTTATCTAA AGCAGTAATT ATTACAACGG	1620
	GAACATTTTT ACGTGGTGAA ATCATTTTAG GTAATATGAA GTATTCAAGT GGACCAAATC	1680
50	ACCAATTACC ATCAATCACA TTATCAGACA ATTTAAGAGA ACTTGGTTTT GATATTGTTT	1740

55

AAATACAACC AGGTGACGAT GTAGGTCGTG CATTGAGCTT TGAAACAACA GAATATATAT 1860
 TAGATCAATT GCCATGTTGG CTAACGTATA CTAATGCTGA AACACACAAA GTTATCGATG 1920
 5 ATAATTTACA TCTATCTGCA ATGTATTCAG GGATGATTAA AGGAACCGGG CCACGTTATT 1980
 GCCCTTCAAT TGAAGATAAA TTTGTTTCGAT TTAATGATAA GCCGCGACAT CAACTTTTCT 2040
 TAGAGCCTGA AGGTGCTAAT ACAAATGAAG TATATGTGCA AGGATTGTCT ACAAGTCTTC 2100
 10 CTGAACATGT GCACGTCAAA TGTTAGAGAC GATACCAGGT CTTGAAAAAG CAGATATGAT 2160
 GCGTGCCGGC TACGCAATTG AATATGATGC GATTGTGCCA ACGCAGTTAT GGCCTACACT 2220
 TGAAACGAAA ATGATTAAAA ACTTATATAC TGCAGGTCAA ATTAATGGTA CATCTGGTTA 2280
 15 TGAAGAAGCA GCAGGACAAG GATTGATGGC AGGTATTAAC GCTGCAGGTA AAGTGTTAAA 2340
 CACAGGCGAA AAGATATTAA GTCGTTTACA TGCATATATT GGTGTCTTAA TCGATGATCT 2400
 TGTAATAAA GGTACTAATG AACCTTATCG TTTACTAACA TCACGTGCAG AATATCGTTT 2460
 20 GTTACTACGT CATGATAATG CTGATTTGAG ATTGACGGAT ATGGGATATG AACTTGGTAT 2520
 GATTTCTGAA GAAAGATATG CACGTTTTAA TGAAAAACGT CAGCAAATTG ATGCGGAAAT 2580
 TAAGCGTTTA TCAGATATTC GTATTAAACC AAACGAACAT ACGCAAGCGA TTATTGAACA 2640
 25 ACATGGTGGT TCTCGCTTAA AAGATGGTAT TTTAGCTATC GATTTATTAC GCAGACCTGA 2700
 AATGACTTAC GATATAATTT TAGAACTTTT AGAAGAAGAA CATCAATTGA ATGCAGATGT 2760
 30 TGAAGAACAA GTAGAAATAC AAACAAAATA TGAAGGTTAT ATCAATAAAT CACTACAACA 2820
 AGTTGAGAAA GTTAAGCGTA T 2841

(2) INFORMATION FOR SEQ ID NO: 178:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3025 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

ATCTAATTTT AAACCCGGTG ATAAATTGCC AAGCGTGACG CAATTAAAAG AACGTTATCA 60
 45 AGTAAGTAAG AGTACTATCA TTAAAGCATT AGGCTTATTG GAACAAGATG GTTTGATCTA 120
 TCAAGCACAA GGCAGTGGTA TTTATGTGAG AAATATTGCT GATGCCAATC GTATCAACGT 180
 CTTTAAGACT AATGGTTTCT CTAAAAGTTT AGGTGAACAC CGAATGACAA GTAAGGTACT 240
 50 TGTTTTTAAG GAGATTGCAA CGCCACCTAA ATCTGTACAA GATGAGCTCC AATTAAATGC 300

55

	CGAATATTCT TATTATCATA AAGAAATCGT GAAATATTTA AATGATGATA TTGCTAAGGG	420
	CTCTATCTTC GACTATTTAG AATCAAACAT GAAACTTCGT ATTGGTTTTT CAGATATTTT	480
5	CTTTAATGTA GATCAACTCA CTTCAAGTGA AGCTTCATTA CTACAATTGT CTACAGGTGA	540
	ACCATGTTTA CGTTACCACC AGACTTTTTA TACAATGACT GGCAAACCCT TTGATTCATC	600
	TGACATCGTA TTTCAATTATC GTCATGCACA GTTTTATATT CCTAGTAAAA AGTAATAAAT	660
10	ACATAAAAAC GTCTATATCC CAGTTATAAA CTGGAGTATA GACGTTTTTT TACGATAATA	720
	ACAATGGCTC AAATTGCTAT TATCTTGCTT AGGTTTTTCG TTTTAGAAGA ATATTGCTAC	780
	AAAGACAGGC ACAACTGCTA CAACAACTAC ACCAACTAAC ACTAAAGCTA TACTTGCCAT	840
15	TGATTCTTCT ACAGGTCCTA ATTCTTTGGC TGGTGCTACA CCTAATGTGT GACCACCTGT	900
	TCCAAGTGCT AATCCTCGGG CAATAGGGTT AGTAATTCGG AAAAGCTTTA AGAATTTATT	960
20	ACCTAGGGCA TAAATAATGA CACCATTAA AATAACTGCT AATGATGTTA ATTCTTTTAT	1020
	ACCACCGATA CCAGCTGATA CTGGTAACGC AATCGCTGTA GTTGCTGCTT GAGGTAACAT	1080
	TGATAAAATA ACATCATTGG CAAATTGTGC TAACTTCGCA AAAGTTAAAA TAATTAATAA	1140
25	CGCTACAAC GTACCGATAC CAATACCTCC GATGATACGA TGCCAATGTT TAACAAGCAC	1200
	TTACAGCTTT TTATATAACG GAATCGCAAA ACAGATTGTT GCCGGTTCTA AGAAGAAGTA	1260
	AATAATGTCT CCACCTATTT TGTAAGTCTT ATACGGAATG CCTGTTAAAT AGAGGAAGGC	1320
30	CACACCAAAT ACCATACTGA CAAATAGCGG TGCGAATAAG AAGAAACGAT TAGTTTTTTC	1380
	AAATAATATG GTCGCTAAGA AAAATGGTAT AACGGATAAC AGTATTCCGA AGTAAGGTGT	1440
	GTTTAgTGCT AAGTGGTTAA TCaTGAGCTT GTGCCTCCTC TATTTTGATC TTTTTTGTGA	1500
35	CTTTGTCAAC TTTAGATCTC GAAGTAACCT TCATAATAAT TTgTGTGACA TAGCCAGTAC	1560
	AAATAgTAA TAGTATTGTT GAGACGATTA TTAGTCCAAT GATTAAAAAT GGTGCTTGGC	1620
	TAATGACACC TAAAGAGTTA ACAACTGAGA TACCGGCTGG TACGAAGAGT AAGCCAATGT	1680
40	TATTTGTTAG TGTCGTTCT ACTTTTTCGA CTTGCGCTAA CTTAACAGCA CCAGTACATA	1740
	ATAATACAAA TAATAATACT AAACCGATTA CTGATGCAGG CATAGGAATT GGCATAAATG	1800
	ATTCAATTAT TTTGATACA AAGAGTACTA AAGCAATTAC AATGACTTGG TGAAAAAGT	1860
45	GTGCTGGTTT TGATGCGTCT TTTTGTGTT TCACGACCAT TGCCTCCTAC GTTTGATTTA	1920
	ACTAAAGTAT AGATGGCTCA CTTGATTTG CGTGATTTTT AGTCCGAAAT ACAAATATC	1980
50	ATAGGTAAAA TGCATAAAAA AAAGGATTAC TGTTAAAGTA ATCCTATCGA CGCTTTAAAA	2040
	TCTTTCATAA ATGAACGTCC AACTTGCATC TTGACACCAT TTGTCAATAT TACCATATAA	2100

TGAATACGTA TAAAATAAGT GGGATTCAAT CGTTTTTCAT AACGATTCAA TGGCTCTGTT 2220
 GTTTCGTATT TATGATTCTG TGTATGTATG GTTGAATAC CATTATGTGT GCCAATCCCA 2280
 5 ATAATATTTT GTTGCTTTAA CATGTGAATT TTATCGTCAA TTTCAACAGG TAAGCTTTGA 2340
 TCAAAATTCG CCGACATATC ATTCGCAATT GCACTTGCCT TATTATCATC TTTGGCTTTA 2400
 GTCGCACGCA CTTTATTGAC TGCTTGTTCA ATACGTTTTT GACCAAACGG TTTCAAAATA 2460
 10 TAGTCTGTCG CATTTAATTC AAATGCCTGT ACTGCGTATT GGTCACTGTC AGTTGCAAAA 2520
 ATAATCGCAG GTGGCTCTTT CATCTTTTGA ATCTTAGCTC CTAATTCGAT CCCATTTTCA 2580
 TCCATTAAAT TGACATCTAA AAATATAATG TCATATTGAT TGATCAGTAG TGCTTCCAAT 2640
 15 GTTTCTTTTA CATTCTTCTG CTCATTAATT TCTTCAAAAC CACCAATTTT ATTTAATAAA 2700
 TATGTTAATT CATTACGTGC TAATGGCTCA TCATCTATGA TTAATGCTTT CATATTTATT 2760
 CCTCCTCTTG TCTTTCATAA GGAAGTACAC ACCAAAAAGT GGTACCGCTC GATGTCGATT 2820
 20 CAAATGTAA TGCTGCGGAT TTCCAAATA ATCCTTTTAG GCGTAAGTTT AAATTTTCTA 2880
 AAGCACTACC AGTTCAGAC TCTGATTCTA CAGATGTC TCCTAACAAA TGCAATTTAT 2940
 25 CTTTAGAAAT ACCCTGACCA TTATCTTGTA CAATAATACG TACATGTGTT GCAGTTTCTT 3000
 TAATCACTGA CACGTCAATA TCGTT 3025

(2) INFORMATION FOR SEQ ID NO: 179:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1689 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

ACAGAATTTT ACAGCATTTT TAGATGAAAA AATAAGCCAG TCATAGCGTT GATTTAACAA 60
 40 ATGAATATCA AAATTTAGTG GCTTTATATC AATAAAGGGT TTGTGAATAA TTGATACTAA 120
 ATCACTTTGC ATGTCATTTG TTTGTGTCAT AACTACAACG GGCTTCATAT TTAAACGTCA 180
 CTCCATTATT TAATGTTGTT CATTTAAGCG TTTTATAATT TCATAAGCAC CTTGCTCTTT 240
 45 TAATTTGTTA CTCACTGTTT TGCCTAACTC AACCAGGATCT GTTCCGTTCA TTGTATATTC 300
 AAATCGTTCT TTACCATCTG GGGTCATAAT TAAACCTGTA AATTCGATTT CGTTTTGATC 360
 TGAGATTGTA GCATATCCTG CAATTGGCAC CTGACAACTA CCATCCATTT CTGCTAAAAA 420
 50 CGTTCGTTCA GCAGTCACAC ATTTTGCAAC CTCATCATTA TGTACTTTGC TTAATAATGT 480

55

	TAACAATGTA TCTCTATCAA GATAAGATGT TnCAATATCA TCTGACCAGC CCATTCTTCT	600
	TAAACCAGCT GCAGCTAAAA TAATCGCATC ATAATCTTCA GTTTGTAAGT TTTCTAATCG	660
5	TGTATCTATA TTACCTCTAA TCCATTTAAT CTCTAAATTA GGATACCTAG ATAATATTTG	720
	TGCACCACGA CGTAATGAAC TAGTACCAAT AATACTGCCT TCTGGCAATT GGGATAGTGG	780
10	TGTATGTGTT TTAGAAATAT ACGCATCAAA AGGTAATTCT CTATCAGGGA TACAACCTAA	840
	TGTTAAACCT TCCGGAATTA CACTTGGTAC GTCTTTAAGC GAGTGTATTG CCAATATCGAT	900
	ATTTTTTTCA AAAAGTTCAT GTTGTATTTT TTTAACAAAT AAGCCTTTGC CTCCGACTTT	960
15	AGACAATTGT TTATCTACTA TACGATCGCC TTTCGTGACA ATTTCTTTAA TTTCAATTTT	1020
	TAGATTTGGC TCGACAGCTT TTAATTTATC AATAAATTGC TGGCTTTGTG TTAAAGCTAA	1080
	TTTACyCTT CTGGAGCCAA CGACTTATTT ACGCATGTTT AATTCCTCCT AGGAACGGAT	1140
20	TGCTCTAGAT TATTTTCTCA ATTCACAAAA TGTGTTGCAA AAAATAAATT AATCATATTT	1200
	AAGCAAAATA AAATAATGTT ATAGTATATT AAATATCTTG AATTCAACCA TTTGTTGATT	1260
	CTAAGTAAAA TATAACTTCC ATATAATACT GTAATAATTG AAGAGAGTAT TACCTTCGGG	1320
25	TCAATGAATA TACGTTCAAC AACTGAAATT ACACCCCACT GTGTACCTAA AATAATACTA	1380
	AATATGAGAA TTATCCACCC ACTTAACGTT GAGTAAAACA CAATTGATTC AAGTGTAGCA	1440
	ACGCTACCAA TTCTAAAGTA TTTTGTATCA AAACGTTTTT CCTTCAAATT ACGGTATTGC	1500
30	ATGATATACA GTAATGCATT GACAAAAGCT AAGGCAAAGA AGACATAACT TAACACAGCT	1560
	AGACCGATAT GGAATAACAG TAACCTGTCT ACAACAGCAA TTTTCTGAAC CTTATTAGTA	1620
	TAATGTGTCG GTTGAAATGT ATTCATCCCT AAGAGTGTTA ACCCTATTAA ATTCCAAGGA	1680
35	AAAACACAG	1689

(2) INFORMATION FOR SEQ ID NO: 180:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

	nTGGnTGGCT TTTCTATTG GACCAAATGG ACCnTTTACC TGGCCnTTCC CAGGACACCC	60
50	CGCTTGCGCC CACATTCCAA TCGGAAAAGG TGTATGTGGT ACAGCCGTTT CAGAACGTCG	120
	TACACAAATT GTAGCTGATG TTCATCAATT CGAAGGACAT ATCGCTTGTC ATGCTAATAG	180

CGATGCCCCT ATAACGGATC GATTTGATGA CAATGACAAa GAaCATCTTG AaGCAATTGT 300
TAAAATTATT GAAAaGCAAC TCGCATAAAA GGACATCAGC ATTTTCAATA AAGTGTGAC 360
5 AGTTAGCAGG AAAATGTTAC AATAATCTTT GTGTGAATTA ACGAAAGTAG CAGTTGTATA 420
TTATTGAGCG CTATGTTGTT CCCAATGCGG ACGTGTACG TAACTGTCGC TATAAGGTGA 480
AGACACATAA AACAATATAT CTTAGTAAGC ATGCAACACT CTTTTTGTGTT TATTCATAAC 540
10 AACAAAAAAG AATTAAAGGA GGAGTCTTAT TATGGCTCGA TTCAGAGGTT CAAACTGGAA 600
AAAATCTCGT CGTTTAGGTA TCTCTTAAAG CGGTACTGGT AAAGAATTAG AAAAACGTCC 660
TTACGCACCA GGACAACATG GTCCAACCA ACGTAAAAA TTATCAGAAT ATGTTTACA 720
15 ATTACGTGAA AAACAAAAAT TACGTTACTT ATATGGAATG ACTGAAAGAC AATTCCGTAA 780
CACATTTGAC ATCGCTGGTA AAAAATTCGG TGTACACGGT GAAAACTTCA TGATCTTATT 840
AGCAAGTCGT TTAGACGCTG TTGTTTATTC ATTAGGTTTA GTCGTACTC GTCGTCAAGC 900
20 ACGTCAATTA GTTAACCAGG GTCATATCTT AGTAGATGGT AAACGTGTTG ATATTCCATC 960
TTATTCTGTT AAACCTGGTC AAACAATTC AGTTCGTGAA AAATCTCAA AATTAAACAT 1020
25 CATCGTTGAA TCAGTTGAAA TCAACAATTT CGTACCTGAG TACTTAACT TTGATGCTGA 1080
CAGCTTAACT GGTACTTTTCG TACGTTTACC AGAACGTAGC GAATTACCTG CTGAAATTAA 1140
CGAACAATTA ATCCGTTGAG TACTACTCAA GATAATACGG TCAATACCAA CACCCACAAT 1200
30 TGTGGGTGT 1209

(2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 698 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

AAATCCCTTt GTtaAaGtSc AAAtTTTTTc AACrgCTTTA AtArGACCCA TATTACctTC 60
TTGGATTAAA tCmAGGaATG AcATACCAG ACCaCGTATC TTTTAGCAAT ACTTACAAC 120
45 AAACGTAAGT TCGCTTCTGC AAGTCTTGAT TTTGCTACTT CATCACCTTG TTCAATACGT 180
TTGGCTAATT CGATTTCTTC TTGTGCACTT AATAAGTTAA CACGCCCAAT TTCTTTAAGG 240
TACATACGAA CTGGGTCAAT TATTTTAACA CCTGGAGGGG CACTAAGATC ACTTGGATT 300
50 AGTTTCTCGT CAGTATCTGA ACTATCTTTT TCATTAACTA GTGAAATATC ATTATCATT 360

5 GCAATTTCTT CATGACTTAA ATGACCCTCT TTTTACCTT TTTCAATTAA TTGCTTCTTA 480
 ACATCTTCTA ATGTTAATGT CGGATCAATT GTTTGTTTTT TAATTTTAAC TGTGTTATCA 540
 10 GACATGAAAC GGCCTCCCGA TTTTAAATAT GAACATTCGA AATTTATTCA ATATTGCTAT 600
 TTTAAACGAA ATTCTTAATT AATTCCATCC ATATTTTnAA TTTTATTTTA CAAATTGGGA 660
 ACTAAATCCC CAATATTTAT TTTTCAATAG TGGTGGTT 698

(2) INFORMATION FOR SEQ ID NO: 182:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5147 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

ACTTGATGAT GTATACAATG TATTTCAAGA ATATTATCAA AAAACATCTA ACATTAAGTT 60
 TTGTAGAATT CACAATTCTA GCTATTATCA CTTCTCAAAA TAAAAACATC GTTCTTCTTA 120
 25 AAGATTTAAT TGAACAATC CACCATAAAT ACCCTCAAAC TGTTAGAGCT CTCAATAATT 180
 TAAAAAGCA AGGCTATCTA ATAAAAGAAC GCTCAACTGA AGATGAAAGA AAAATTTTAA 240
 TTCATATGGA TGACGCGCAG CAAGACCATG CTGAACAATT ATTAGCTCAA GTGAATCAAT 300
 30 TATTAGCAGA TAAAGATCAT TTACATCTTG TTTTGAATA ATATCTCTAT TACGCAAGTG 360
 TGCTGTATTC TAAAGTGCAC TTGTGTTTTT TATTTTTTAA TAAACCTCA GCACATAATG 420
 AACAACTTTC TATTTTCTAT ATCACTTAAA ACCATTTCGG AAATTAAACC TCAGCACATT 480
 35 CAAAGCCCCA CTTTATTCTT AAAAATATTT TTTAACTCAT ATGTATTAAA CCGCTTTCAT 540
 TATAAAAAAT ATCTCTATAT TtTATCTGtT TtTATTAATC GAAATAGCGT GATTTTGCGG 600
 TTTTAAGCCT TTTACTTCCT GAATAAATCT TTCAGCAAAA TATTTATTTT ATAAGTTGTA 660
 40 AAACCTTACCT TTAAATTTAA TTATAAATAT AGATTTTAGT ATTGCAATAC ATAATTCGTT 720
 ATATTATGAT GACTTTACAA ATACATACAG GGGGTATTAA TKTGAAAAAG AAAACATtT 780
 ATTCAATTCTG TAACTAGGT GTAGGTATtG CATCTGTAAC TTTAGGTACA TTACTTATAT 840
 45 CTGGTGGCGT AACACCTGCT GCAAAtgctG CGCAACACGA TGAAGCTCAA CAAATGCTT 900
 TTTATCAAGT CTAAATATG CCTAACTTAA ATGCTGATCA ACGCAATGGT TTTATCCAAA 960
 50 GCCTTAAAGA TGATCCAAGC CAAAGTGCTA ACGTTTTAGG TGAAGCTCAA AAACCTAATG 1020
 ACTCTCAAGC TCCAAAAGCT GATGCGCAAC AAAATAACTT CAACAAAGAT CAACAAAGCG 1080

55

	AAAGTCTTAA AGACGACCCA AGCCAAAGCA CTAACGTTTT AGGTGAAGCT AAAAAATTAA	1200
	ACGAATCTCA AGCACCAGAA GCTGATAACA ATTTCAACAA AGAACAACAA AATGCTTTCT	1260
5	ATGAAATCTT GAATATGCCT AACTTAAACG AAGAACAACG CAATGGTTTC ATCCAAAGCT	1320
	TAAAAGATGA CCCAAGCCAA AGTGCTAACC TATTGTCAGA AGCTAAAAAG TTAAATGAAT	1380
	CTCAAGCACC GAAAGCGGAT AACAAATTCA ACAAAGAACA ACAAATGCT TTCTATGAAA	1440
10	TCTTACATTT ACCTAACTTA AACGAAGAAC AACGCAATGG TTTCATCCAA AGCCTAAAAG	1500
	ATGACCCAG CCAAGCGCT AACCTTTTAG CAGAAGCTAA AAAGCTAAAT GATGCTCAAG	1560
	CACCAAAGC TGACAACAAA TTCAACAAAG AACACAAAA TGCTTTCTAT GAAATTTTAC	1620
15	ATTTACCTAA CTTAACTGAA GAACAACGTA ACGGCTTCAT CCAAGCCTT AAAGACGATC	1680
	CTTCAGTGAG CAAAGAAATT TTAGCAGAAG CTA AAAAGCT AAACGATGCT CAAGCACCAA	1740
	AAGAGGAAGA CAATAACAAG CCTGGCAAAG AAGACAATA CAAGCCTGGC AAAGAAGACA	1800
20	ACAACAAGCC TGGTAAAGAA GACAACAACA AGCCTGGTAA AGAAGACAAC AACCAAGCCTG	1860
	GCAAAGAAGA CGGCAACAAG CCTGGTAAAG AAGACAACAA AAAACCTGGT AAAGAAGATG	1920
	GCAACAAGCC TGGTAAAGAA GACAACAAAA AACCTGGTAA AGAAGACGGC AACCAAGCCTG	1980
25	GCAAAGAAGA TGGCAACAAA CCTGGTAAAG AAGATGGTAA CGGAGTACAT GTCGTTAAAC	2040
	CTGGTGATAC AGTAAATGAC ATTGCAAAAG CAAACGGCAC TACTGCTGAC AAAATTGCTG	2100
30	CAGATAACAA ATTAGCTGAT AAAACATGA TCAACCTGG TCAAGAACTT GTTGTGATA	2160
	AGAAGCAACC AGCAAACCAT GCAGATGCTA ACAAAGCTCA AGCATTACCA GAAACTGGTG	2220
	AAGAAAATCC ATTCATCGGT ACAACTGTAT TTGGTGGATT ATCATTAGCC TTAGGTGCAG	2280
35	CGTTATTAGC TGGACGTCGT CGCGAACTAT AAAACAAAC AATACACAAC GATAGATATC	2340
	ATTTTATCCA AACCAATTTT AACTTATATA CGTTGATTAA CACATTCTTA TTGAAATGA	2400
	TAAGAATCAT CTAAATGCAC GAGCAACATC TTTTGTGCT CAGTGCATTT TTTATTTTAC	2460
40	TTACTTTTCT AAACAACCTC TGAAACGCCT CAACACTTTC TACTCTGATT ACATATATGA	2520
	CATTTTTAGG CATTAAAAAA TCGAACTAGA CAAGATGCTC ATTGCATTTT GTACTAGTTC	2580
	GATTCATGAA TAATTAGATT TAAAATGTCA TTTGAATCCA AGTGACAACA TTATTTATAT	2640
45	TTAGAATATT AACGTTAGTA TAAACGTCCA AACACAAATA AAAGCAACAA ATATAATACT	2700
	GTATTTTAAC GTCATTTTAA ATAATGCAGA TTCTTCACCA ACTTTTTTAA CAGCTGCAGT	2760
	CGCAATGGCA ATTGATTGTG GTGAAATAAG TTTCGCTGCT ACACCACCTG CAGTGTAGC	2820
50	TGCCACAAGT AATGAACCGC TTGTTGAAAT TTGTTGTGCC ACTGTCGCTT GAATAGGTGC	2880

55

	TGGAGAGAAT AATGGGAAAA TTGCTCCCGC TTTAGCAATA CCTGTGCCAA TTGCTACAGT	3000
	CAAACCACCG TATGTCATAA CTTTAGCAAT AGCTAGGATA GCTGAAATTG TAAGGATCGG	3060
5	TAACCATAAT TCTTTAATTG CTTGACCAA TAAAGCACCT GCACTTTTCC ATTTTAACTT	3120
	CGTAATTAAA ATTGTAATAA TTACTGTTAA TAAAATCGCT GTCCCAGTTG CACCAATTAA	3180
10	ATCGAGACGC AACGCAATTC CTTTAGGCGA TAAATCACTC ACAGTATTTG GAATTGGCAA	3240
	TTTTATTACT AAACCTTTCAA GTGCACCTCC AGGTTGGAAT AATTTTTTGA AGAATGGTGC	3300
	ACTCCATACT AATACAAAGG CAGTTAAAAT TACGAACGGA CTCCAAGCAA AGACAATTTT	3360
15	TTTAGGCGTT CGTTTTTGAA TTTTATGTTT AGACGCTTCC AATCTGAAAA TGTTTTTTCGG	3420
	TTTAAATTTA CGACAAACAA ATGCTAACAC CACCATTGTT GCTAGTGATG GAATAATGTC	3480
	TGCTAGTTCT GGACCATGGA ATATTGTTAA TAATAATTGT AATCCAGTAT ATGTACCACT	3540
20	CACTGTTAAA ATGACAGGTA AAATTTCTTT AATACCTTTC ATACCATCTA CAATGAATAC	3600
	TAAAACAAAT GGAATAATAA AGTTTAAAAT TGGAAGTGTT AATGCTGAGT ATCTCGCAAC	3660
	ATCTAATGTT GTAACGCCTC CACTTAAGTT AAACGTATCA ATAATACTAA CTGGTAAACC	3720
25	AATTGCACCA AAGGCACCCG CCGCACCATT AGCAATTAAA CATAACATCG CTGCTTTTAA	3780
	TGGTTCAAAT CCAAGTTGAA TTAATAATAC TGCACAAATC GCAATTGGCA CACCAAATCC	3840
30	TGCTGCACCT TCTAAAAATG CGTTGAAACA AAATCCAATT AATAATAGTT GGATTCTTTG	3900
	GTCCACTGAA ATACTTGCAA TACTATCTTG AATAATAGAA AATTGTCTCG TTTTAATAGA	3960
	AACCTTATAT AACCAAACCTG CCATTAAAAC GATATATCCT ATTGGGAAAA TACCGGCAAC	4020
35	AACGCCTTCT GTAATCGCAC CTGCTGATAC ACGCGCTGGT AATTCAAATA CAAATAAAGC	4080
	CACAATCAAT GTAACAACCA AAGTTGTCAA TGCTGCATAA ATGCCCTTCA TTTTAAAAAC	4140
	GGTTAAGCAT AATAAAAATA AAATAATAGG TACTGCTGCA ACTAAGGCTG ATAATCCGAC	4200
40	ATTATCGAAT GGATTTACAG TAAGTAGTGT CATAATGACT CCCTCTCTTT ATATAAATA	4260
	TTTATCATTC TGATTAATCT ACAACCTATT TCAACTTATA TTTTGGGATG ATCACATATT	4320
	TAAATGTAA CACTCCTATA TGTGACAGGC AATCGAATTT TTACAAAAAG TTCACAAAAT	4380
45	ATACACAATA TTTAACTATA ATAmATAATA TATCaTntTA ATTATAAATA CTAGATATTA	4440
	TTTATAATAA TCTCAGGAAT TCGCTTCAAA ACTGCATCAT GAGAGTTTAT ATTTTTATTG	4500
50	AGAATCTCTC ATTTTATGAA TTGTAGGAAG TAAACAAAAT ATGACAAGCG TCAAACCAAT	4560
	GATAATGATA AATATCATAT TAAACCATAG TAAATTGAAT TGATGATGGT GTTGTATTG	4620
	CCAAATTTCT AATACTGTGA AGATAGACAT ATAGCTCATA ATCTCTAAAT TTAACGTACT	4680

55

AAATCGTTCA TAGTATCTAC CTGCAATGAA AAATATAAGC CAAATCACTA TAAATGCGCT 4800
 ATTAATCAAA AGCAGCACCC ATTTATCAGC AAAATTATCA GCATCCCCTG CTAAATTATA 4860
 5 ATGAATAGGC ACTTTGGTTG GTAATTTTGG ATAGGTCACT ACTGTATAGC ACATCATAGC 4920
 TAAGTAAATA AGTAGACTTA ATATTGTAAA AGACCTGATT TTAGACATTC TATCGCCTcT 4980
 TcTTTACATT TTATGTATAA CACTCTGCCT ATTTTACCTT TTAATaCATT ACCCCAACGA 5040
 10 TtAAaCAATA tGTAAaTGATA CTATAATTGC GTCAGGAGTA TCCGCTTGTT AAATGTGCAT 5100
 AGCTTATATT TAGCTGTTTA ACATGCCACA TAATGATTCTG AATTATT 5147

15 (2) INFORMATION FOR SEQ ID NO: 183:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1312 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

25 CACTTACTTC CACCATTATC ATAACCTTAA AATGGATATA nttCATCAA CATTATCTAA 60
 AGGCGTCGCA CCTACACCAA CACCATCCAA CAATTAACCT ACAACTCTGC GATTACTTCT 120
 TCAGCAGCAA CTTTCACnTG CGTAATACAA TCAGGTAGTC CAACCGCTTC AAAAGATGCA 180
 30 CCAGTTACTC TAAGTCGTGG ATATGTTTGT TTAATATGTG CTTGAATCTG TCTAATTTGT 240
 TGAATATGAC CGACATGGTA CTGTGGCATA CTTTTCGGCA AACGATTGAC AATTGTAAAT 300
 35 TCAGGATCAC CTTTAAATGT CATCATTTGA CTTAAATCTC TACGTACAAT CGATACTAAT 360
 TCATTATCTG TATGATCATC AACCACAGTA TCACCTGGTT TACCTACATA CGCACGAATC 420
 AAAACCTTAC CTTCCGGTGT AGTAAATGGC CATTTTTTTCG ATGTCCAAGT ACATGCGGTA 480
 40 ATGTCTGTAT CACTCGTTCT CGCAATTACG AAGCCAGTAC CATCATGGGT ATTTTCAATG 540
 TCTTTTTTCAT CAAATGCCAA TACAACAGTT GCAACAGTCG TACTATCCAT CGTTTTAAAG 600
 TAATCAAATG CTGGATCTTG TCCGAACCAA TTTAAAAACA CTTGATGTGG TGTCGTTACT 660
 45 AATACGCCAT CATACACTTC TTCTAGTTGA TCAITGTAAA CAATTTTATA TTGTTTTTGA 720
 GATGTAATTA TATCATCCAC TGACGTATTG TAGCGTATTG TCACACCTTT ATTTTAAACA 780
 50 TCTTGTTCTA ATGCTTCAAT AAATGAGCTT AAACCATGCT TAAATTGTTT GAATTGTCCT 840
 TTCGGTGCGC CAGGATATAA TTGTCTTTGT TTCAGACGCT TATTTTTCTC ATCCTTCATA 900
 CCTTTTATCA GACTTCCGAA TGCCTCTTCT TTTTCTTTAA AATTAGGAAA CGTACTCATC 960

55

TCAAGTACCT CATTACCTAA TCTTGCTCTG AAAAATGCAC CAACAGAAAT GTCACCATCC 1080
 TGCATTTGAG TAGGTTTTTT TAATAAATCA AACCCTGCTC TTAATTTACC AAGTGGCGAT 1140
 5 ATTAATTTTG TAGTAACAAA TGGTTTAATA TCTGTTGGAA TACCCATAAT TGAACCACCT 1200
 GGAATCGGAT ATAATTTATT TTTCGCAAAA ATATATGATT GTCCAGTCGT ATTTGTAACA 1260
 10 ATATCTTGTT CTAATCCAAT ATCTTTCGCT AATTCTGTCA TAATCGTTTT TC 1312

(2) INFORMATION FOR SEQ ID NO: 184:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6157 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

TTTTACAATA AAAATATGAT ATACTACTTG TCGTATATAA GGAACGGAGG ACAATTTATG 60
 CATACATTTT TAATCGTATT ATTAATCATT GATTGTATTG CATTAAATAAC TGTTGTACTA 120
 25 CTCCAAGAAG GTAAAAGCAG TGGACTTTCA GGTGCCATCA GTGGTGGTGC TGAGCAGTTA 180
 TTCGGTAAAC AAAACAACG TGGCGTCGAT TTATTCCTAA ATAGATTAAC AATTATTTTA 240
 30 TCAATATTAT TTTTGTACT TATGATTTGC ATAAGTTATC TTGGTATGTA AGGTCCGGCG 300
 ATGTAAATGT CGGGCTTTTT TATTTATAAT TAAGAATGTA ATAGTTTAAC AATAAGCTAT 360
 GTAAATATA TAGCCTAGTT AAGTATGCAA AGGGAGCGTT AGATTTATGC AGATAAAATT 420
 35 ACCAAAACCT TTCTTTTTTG AGGAAGGTAA ACGTGCCGTG TTATTACTAC ATGGTTTTAC 480
 AGGCAATTCG TCTGATGTTT GTCAATTAGG TCGATTTTA CAAAAGAAAG GTTATACATC 540
 ATATGCACCG CAATATGAAG GCCACGCGGC ACCACCAGAT GAAATACTGA AATCTAGTCC 600
 40 TTTGTTTGG TTAAAGATG CGTTAGATGG TTATGATTAT CTTGTTGAAC AAGGTTATGA 660
 TGAAATTGTT GTTGCTGGTC TATCATTAGG TGGGGATTTT GCTTTAAAT TAAGCTTAA 720
 TAGAGATGTA AAGGGTATTG TAACGATGTG TGCTCCTATG GGTGGCAAAA CTGAAGGTGC 780
 45 CATTTATGAA GGCTTTTTAG AATATGCACG CAATTTTAAA AAGTATGAAG GTAAAGATCA 840
 AGAGACTATT GATAATGAAA TGGATCATTT TAAACCAACT GAACTTTAA AAGAATAAG 900
 50 TGAAGCATTA GATACGATTA AAGAGCAAGT TGATGAAGTG TTGGATCCTA TTTTAGTGAT 960
 TCAAGCAGAA AACGACAATA TGATTGATCC ACAATCCGCA AATTATATAT ATGACCATGT 1020
 AGATTCTGAT GACAAAAATA TCAAGTGGTA CAGTGAATCT GGACATGTTA TTACGATTGA 1080

	AGAATAAAAA	GAGATTTTAA	CATTAGAAAAG	GAGGGGCATA	ATGAATTTAA	AGCAATCTAT	1200
	AGAAGAGATT	ATTAATCAAC	CTGAATATGA	ACCTATGTCA	GTGTCAGATT	TTCAAGATGC	1260
5	ATTAGGTTTA	AGCAGTGCCG	ACTCGTTTAA	AGATTTAATT	AAGGTGCTTG	TGGAGTTAGA	1320
	ACAATCAGGA	TTAATCGAAC	GTACAAAAAC	AGACAGATAC	CAAAAAAGC	ATAGTTATAG	1380
	AGGTCAATCA	AAATTGATAA	AAGGAACGTT	AAGTCAAAAT	AAAAAAGGCT	TTGCATTCTT	1440
10	AAGACCTGAA	GATGAGGATA	TGGAAGATAT	ATTTATTCCC	CCGACGAAAA	TTAATCGTGC	1500
	CTTGATGGA	GATACTGTTA	TTGTAGAAAT	CCATCAATCA	AAAGGTGAAC	ATAAAGGTAA	1560
	AATCGAAGGG	GAAGTTAAGT	CGATTGAGAA	GCATTCTGTA	ACTCAAGTTG	TTGGTACGTA	1620
15	TAGTGAAGCT	AGACATTTTG	GCTTTGTTAT	TCCGGATGAT	AAACGTATTA	TGCAAGATAT	1680
	TTTCATTCCCT	AAAGGTCAAA	GTTTAGGCGC	AGTCGATGGT	CATAAGGTAC	TTGTACAAAT	1740
20	TACTAAGTAT	GCTGATGGTT	CAGATAATCC	AGAAGGACAT	ATTTCTGCTA	TTTTAGGACA	1800
	TAAAAATGAT	CCTGGCGTAG	ATATTTTATC	TATTATCTAT	CAACATGGCA	TAGAAATTGA	1860
	ATTTCTGAT	GAAGTGTTAC	AAGAAGCTGA	AGCAGTACCT	GATCATATTG	AAAATACTGA	1920
25	AATTAAAGGC	CGTCATGATT	TACGTGATGA	ATTGACAATC	ACAATTGATG	GTGCTGATGC	1980
	TAAAGACTTA	GATGACGCAA	TTAGTGTTAA	AAAGTTAGCG	AACGGTAATA	CGCAATTAAC	2040
	TGTAAGTATT	GCTGATGTCA	GCTATTATGT	AACAGAAGGT	TCTGCATTGG	ATAAAGAGGC	2100
30	ATATGATAGA	GCGACAAGTG	TATATCTTGT	TGACCGTGTA	ATTCCAATGA	TTCCACATCG	2160
	ATTAAGTAAT	GGTATTTGTT	CATTGAATCC	TAATGTTGAT	CGTTTAACTC	TAAGCTGTGC	2220
35	CATGGAAATC	GATGCTAGTG	GTCCGTTTGT	TAAACATGAA	ATTTTGTGATA	GTGTTATACA	2280
	TTCTGATTAT	CGAATGACGT	ATGATGCGGT	AAATCAGATT	ATTACTGAAA	AGGATCCTAA	2340
	CATTCGCGAA	CAATATAATG	AAATTACGCC	TATGCTAGAT	TTAGCACAAG	ATTTATCTAA	2400
40	TCGTTTGATT	CAAATGAGAA	AACGACGTGG	TGAAATCGAT	TTTGATATTA	GTGAAGCAAA	2460
	AGTATTAGTT	AACGAAGACG	GTATACCAAC	AGATGTTCAA	TTAAGACAAC	GTGGCGAGGG	2520
	TGAACGTCTA	ATTGAATCAT	TTATGTTAAT	TGCAAAATGAA	ACAGTTGCTG	AACATTTTAG	2580
45	TAAGTTAGAT	GTACCTTTTA	TTTACCGAGT	GCATGAGCAA	CCTAAATCAG	ATCGCTTAAG	2640
	ACAATTCTTT	GATTTTATTA	CAAACCTTGG	CATCATGATT	AAGGGTACTG	GCGAAGATAT	2700
50	TCATCCAACA	ACACTTCAAA	AGGTTCAAGA	AGAAGTAGAA	GGTCGACCTG	AACAAATGGT	2760
	CATTTCAACA	ATGATGTTGC	GTTCAATGCA	ACAAGCGCAT	TATGATGATG	TGAACTTGGG	2820
	ACATTTTGGC	TTATCAGCTG	AATATTATAC	GCAITTTTACA	TCACCAATTA	GACGTTATCC	2880
55							

	AGAAGTGAAG CGTTGGGAAG ACAAATTGCC TGAGTTAGCT GAACATACTT CTAAACGTGA	3000
	ACGTCGTGCT ATTGAGGCAG AACGTGATAC TGATG a ATTG AAAAAAGCAG AATATATGAT	3060
5	TCAACATATT GGTGATGAAT TTGAAGGTAT TGTCAGCTCA GTAGCTAACT TCGGTATGTT	3120
	CATTGAATTG CCAAATACGA TAGAAGGTAT GGTTCATATT GCGAATATGA CTGATGATTA	3180
	TTACCGTTTT GAAGAGCGTC AAATGGCATT AATTGGTGAG CGTCAAGCTA AAGTATTTAG	3240
10	AATTGGTGAC ACAGTTAAGG TTAAAGTGAC GCATGTTGAT GTAGATGAAC GATTAATTGA	3300
	TTTTCAAATT GTAGGTATGC CTTTACCGAA AAATGATCGA TCACAGCGCC CAGCGCGAGG	3360
	TAAGACAATT CAAGCCAAAA CGCGTGGTAA ATCATTAGAT AAATCAAAAT CTGATGATAA	3420
15	GGGTCGTAAG AAAAAAGGTA AGCAACGTAA AGGTAAAAAC CAACGTAATA ATGATAAATC	3480
	AGGTAATAGT AAGCATAAGC CATTTTTATAA AGATAAAAGT GTGAAAAAGA AAGCACGTCG	3540
20	TAAGAAAAAA TAAGCAGCAA TGAGGTGAGT ATGAATGGCT AAGAAGAAAT CACCAGGTAC	3600
	ATTAGCGGAA AATCGTAAGG CAAGACATGA TTATAATATT GAAGATACGA TTGAAGCGGG	3660
	AATTGTATTG CAAGGCACAG AAATAAAATC AATTCGCCGA GGTAGTGCTA ACCTTAAAGA	3720
25	TAGTTATGCG CAAGTTAAAA ACGGTGAAAT GTATTTGAAT AATATGCATA TAGCACCATA	3780
	CGAAGAAGGG AATCGTTTTA ATCACGATCC TCTTCGTTCT CGAAAATTAT TATTGCACAA	3840
	GCGTGAAATC ATTAAATTGG GTGATCAAAC ACGTGAGATT GGTTATTCTGA TTGTGCCGTT	3900
30	AAAGCTTTAT TTGAAGCATG GACATTGTAA AGTATTACTT GGTGTtGCAC GAGGTAAGAA	3960
	AAAATATGAT AAACGTCAAG CTTTGAAAGA AAAAGCAGTC AAACGAGATG TTGCGCGCGA	4020
35	TATGAAAGCC CGTTATTAAG CGATTTAGTT GCTTAATCGG GCTATATTTG ATATAGTTAT	4080
	ATGTGCTTTT GTAAATTACA AAAGTATGAT TTGTTTGATT TATTATTTTCG GGGACGTTCA	4140
	TGGAFTCGAC AGGGGTCCCC CGAGCTCATT AAGCGTGTGCG GAGGGTTGTC TTCGTCATCA	4200
40	ACACACACAG TTTATAATAA CTGGCAAATC AAACAATAAT TTCGCAGTAG CTGCCTAATC	4260
	GCACTCTGCA TCGCCTAACA GCATTTCTTA TGTGCTGTTA ACGCGATTCA ACCTTAATAG	4320
	GATATGCTAA ACACTGCCGT TTGAAGTCTG TTTAGAAGAA ACTTAATCAA ACTAGCATCA	4380
45	TGTTGGTTGT TTATCACTTT TCATGATGCG AAACCTATCG ATAAACTACA CACGTAGAAA	4440
	GATGTGTATC AGGACCTTTG GACGCGGGTT CAAATCCCGC CGTCTCCATA TTTGTAGCCT	4500
50	ACAGCCTTTG TGGTTGTGGG CTTTTTTATT TTGTGTTTTT CAGGGGATAA TGCATTGCAG	4560
	AATTTGTTGT GAGTATTGAT ATAGCAGTGT TTGTATAGGT GTTTATTTGA TGGAGGAAAG	4620
	AGTAATAAGT GATTATGAAT TAGTTTTTGA GATATAAGGG GACAGTGATG TGTGTCAAAT	4680

55

	TTATACGCAA AAAATTCTCC ATGTTATATA TGTCAATATA AAAATGTGAA TCGTCTACAC	4800
	TTAATTGGAT AAATGGCTAC TGAAAAAGAA CTTTTCATTT TTGTTACGTC ACTAAGTGGG	4860
5	TGTAGTTATA AAGAGATGAG CCGAGTTTTG ATATTTTCAT TAGAATCAAT ATGCCTATTA	4920
	ACACAATCAG CAATAGTTGA CGAGACGGAA ATAAAAGAAG TCGTAGTTAA GAAATGCATT	4980
	TCACAACATA CCATTGTAGC CATTTTTTATT GTTTTGGATG ATAAACTCTT TTTGGAATTT	5040
10	TTAGTTTTTA TAATTTGCAA CTACACTACT TCTTTTACTA ATATTAATGT CTAAGTAATC	5100
	GATAAAAAAT TTTCCATTGA ATAAATGAGA AGTTAAAAAC TTTACTTAAC CTTTCyCATT	5160
15	GCATTTTCCT ATTCACGATT TTAAGAACCC AACATACTAC AAACGAATTT TAAAAGGCGA	5220
	GAGTAAAGCT TACTTGTTTA TTATACATAT TTAAAATCCA AGAGTCAGAA CAGACTACTC	5280
	CTCTTTATAA CTATAAAAAA TAGCTATGAA AAAATCTATC GTCATAGATT CCTTCATAGC	5340
20	TAATCTTAGT ATGTTTATTT TTATTTTAGG ATGCTATTTA TCAACTCAAC ATATAACTCA	5400
	CTATTTTTAT AACCTTCTAA TATATCATTa ACTTGTCTAA TAGGTATTTT TGGTACTTCT	5460
	CTAATGTTTT CCAATTTTGT TTTAAATTGT TTTTTTGTTA TTTGCTCTTT ATTTGTAGCC	5520
25	AATTGGAACA AGTAAGAATC TAGCATATTA ATTTCTTTAT ATGAATACAT ATATCTTAAT	5580
	AACACTAAAT CTCTAGTTTT TAAGTTAGGC GCTAGTTCTT CTGTGAATTG TTCTATTGAT	5640
30	TGTYTCATTa ATAACAATCT CATTTCTAAT TCTTCATTAT TCATTTTATC ACACTCTTtT	5700
	TATATTAATG CTTGACCAAC TTGGGAAACC CAAAACCCTA TGCTTCTTGC AGTAGAATCT	5760
	TTAATACCAG TTCCCATCAA TGCTTGTGAA ACTTGACCTT GTACATTTCC CCATGTAGCC	5820
35	TCTTCTTGTT TTAATGCATT ATTCAATGCG GGATTTACAA ATTTATCCCA TCTTTTTTTT	5880
	ATGATTTTCC GGCACGGGGA CTGATTTCTT TAACACCATT AAACACAGAT TTTTTATTTT	5940
	TAATCATAGC TTTATAGTAT CATGTTGGCT AAGCTATAAA TAAGTCAGTT TCTCTAAAAA	6000
40	TTAAATAACT GAATGTAAGA CAATCAACAA WCCAAATTTA TACTTCATCT AAACCACTGT	6060
	GGTCGTCATC TTTTGTCTT TCTTTTTCTT TCTCTCGTTC TGTTCTTTT TTGTACTCTT	6120
45	CTTCAAATTC TTTTCTTTT TTTTCTACTT CTTCTCT	6157

(2) INFORMATION FOR SEQ ID NO: 185:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 884 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

CATTGTGTTAT TCTGAGTAGC CAATTTGGCA AAGATGAACA AACGTCTGAA CAAACGTATC 60
 AAGTTGCAGT CGCATTAGAG TTAATTCATA TGGCAACACT TGTTTCATGAT GACGTTATTG 120
 5 ATAAAAGCGA CAAGCGTCGA GGCAAGTTAA CCATATCAAA GAAATGGGAT CAGACAACCTG 180
 CTATTTTAAC TGGGAATTTT TTATTGGCAT TAGGACTTGA ACACTTAATG GCCGTTAAAG 240
 ATAATCGTGT ACATCAATTG ATATCTGAAT CTATCGTTGA TGTTTGTAGA GGGGAACTTT 300
 10 TCCAATTTCA AGACCAATTT AACAGTCAAC AGACAATTAT TAATTATTTA CGACGTATCA 360
 ATCGCAAAAC AGCACTGTTA ATTCAAATAT CAACTGAAGT TGGTGCAATT ACTTCTCAAT 420
 CTGATAAAGA GACTGTACGA AAATTGAAAA TGATTGGTCA TTATATAGGT ATGAGCTTCC 480
 15 AAATCATTGA TGATGTATTA GACTTCACAA GTACCGAAAA GAAATTAGGT AAGCCGGTCC 540
 GAAGTGATTT GCTTAATGGT CATATTACGT TACCGATTtTT ATTAGAAATG CGTAAAAATC 600
 20 CAGACTTCAA ATTGAAAATC GAACAGTTAC GTCGTGATAG TGAACGCAAA GAATTTGAAG 660
 AATGTATCCA AATCATTAGA AAATCTGACA GCATCGATGA GGCTAAGGCA GTAAGTTGCA 720
 AGTATTTAAG TAAAGCyTTG AATTTGATTT CyGaGTTACC aGATGGACaT CCGaGatCAC 780
 25 TACyTTTAAG TTTGACGAAA AAAATGGGTT CAAnAAACAC GTAGTATTTA TgNAAAAGTA 840
 TTGAAAGCGC TTTACCAACC TGTTAATATA TAATAGTAAT ATAC 884

(2) INFORMATION FOR SEQ ID NO: 186:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6876 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

AATTTTCATCT GCTCGTGCAA AATCTTTGTT TTTCCTTGCT TCATTACGCT CTTGATTAA 60
 TTTTCAACA TCTTCATCCA ATAATTCATC TGCATTTTTA GATTTTAACG GTACACCTAA 120
 AACATCGCTG AAAATTTGAT AAATGCTTTT AAATTTATCA ATTACTTCTG TTGATGTTGT 180
 45 GTTCTCTAGT ACATATTTAT TCGCAAGTKt TGCTAAATCA TACCAAGCTG TAATTGCATT 240
 AGCTGTATTA AAATCATCAT TCATAACTGT TTCAAAACGA TTTAAATCG CATCAATTTG 300
 ATCAATATAT GTCTGTTGAT TTTCAATATT AGTAGCAATT TGTGCGCGCT CTTCAATTAA 360
 50 TTGATAACTA TTGCGAATAC GCTCTAGTcC aCTACGTGCT GATTCTACCA ATTCTAGATT 420
 ATAGTTAATT GGGCTTCTAT AATGTACGCT AATCATAAAG AATCTTAGTA CATCTGGATC 480

	ATTATCAATA TTAATGAAAC CATTATGCAT CCAATAATTA GCAAATGGCG CATGATTATG	600
	TGCTTCTGAT TGTGCTATTT CATTTTCATG ATGTGGAAAT TGTAATCTG AACCACCCGC	660
5	ATGTATATCA ATTGTAGGTC CTAGCTCATG AAATGCCATT ACAGAACATT CTATATGCCA	720
	TCCTGGTCTA CCTTCACCAA ATGGGCTATC CCAACTAATC TCGCCAGGTC CGCTTTTTTC	780
10	CACAATGTAA AATCAAGTGC ATCTTCTTTA TGCTCTCCTG CATCTATACG AGCACCCACT	840
	TTTAAGTCAT CTATGGATTG ATGACTTAAT TTACCATAAC CTTCAAATTT ACGTGTCTA	900
	AAGTAAACAT CGCCACCACT TTCATATGCA TAACCTTGAT CCACCAAATC TTTAATAAAT	960
15	TGAATAATGT CATCCATATG GTCCATTACC CTTGGATTTC AAGTCGCTTT TCTAACATTT	1020
	AACGCACCAA CATCTTCATG AAAAGCAGCG ATATATTTTT CTGCAATTTT GGGAACAGAC	1080
	TGATTTAATT CTTGAGAACG TTTAATTAAT TTATCATCTA CGTCTGTAAA ATTTGATACA	1140
20	TATTCTACAT TATATCCTTG GTATTCAAAG TAACGTCTCA CTACGTCATA ATTAATTGCW	1200
	GGTCTTGCGT TACCAATATG AATGTAGTTA TATACAGTAG GACCACATAC ATACATTTTT	1260
	ACTTTCCCTG GTTCTATAGG CTTGAACACT TCTTTTTGAC GTGTAAGCGT ATTATATAAT	1320
25	GTAATCATCT TGAATCTCTC CATTCCTAGT CTTTTCAAGT TGTCGTTCTA AATGCTTAAT	1380
	TTGTTTCATAA ATTGGATCAG GTAGATGGCG ATGATCAAAT GTTTTTCCAA CTCGAACACC	1440
30	ATCTTGCTTA ACAATATGTC CTGGTATACC AACAAACGTT GAATAACTTG GAACTGATTG	1500
	TAAAACAACCT GAATTTGCAC CAATATTTAC ATTTGAATTT ATTTTAATAT TTCCTAAAAC	1560
	TTTCGCACCG GCTGCTATTA AAACATTGTC TCCTATATCT GGGTGTCTTT TCCCTCTTTC	1620
35	TTTCCCTGTC CCACCAAGTG TCACGCCTTG ATAGATTGTC ACATTATCAC CAATTGTACA	1680
	TGTTTCTCCT ATTACAACGC CCATACCATG ATCTATAAAT AGACGCTTTC CAATTTTAGC	1740
	ACCTGGATGG ATTTCTATAC CTGTGAAAAA TCTTGAAATT TGAGATATCG CGCGTGCTGC	1800
40	AACATATTTT TTTTGGTTGT ATAACCTATG TGCAATCAAA TGACTCCAAA CTGCATGTAA	1860
	ACCTGCATAC GTTGTAATGA CTTCTAATGT TGAACGTGCC GCTGGATCCT GCTCAAATAC	1920
	CATTTTTATA TCGTCTCTCA TTCTTTTTAA CAAGATCATT TCCTCCTCAA TGATTGAACT	1980
45	ACGTAAATAC ATAATTGAAG TACCTGCGAA ATTAAATATC AAAAAAGCAC CACTAACATA	2040
	CAAATTGTAT TGTTAGAGGC GCTTCCGCAC GGTCCACTC TGAATTTAGC GAATAACATT	2100
50	AATAATATTG CGGGCGCTTC CAAATTATCA AGGAAACTAA GTCAACTTAA TGCTCATCAC	2160
	TCTCATTATA TATTTAATTC ATTTTACGAA GGTGCATTCA TTAATTTCTA CGTTGTACTC	2220
	ACAGCAACCG TACACTCTCT GCATCGTATA AATTTAATTA CTAATCCTTC GTTTTATATA	2280
55		

	ATAAAATTCA AGTATATACT ACCTTGATCT TGTCTATTTT ATTACTTATA TTGTTTTTAAA	2400
	CGGTTTAGCA CTTTTTCTTT ACCAAGTACT TCAATTGTAT TTGGTAATTC AGGACCATGC	2460
5	ATTTGGCCTG TTACAGCAAC ACGAATAGGC ATAAATAATT GCTTGCCTTT TATTCTGTGT	2520
	TCTTTTTGAA CTTCTTTAAT TGTCTTTTTA ATTTTCAGCCG CTTCAAATGG TTCAAGTGCT	2580
10	TCTAATTTAC TGAATAAGTG CGTCATTAAC TCTGGTACTT GCTCTCCATT AATCACTTGT	2640
	TGTTCTTCTT CACCAAGAGC TGGCATTCTT TTAAAGAACA TTTCTGATAA AGGTACAATT	2700
	TCACCGGCAT AACTCATTTT TTTTGTATAA AGCGCAATTA ATTTGCGTCC CCAAGATAAA	2760
15	TCCTCTTCTG ACGGCACCTC AGGAATCAAA TTTGCTTTAA TTAAATGAGG TAATGCTAAT	2820
	TGGAATACTG TTTCACTATC TTTTGTGTTT ATATATTGGT TATTAACCCA TGCTAATTTT	2880
	TGCTTATCGA AAAATGCTGG TGATTTTGAC AAACGCTTTT CATCAAAGAT TTTGATAAAT	2940
20	TCTTCTTTAG AAAAGATTTT TTCTTCACCT TCAGGAGACC AACCTAATAA CGCAATAAAA	3000
	TTAAATAACG CTTCAAGTAA ATAACCTAAG TCACGATATT GCTCAATAAA TTGTAAAT	3060
	TGCCCCATCAC GTTTACTTAA CTTTTTACGT TCTTCATTAA CAATTAATGA CATATGACCA	3120
25	AAACGAGGTG GCTCCCAGCC AAATGCTTCA TAAATCATAA TTTGTTTAGG CGTGTGTTGAA	3180
	ATATGATCAT CACCACGAAT TACATCTGAA ATTTGCATGT AATGATCATC TATAGCTACT	3240
30	GCAAAATTGT ACGTTGGAAT GCCATCTTTT TTTACGATAA CCCAGTCACC AATACCATTT	3300
	GAATCAAATG AAATATTTCC TTTTACCATA TCATCAAATG AATACGTTTG GTTTTGAGGT	3360
	ACTCGGAAAC GAATTGATGG TTGGCGTCTT TCTGCTTCAA ATTGTTGACG TTGTTCTTCA	3420
35	GTCAAATGCG CATGTTGACC ACCATAGCGA GGCATTTTAC CACGAGCGAT TTGCGCTTCA	3480
	CGTTCAGCTT CTAATCTTTC TTCTGTCATA TAGCATTTAT ATGCTTTATC TTCTGCTAGT	3540
	AACGATCTA TTAATGGTTG GTAGATATGT TGACGTTTCA ATTGACGATA TGGTCCGTAG	3600
40	CCATTGTCTT TATCTACAGA CTCATCCCAA TCTAATCCTA ACCATTTAAG ATTATCAAAT	3660
	TGTGATGTTT CTCCATCTTC TAAATTACGT TTTTATCAG TATCTTCAAT TCGAATCACA	3720
	AAATCTCCGT TGTAATGTTT AGCATACAAG TAATTGAATA ATGCTGTTCT TGCATTACCA	3780
45	ATATGAAGAT ACCCAGTTGG ACTTGGTGCA TATCTTACTC TTATACGATC GCTCATTTT	3840
	TTCACTCCTA AATTAAATAT CAGATTTTCA AGTTAGTTCA TATAAATTGT TCATTTGCTA	3900
50	TCTTCGACCG TCATAACAAA TGTCTAACTC GTCTTATTGT TAAAACGAAA CAATGCTTTT	3960
	TAACATGACC TTAAATAAAT TTCATTGTTT AATCATAACA TAATCCCTG GGTAATATGC	4020
55	TTAAATTTTA AATAGAAAGC TGTGTTTTT TCAACACTTT AAAAAAGCTA TCCCTAAGAA	4080

	TTAAACTTCA AATTAACTAT TCAAATACGT TAAAATTGAT TCTAATTTTG TATGTCTTGA	4200
	TTGCTATAAG AATAACTTTA TTAATATCTA AAATTTAACA CTTAATGAAC TTGTTTCAAT	4260
5	GATATATTAG CACTATTTGT ATTTTTTGAT AACTAATATG TTTTGCATT TTTTATAGTT	4320
	ATACTTCAAA TTACAAACTt CGCCATTTC A TATACCTTTT AATATCTATT TTGTTTTCGT	4380
10	CAACTACAGT TTTTATAATG ATACTGTATC TTCGATTTTT TTAGCAAAAA CAATTCTTCC	4440
	TGAAGATGTT TGCAATAAGC TGACTIONC TAAATTGACA TGACTIONC TAAGATTTTT	4500
	AGCATTATCA ACAACTACCA TCGTACCATC ATCTAGATAT CCTACTGCCT GACCAGGCTC	4560
15	CTTACCCATT TTTGTGAGTA AAATATGCAG TTGATCACCT TGATGTACAT TAGGTTTGAT	4620
	TGCTTCTGAT AAATCATTAA CATTTAATGC TTTGATACCA TGTACATGAC AAACCTTATT	4680
	TAGGTTGAAA TCTGTCGTTA TAATACTTGC ATGATATTGT TTTGCTAATT TTAATAACAT	4740
20	CGTATCAATA TCACTATGTG TTTTAGTTGG ATGTATAACC TTTGTAGGAT AGTCTAAATC	4800
	ATACAATTCA TTTAAAATAT CTAAGCCTCT TTTACCCTTT TCaCGTTTAA CACTGTCAAT	4860
	TGAATCTGCA ACAATTTGTA ATTCAATTAAT AACACCTGT GGAATTAAAA TATTGCCATC	4920
25	GATAAAACCG CAACGAATGA CTTCTAAAAT ACGACCATCA ATAATTGCGC TTGTGTCGAT	4980
	AATTTTTGGC GTAGcaCTTT TaGTATGTTG TGACATGGAA CGCGCTATAT TCTCAGGTAA	5040
	AAACATTAAC ATTTTCATCTC GTTTTTTAAG GCCAAATTGG AAACCGAAAT AACATAGTAA	5100
30	TATCGTAATT ATGACAGGAA TGAAATGATT AAAAATAGAG TTGCCAATTG ATTCTAATAT	5160
	AAACGACACC ATAACAGAAA TAAGTAATCC GATTATTAAA CCTATTGTTG CGAATAGTAT	5220
35	TTCAACAGCA CTTCTACGCA TAATAAAATG TTCTAAACCT TTTATAGCGT TAGTAACTCG	5280
	TCTAATAAAT ACACCAAAAA TTAAGAACAT AAAAATACTA CCGATAATGC CATCTACATA	5340
	GTGATTTTTT AAAAAGCTGG AGTTTTGTAA TCCAAGATCA TTTGCAATTT CAGGAATAAT	5400
40	AATTATTCCT AATGCGCTCC CAATAATTAA GTAAATAATA ATAACCATTA GTTTAACGAT	5460
	ATTCACACAA TGTCTCCTT TCTTGATGTT TTATGAATGA AGAGCAAATG ACAATACTTC	5520
	ATGTACAGTA GTTACACCTA TTACTTGTAT ACCTTCAGGA TATGTCCATC CGCCTATATT	5580
45	ATTTTTAGGA ATAATTACAC GTTTGAAACC TAGTTTTGCA GCCTCTTGCA CGCGTTGTTC	5640
	TATCCGAGAT ACACGACGTA CCTCACCCGT TAAACCAACT TCTCCAATAT AGCAATCTAA	5700
	TCCGTCGACA GCTTTATCTT TAAAGCTAGA TGCAGTTGCT ACAATTACAC TTAAATCAAC	5760
50	TGCTGGCTCC GTTAACTTTA CACCGCCAGC TACTTTGATA TAAGCATCTT GTTGTGTAA	5820
	TAGATAATTT TCTTTCTTTT CCAAAACAGC CATCAACAAA CTTAATCGAT TATGATCAAT	5880

55

5 TATTAAAAGT GGTCTGGTTC CCTCCATGGT TGCAACAATT GTTGAACCTG GAACATTTGT 6000
 TGAACGTTCT TCTAAAAACA TTTCAGATGG ATTATTTACA CCTTTTAATC CACTTTGCTT 6060
 CATTTCGAAG ATTCCcATTt CATTCGTTGA ACCAAAACGG TTTTAAACAG CTCGCAAAAT 6120
 TCGATATGCG TGGTGTTCAT CGCCTTCAAA ATAAAGCACA GTATCaACCA TGTGTTCTAG 6180
 10 CAATCTTGGG cCCAGCAATT TGACCTTCTT TCGTTACATG ACCCACTATA AAAGTTGCaA 6240
 TGTTCATTG TTTAGCAATA TTCATTAAAC TTTGTGTACT TTCACGAAct TGTGAAACAG 6300
 AACCTGGCGC AGAGCTGATT TCAGGATGAT ATATTGTTG AATCGAATCC ACTACTAATA 6360
 15 AATCAGGTTG TTCTTCTTTT ACTGTTTGAT AAATAACTTC AAGATCTGTT TCAGCTAATA 6420
 CTTGCAATTC ACTTGAATCT TCATCTAATC GCTCTGCACG TAATTTAGTC TGACTAAGCG 6480
 ATTCTTCTCC AGTAATATAT AGTACTTTTT TCTTTTGAGA TAACGATGCA CAAATTTGTA 6540
 20 AAAGTAACGT TGACTTACCA ATACCTGGAT CCCCAACCAAT AAGTACTAAC GATCCGCTCA 6600
 CAATACCTCC ACCTAATACA CGGTTGAATT CTGCTGAATC TGTTAACACT CTCGGCGTTG 6660
 TTTCAATGTTT AATACTATTT AATTTTTGTA CTTTACCTGC TAATTCCTTG GTTTTAACTC 6720
 25 CATGTTTAGG ATTGGCTGCT TTTTCAACAA TTTCTCCAT TTGATTCCAA GCGCCACAAT 6780
 TAGGACATTT CCCCATCCAT TTAGGAGATT GATAACCACA AGCCATACAT TCAAAAATCA 6840
 CTTTTTCTT GGCCAArAATT GCACCTCCAC TTTCTT 6876
 30

(2) INFORMATION FOR SEQ ID NO: 187:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1193 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

40 CAACTCAAAC AGCAGAACAA CGTCGTGAGT TGATTAATGG TGTATTTACT GACATTAATC 60
 CCATACATTA AAAATATGAT GTACGTGTTA GCAGATAATA GACATATCTC ATTAATAGCT 120
 45 GACGTATTCA AGGCGTTCCA AAGCTTATAT AACGGACACT ACAATCAAGA TTTTGCAACA 180
 ATTGAGTCAA CATATGAATT GAGTCAAGAA GAGTTAGATA AGATTGTCAA ACTAGTAACT 240
 CAACAAACGA AGTTATCTAA AGTTATTGTA GATACAAAAA TTAATCCAGA TTTAATTGGT 300
 50 GGATTTAGAG TTAAAGTCGG CACAACTGTA TTAGATGGTA GTGTTAGAAA TGATCTTGTC 360
 CAATTACAAA GAAAATTTAG AAGAGTTAAT TAATTATAAA GAGGAGTGAC ATAGATGGCC 420

5 ATGTCGGTAA CTGATGTAGG TACTGTATTA CAAATTGGTG ATGGTATTGC ATTAATTCAC 540
 GGATTAAATG ACGTTATGGC TGGTGAGCTA GTAGAATTCC ATAACGGCGT ACTTGGTTTA 600
 10 GCCCAAAACC TTGAAGAGTC AAACGTGGGT GTGGTTATTT TAGGACCATA CACAGGTATT 660
 ACTGAAGGTG ACGAAGTTAA ACGTACTGGT CGTATCATGG AAGTACCACT AGGTGAAGAA 720
 CTAATCGGAA GAGTTGTAA TCCATTAGGA CAACCTATTG ATGGACAAGG ACCGATTAAAC 780
 15 ACAACTAAAA CACGTCCaGT AGAGAAAAAA GCTACTGGTG TAATGGATCg TAAATCAGTA 840
 GATGAGCCAT TACAAACAGG TATCaAGCA ATTGATGCTT TAGTACCAAT TGGTAGAGGT 900
 CAACGTGAGT TAATCATCGG TGACCGTCAA ACAGGTAAAA CAACAATTGC AATTGACACA 960
 ATTTTGAACC AAAAAGATCA AGGTACGATT TGTATCTATG TTGCTATTGG TCAAAAAGAT 1020
 TCAACAGTAA GAGCAAATGT TGAAAAGTTA AGACAAGCAG GCGCTTTAGA CTACACTATT 1080
 20 GTTGTAGCAG CATCAGCTTC TGAACCTTCT CCATTATTAT ATATTGCACC ATATTCAGGT 1140
 GTAACATGG GTGAAGAATT CATGTTTAAAC GGTAACATG TTTTAATCGT TTA 1193

(2) INFORMATION FOR SEQ ID NO: 188:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5549 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

35 TGCTAAGAAG TCAAAATAAA CTAACtATnA AACATCTAGT ACGATTATTA AAGTGACAGA 60
 TnATAAAATT GAATTATTnA GAGAAGGAGA TATAAAGTTT GAAGAAATAA AAGAAAGACT 120
 AGGTACAGGT ATTATTTATG AATAAGTTAA TACTTGGGAT TTATTTATAC CGAATTTTTT 180
 40 CACGAGCATA CTTTTATTTA CCGTTTTTAT TAATTTACTT TTTGATTCAA GGTTATTCCA 240
 TAATACAATT AGAAATATTA ATGGCGTCTT ATGGCATTGC AGCATTTTTA TTCTCTCTAT 300
 ACAAAGAGAA GTGTTTTTAA ATTTGTAAC TAAAAGATTC TAATAAATTA GTTGTTAGTG 360
 45 AAATATTCAA AATCATCGGT TTATTGTTGT TATTATATCA AAATCAATAT TTAATTTTAG 420
 TAGTGGCACA AATATTATTA GGGTTAAGTT ACTCAATGAT GGCGGGTGTT GATACCGCAA 480
 TAATTAAAAG AAATATAACA AATGAGAAAT ACGTACAAA TAAGTCAAAT AGCTATATGT 540
 50 TCCTATCATT ATTAATTTCA GGGATTATAG GTAGTTATCT TTATGAATA AATATTAAAT 600
 GGCCTATAAT AATGACTGGT ATATTTTCAA TTCTAACAAT TATAATTATT CGATGCACAT 660

	TACCAGAAGA GAAGTTTTGG ATATTGCATT ATTCTTTTTT AAGAGCGTTA ATATTAGGAT	780
	TTTTTATAGG ATTTATTCCA ATTAATATAT ATAATGATTT AAAACTGAAT AATTTACAAT	840
5	TTATTTTCAGT ATTAACCTTGT TACACAGTTA TGGGTTTTGT ATCTTCACGT TATTTAACTA	900
	AATACTTGAA TTATAAGTTT GTGTCAGAAA TTTGTTTAGT AATATTTTTA ATAATATATA	960
10	CATATCAAAG TTTCATAGCA GTTACTATTT CTATGATATT TTTAGGTATT TCTTCAGGGT	1020
	TAACTCGTCC ACAAACCTATA AATAAACTTT CTAGCAGTAG TAACTTAAGA GTGATGCTTA	1080
	ATTATGCAGA AACGTTATAT TTTATTTTTA ATATCGCATT TTTACTTATG GGTGGTTACT	1140
15	TATATACAAT AGGAACTATT CAATACTTAA TATTATTTAT TTCGTTATTA ATTTTTATAT	1200
	ATTTAATAAT AATATTTyAT TTTACAAGGA GAGAGCAACA TGAAAATAAA AACTGAATTT	1260
	AAAGGGAACA ATATACCATA TGAATACGCA GCAGGTGCAG ATGTGAGTGA TTCTATTAAC	1320
20	GGGAATCCAA TTAAGTCATT TCCATTTGAA GTAATTGAAT TACCGGAAGG gACTAAATAT	1380
	CTTGCTTGGT CTTTAATTGA CTATGATGCA ATTCCTGTAT GTGGCTTTGC TTGGATTCAAT	1440
	TGGAGTGTAG CTAATGTAAG TGTTAGTGGC AATTCAATTT CTATAAAAGC AGATTTATCA	1500
25	AGAACAAAGG GCGACTATGT ACAAGGTAAA AATAGCTTTA CTAGTGGGTT GTTGGCTGAA	1560
	GATTTTTTCAG AAATAGAAAA TCACTATGTA GGACCTACAC CACCTGATCA AGATCATCAA	1620
30	TATGAATTAA CAGTTTATGC GTTAGATCAT TCTTTAAATT TGAAGAATGG GTTCTACTTG	1680
	AATGAATTTT TAAAAGAAGT AATCAACAT AAAATTGATC AAACAAGTAT TAACCTTATA	1740
	GGAAGAAAAA TTTAATACTA AATATCTCAT CAATATAAAA TTGTTCAATT AAAAGTACAA	1800
35	AGAAACAAAG GTTTTAAATT ATATATTAGG TACGGCGTTC GCTATAATGC AAAGAAGTAA	1860
	TTAAATTTAA GAAATGTAAA CTTAGTTATT GTAATGTGAA TTTATTTGAA AAAATAGAAA	1920
	GTATTAACAA TTATAGCTTT TACATTAATT AAAATTATT TTTAAAAACA AGTAAACAAT	1980
40	TTACATACTT ATAATTTTTG AAAATTTTCA ATTTGTGTTA TATTGATTTT GTAAGATACT	2040
	TTAACTCACA AAGGAGAGAG AGTATATGAA ATTAAATCA TTTATAACTG TAACTTTGGC	2100
	ACTGGGCATG ATCGCAACGA CTGGCGCTAC TGTGGCAGGT AATGAGGTAT CTGCAGCAGA	2160
45	AAAGGACAAA CTACCGGCAA CTCAAAAAGC TAAAGAAATG CAAAATGTTT CATATACAAT	2220
	TGCAGTAGAT GGCATTATGG CTTTCAATCA ATCTTACTTA AATTTACCAA AAGATAGCCA	2280
50	ATTATCATAT TTAGATTTAG GAAATAAAGT TAAAGCTTTG TTATATGATG AACGCGGTGT	2340
	AACACCTGAG AAGATTCGAA ATGCAAAATC TGCCGTTTAC ACGATTACTT GGAAAGATGG	2400
55	TAGTAAAAAA GAAGTGGATC TTAAGAAAGA TAGCTACACA GCAAACCTGT TTGATTCAAA	2460

	CAACATGAAG	CATTTAATTT	TACAGTGATG	ATTATAAAAT	AATTGCCTTG	ATACAAAGAT	2580
	TACTCGTAAA	TGACATCTTT	GTATTAAGGC	TTTTTCTAAA	TTTAAAAGTG	ATGGGTTAGA	2640
5	GGTCATTGAG	CTTTAAAATA	TTCAAAATAC	AAAACATTAA	TGGCCAAAAA	TAAAAGCCGC	2700
	CTTTATCTGG	GCAGCTTCAA	TAATAAGAAA	GACATATTTT	ATTTTATACT	AAATAGTTAT	2760
	TGTGATGAAT	CTTTCGGCGG	TTTAATTACT	GCAGCAAAAA	TTGCTGTGAA	AATCGTGAAC	2820
10	AATACTGCCA	TGATAATTGG	ATTCACTACA	TTTAAGCTGT	CTCCACCTAC	TAGGCTATTA	2880
	AGTACAAAGT	TAACCATTTG	CATTAATAAT	AATGCCCAAA	AGAATGTTAC	GAGGTGTTTT	2940
15	ATGTCATTCT	ACCTCCACTT	TAATTATATA	TATTTTATTT	TAAGTGAAAG	TTAGAAATTT	3000
	GTATAGTAAC	ATCTCATATA	TTTTGACCAT	ATTATACAGT	TTAAATAAAT	GATTTTATCT	3060
	GAATGGCTAT	TCTAAATTAA	GCGCATTAAA	ACCAATTTCA	TACTGAAATT	TGACGATAAT	3120
20	AAAGCATTAA	AATTTTATTA	ACTAGTCAAT	ATTCCTACCT	CTGACTTGAG	TTTAAAAAGT	3180
	AATCTATGTT	AAATTAATAC	CTGGTATTAA	AAATTTTATT	AAGAAGGTGT	TCAACTATGA	3240
	ACGTGGGTAT	TAAAGGTTTT	GGTGCATATG	CGCCAGAAAA	GATTATTGAC	AATGCCTATT	3300
25	TTGAGCAATT	TTTAGATACA	TCTGATGAAT	GGATTTCTAA	GATGACTGGA	ATTAAAGAAA	3360
	GACATTGGGC	AGATGATGAT	CAAGATACTT	CAGATTTAGC	ATATGAAGCA	AGTTTAAAAG	3420
	CAATCGCTGA	CGCTGGTATT	CAGCCCGAAG	ATATAGATAT	GATAATTGTT	GCCACAGCAa	3480
30	CTGGaGATAT	GCCATTTCCA	ACTGTCGCAA	ATATGTTGCA	AGAACGTTTA	GGGACGGGCA	3540
	AAGTTGCCTC	TATGGATCAA	CTTGCAGCAT	GTTCTGGATT	TATGTATTCA	ATGATTACAG	3600
35	CTAAACAATA	TGTTCAATCT	GGAGATTATC	ATAACATTTT	AGTTGTCGGT	GCAGATAAAT	3660
	TATCTAAAAT	AACAGATTTA	ACTGACCGTT	CTACTGCAGT	TCTATTTGGA	GATGGTGCAG	3720
	GTGCGGTTAT	CATCGGTGAA	GTTTCAGATG	GCAGAGGTAT	TATAAGTTAT	GAAATGGGTT	3780
40	CTGATGGCAC	AGGTGGTAAA	CATTTATATT	TAGATAAAGA	TACTGGTAAA	CTGAAAATGA	3840
	ATGGTCGAGA	AGTATTTAAA	TTTGCTGTTA	GAATTATGGG	TGATGCATCA	ACACGTGTAG	3900
	TTGAAAAAGC	GAATTTAACA	TCAGATGATA	TAGATTTATT	TATTCCTCAT	CAAGCTAATA	3960
45	TTAGAATTAT	GGAATCAGCT	AGAGAACGCT	TAGGTATTTT	AAAAGACAAA	ATGAGTGTTT	4020
	CTGTAAATAA	ATATGGAAAT	ACTTCAGCTG	CGTCAATACC	TTTAAGTATC	GATCAAGAAT	4080
	TAAAAAATGG	TAAAAATCAA	GATGATGATA	CAATTGTTCT	TGTCGGATTG	GGTGGCGGCC	4140
50	TAACTTGGGG	CGCAATGACA	ATAAAATGGG	GAAAATAGGA	GGATAACGAA	TGAGTCAAAA	4200
	TAAAAGAGTA	GTTATTACAG	GTATGGGAGC	CCTTTCTCCA	ATCGGTAATG	ATGTCAAAAC	4260
55							

TGAACCTTAT AGCGTTCACT TAGCAGGAGA ACTTAAAAAC TTTAATATTG AAGATCATAT 4380
 CGACAAAAAA GAAGCGCGTC GTATGGATAG ATTTACTCAA TATGCAATTG TAGCAGCTAG 4440
 5 AGAGGCTGTT AAAGATGCGC AATTAGATAT CAATGAAAAT ACTGCAGATC GAATCGGTGT 4500
 ATGGATTGGT TCTGGTATCG GTGGTATGGA AACATTGAA ATTGCACATA AACAAATTAAT 4560
 10 GGATAAAGGC CCAAGACGTG TGAGTCCATT TTTCGTACCA ATGTTAATTC CTGATATGGC 4620
 AACTGGGCAA GTATCAATTG ACTTAGGTGC AAAAGGACCA AATGGTGCAA CAGTTACAGC 4680
 ATGTGCAACA GGTACAAATT CAATCGGAGA AGCATTTAAA ATTGTGCAAC GCGGTGATGC 4740
 15 AGATGCAATG ATTACTGGTG GTACAGAAGC ACCAATTACT CATATGGCAA TTGCTGGTTT 4800
 CAGTGCAAGT CGAGCGCTTT CTACAAATGA TGACATTGAA ACAGCATGTC GTCCATTCCA 4860
 AGAAGGTAGA GATGGTTTTG TTATGGGTGA AGGTGCTGGT ATTTTAGTAA TTGAATCTTT 4920
 20 AGAATCAGCA CAAGCTCGAG GTGCCAATAT TTATGCTGAG ATAGTTGGCT ATGGTACTAC 4980
 AGGTGATGCT TATCATATTA CAGCGCCAGC TCCAGAAGGT GAAGGTGGTT CTAGAGCAAT 5040
 GCAAGCAGCT ATGGATGATG CTGGTATTGA ACCTAAAGAT GTACAATACT TAAATGCCCA 5100
 25 TGGTACAAGT ACTCCTGTTG GTGACTTAAA TGAAGTTAAA GCTATTAAAA ATACATTGTTG 5160
 TGAAGCAGCT AAACACTTAA AAGTTAGCTC AACAAAATCA ATGACTGGTC ACTTACTTGG 5220
 TGCAACAGGT GGAATTGAAG CAATCTTCTC AGCGCTTTCA ATTAAAGACT CTAAAGTCGC 5280
 30 ACCGACAATT CATGCGGTAA CACCAGATCC AGAATGTGAT TTGGATATTG TTCCAAATGA 5340
 AGCGCAAGAC CTTGATATTA CTTATGCAAT GAGTAATAGC TTAGGATTCG GTGGACATAA 5400
 35 CGCAGTATTA GTATTCAAGA AATTTGAAGC ATAACTATAA nAATCTTCAG TAACGTTGTT 5460
 TTAGTTACTG AAGATTTTTT CaGTTTCTTT ATACTAAGAT GAGCGACaCa CAATCGTCAT 5520
 AATAAAATAT GAATATTTAT TAATAATAA 5549

40 (2) INFORMATION FOR SEQ ID NO: 189:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4832 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

AGATTATAGT AAGATTGATA GTTTGGCGAC TGaAGCgCGa GaAAAAATTAT CAGaAGTAAA 60
 mCCTTTAAAT ATTGCACAAG CTTCTAGAAT ATCAGGGGTA AATCCAGCAG ACATATCTAT 120

	TGGTTAGCAG AACAAATTAAG	AGAACATAAT ATTCAATTAA	CTGAGACTCA AAAACAACAG	240
	TTTCAAACAT ATTATCGTTT	ACTTGTGGAA TGGAAAGAAA	AGATGAATTT GACAAGTATT	300
5	ACAGATGAAC ACGATGTATA	TTTGAAACAT TTTTATGATT	CCATTGCACC TAGTTTTTAT	360
	TTTGATTTTA ATCAGCCTAT	AAGTATATGT GATGTAGGCG	CTGGAGCTGG TTTTCCAAGT	420
10	ATTCCGTTAA AAATAATGTT	TCCGCAGTTA AAAGTGACGA	TTGTTGATTC ATTAAATAAG	480
	CGTATTCAAT TTTTAAACCA	TTTAGCGTCA GAATTACAAT	TACAGGATGT CAGCTTTATA	540
	CACGATAGAG CAGAAACATT	TGGTAAGGGT GTCTACAGGG	AGTCTTATGA TGTGTTACT	600
15	GCAAGAGCAG TAGCTAGATT	ATCCGTGTTA AGTGAATTGT	GTTTACCGCT AGTTAAAAAA	660
	GGTGGACAGT TTGTTGCATT	AAAATCTTCA AAAGGTGAAG	AAGAATTAGA AGAAGCAAAA	720
	TTTGCAATTA GTGTGTTAGG	TGGTAATGTT ACAGAAACAC	ATACCTTTGA ATTGCCAGAA	780
20	GATGCTGGAG AGCGCCAGAT	GTTCAATTATT GATAAAAAAA	GACAGACGCC GAAAAAGTAT	840
	CCAAGAAAAAC CAGGGACGCC	TAATAAGACT CCTTTACTTG	AAAAATAATG CATAATCCTT	900
	TACAACTAAC ATAAAAGGAG	CGAATGGATA ATGAAAAAAC	CTTTTTCAAA ATTATTTGGT	960
25	TTGAAAAACA AAGATGACAT	CATTGGACAT ATTGAAGAAG	ATCGCAATAG TAATGTTGAA	1020
	TCCATTCAAA TTGAACGTAT	CGTTCCCAAC CGTTATCAAC	CAAGACAGGT GTTTGAACCA	1080
30	AATAAAATTA AAGAAGTTGC	TGAATCAATA CATGAACATG	GTTTACTACA ACCTATTGTT	1140
	GTAAGACCGA TTGAAGAAGA	TATGTTTGAA ATTATTGCTG	GAGAGCGCCG ATTTAGAGCA	1200
	ATACAATCAC TAAATTTACC	TCAAGCAGAC GTTATTATTC	GTGATATGGA TGATGAAGAG	1260
35	ACGGCTGTTG TTGCATTAAT	TGAGAATATT CAAAGAGAAA	ATTTGTCTGT TGTGAAGAA	1320
	GCGGAAGCCT ATAAGAAATT	ATTGGAAATT GGTGATACAA	CGCAAAGTGA ATTGGCAAAA	1380
	AGTTTAGGTA AAAGTCAAAG	CTTTATTGCA AATAAGTTGC	GTTTATTGAA GTTGGCGCCG	1440
40	AAAGTACTAC TTCGCTTAAG	AGAAGGTAAA ATTACTGAAC	GTCATGCGAG AgcGGtATTA	1500
	TCATTGTCTG ATAGCGAACA	AGAAGCGTTG ATTGAGCAAG	TCATTGCACA AAAGCTAAAT	1560
	GTGAACAGAc TGAAGATAGA	GTACGCCAAA AAACGGGGCC	CGAAAAAGTC AAAGCACAAA	1620
45	ACCTTCGCTT TGCACAAGAT	GTCACCTAAG CACGAGATGA	GGTAGGCAAA AGTATCCAAG	1680
	CGATTCAACA AACAGGATTA	CATGTTGAGC ATAAAGACAA	AGATCATGAA GATTATTATG	1740
	AAATAAAAAAT TCGAATATAT	AAACGTTaGT AGTAGGATGT	CGTATACATG ATGACTAACA	1800
50	CATAAAGAC AAAGCTAAGA	TCATAACAGC TTTGTCTTTT	TTTTTTGTTT TACGTGAAAC	1860
	ATAAAAAATT ATATTTATAT	GTTGATCAGG CTGGTACATA	AATCAATGTT CTATGCTCTA	1920
55				

	TTCTAGTCAA CCTTGCTGGG GTGGGACGAC GAAATAAATT TTGCGAAAAT ATCATTTCCTG	2040
	TCCCACTCCC TAATTTGAGC TGGATATACT TTCATTTGAA CCCTTTATTG CTAGTTTATG	2100
5	AAAGTATCAT GAAAGCTTTA TGAACATCGC TTGAGTTGCC TTTACAGTAG AAAATTTAAG	2160
	TTTTACACTT TGTGTGAATG ATACGTTTTG TATTGAATTA ATTATAGAAA GGTACGTTGA	2220
	AGATGTTTTT AATTGGAAGT GCAATTCCTC ATTTTGTTCAT TGGTGGTATC GCTGTTGCAT	2280
10	TAGCTTCAAT TATTGCTGAT AAGTAGAGTG GTAAGTTAGG AGGTATTATA GCTACTATGC	2340
	CGGCAGTCTT TCTTGCGGCT ATTATCGCAT TAGCTTTAGA TCATCGTGGT ACGCAATTAG	2400
	TGGAGATGTC GATGAATCTT AGTACTGGAG CAATTGTCGG TATTCTGTCT TGTATATTAA	2460
15	CTGTATTTTT GACATCTCTC TACATTAAGC ATAAAGGTTA TCGGAAAGGC GCAATATTCA	2520
	CAGTTGTTTG TTGGTTTGTG ATTTCCCTCG CAATATTCAG TATTAGACAT TTATAGTTTG	2580
20	GAAATGCGT GATAATTAGT TGTATTCAGT TATTAAGTAA TAAATTATTG GAGGCAGAAC	2640
	ATCATGAAAT TAACATTAAT GAAATTTTTT GTGGGGGGAT TTGCAGTATT ATTAAGTTAT	2700
	ATTGTATCTG TAACACTACC TTGGAAAGAA TTTGGCGGTA TATTTGCaAC GTTTCCGGCA	2760
25	GTATTTTTAG TGTCTATGTT TATTACAGGT ATGCAATATG GTGATAAAGT CGCTGTGCAT	2820
	GTAAGTCGTG GCGCAGTGTT TGGTATGACA GGGGTATTAG TTTGTATTTT AGTTACATGG	2880
	ATGATGTTAC ATATGACGCA CATGTGGTTG ATTAGCATTG TTGTTGGTTT CCTAAGCTGG	2940
30	TTCATCAGTG CAGTATGTAT TTTTGAAGCG GTAGAATTTA TAGCACAAA AAGATTAGAA	3000
	AAGCATAGTT GGAAAGCTGG AAAATCGAAT AGTAAATAGT GTGAACGTAA TCTCTTAACT	3060
	AGGACTAACT TTGCAAGCAT TGAATAGCAT GGAAAGTTG CATCATTAAT AAGTGAAATT	3120
35	CAAGTTGGCA TTGAGAAAAT TACAAGCGCG TAATCATACa GGTCTGTCTT AAGGGAGTCT	3180
	TCGAACCCCG ATGTTGTCGT ATGTCAAAAC ATTTAGTCAA TCATAAAGGT GACTTGATTT	3240
40	AACTTTATCT GATAGTCTGA TTGTAATGAT TGTACTAATT GACTGGAGGC GTATGTAATT	3300
	GAATCTGAGT AAACAAATTA AAAAGTATAG GGAACGAGAT GGTATTTCAC AAGAATATCT	3360
	TGCTGAAAAG TTATATGTAT CTAGGCAGAG TATTTCTAAT TGGGAAAATG ACAAAGCTT	3420
45	ACCAGACATA CATAACTTAT TAATGAYGTG TGAATTGTTT AATGTAACTT TAGATGATTT	3480
	AGTAAAAGGG ACCATTCCAT TTGTACCTGA TATTAAAGCG CAACGAAGTC TTAACCTTATG	3540
	GACATATGTG ATGCTTATTT TCATGACATT AGCTGCAATT TTAATGGGAC CTTTAGTTGT	3600
50	TTATTGGAAT TGGACTTGGG GTGTAACGGT GGCAATCATT TTGGGAATAG GTTTTTATGC	3660
	ATCTATGAAA ATAGAAGATT TAAAAAAGT GCATAAAATG GACAACTACG ATCGAATTGT	3720

55

5 GACAAATGCG CTTTCTATTA TATCAGTAAT TGGTATACTC AGCCTCATAA TTTTCCTTAG 3840
 TGTGTATTTG GCAAATAAGT TTTTATAAAT CATCGTGGTA TCGTCTCATA TTATTTATAT 3900
 TATCCAAAAT AGCATAAAAA AATACCAACA AGATTTAGAA CCTTGTTGGT AATCAAAGCG 3960
 aTTCATTTAT AATGAGTCGT TTTATGTTGT AAGATTAAAC AGTTTGTACG TTAAGTGCTT 4020
 10 GGTCTCCACG TTGACCTTCA GTGATTTTCA AAGTAACTTT TTGACCTTCT TCTAAAGTTT 4080
 TGTAGCCATC GCTAGCGATA CCTGAGAAAT GTACGAATAC GTCTCCGCCA TTTTCTTGTT 4140
 CGATGAAACC AAAACCTTTT TCTGCTTTAA ACCATTTWAC TGTACCGTTA TTCATATWGA 4200
 15 AwACCTCCGT gTGCTTTTGC ACTTAATATT TGTAACAAAT TCATAACTAA AAAAGAGGAT 4260
 ATTCTAAACA AATACACTAC AATTTAATTC ACGAGCTTTT ATTACGTAAG ACCAACTATA 4320
 CGCTCATATT GGCATAATGT ACAGTGT TTTT TGAAAATAA ATTAAAAAG ATTTTTTAAAA 4380
 20 ACCTTAGAAA CGTTGATTTA AAGGGGTTTA TAAAAATwaw AAAATTGTAG TCTTTTATGG 4440
 TGTTTGCTAG TTTTCAAAGT GACATATCGT TTAACATGA TGATTTTATA AGCAATCCAT 4500
 AAAAAACAAG CAGCGATAAA CGCTACTTGT TGATATTAAA ATCTGACTTG AAAGGTCATA 4560
 25 GCAATGTTCT ATACCGATGG AATGTGCTTA CTGCTTTT TCTTCACGAC GTTTTAAATA 4620
 ATAAGAGCCA CCTAATAAAC CAGCTGGAAT GCCTATCATT GGTGTTGTGA ATGAGCTTAA 4680
 TACAATAACA AGTATTGTGA AAGCAATGAC GTTATACCAA GTTACAGTCA AATTTTTTCAA 4740
 30 ATCCTCATAT GATTGTTTTA CTAATTCTCT AAATTTCATG ATTCAATCTC TCCTTTTTTA 4800
 TAAATCTTTA GATTGTCAAA TTAAGCTGGA CA 4832

(2) INFORMATION FOR SEQ ID NO: 190:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5727 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

45 CAAAGCTGTT CAAAAGGCTT ATAATTTAAA TTTAGATAAC ATACGTACAA TGGAACCTAA 60
 GTTGAGATAT CAAGCGATCA ATAAAGGTAA TATTAATTTA ATAGATGCAT ATTCAACTGA 120
 CGCTGAATTA AAACAATATG ATATGTTTGT GTTAAAAGAT GATAAGCACG TATTTCCACC 180
 50 ATATCAAGGA GCACCATTAT TTAAGAAAG CTTTTTAAAG AAACATCCAG AAATTAAGAA 240
 ACCGTTAAC AACTAGAAA ACAAATATC TGATGAAGAT ATGCAATGA TGAAGTATAA 300

	GTTAATCAAA TAACGACCAA CGCCACATAA GATGCGTAAC ACCAAATTAT ATCTTATGTG	420
	GCGTTGTTAT ATTTAAATCT ATAATTATGT TCAATTTAAA CATGCAATAA TGATTAAAAA	480
5	ATATGACATG TTAAACACAA TGTAAGCTAT TATGATGTGA AAATAGTAGC ATTGCATTTT	540
	AGAAACATAG AGCGATATAA TGAATATAAG TTTTTTGAAA TTTCAGTTAA TTCTAAGGAG	600
10	GTTGTTTTTA TTATGAAAGA ACAACTTAAT CAACTATCAG CATATCAGCC TGGTTTATCT	660
	CCAAGGgCAT TGAAAGAAAA GTATGGCATT GAAGGAGATT TATATAAACT TGCATCAAAT	720
	GAAAATTTGT ATGGACCATC GCCTAAAGTT AAAGAAGCGA TATCAGCACA CTTAGATGAG	780
15	TTATATTATT ATCCTGAAAC AGGATCACCG ACATTAAAG CGGCGATTAG TAAACATTTA	840
	AATGTAGATC AATCACGCAT TTTATTTGGT GCGGGATTAG ATGAAGTTAT ATTAATGATT	900
	TCTAGAGCTG TATTACGCC AGGGGATACT ATTGTTACAA GTGAAGCGAC ATTCGGTCAA	960
20	TATTATCACA ATGCGATTGT TGAATCAGCT AATGTGATAC AAGTACCTTT AAAAGATGGT	1020
	GGCTTCGATT TAGAAGGTAT TTTAAAAGAA GTTAATGAAG ATACGTCATT GGTATGGTTA	1080
	TGTAATCCAA ATAATCCTAC AGGTACATAT TTTAATCATG AGAGCTTAGA TTCGTTTTTA	1140
25	TCTCAAGTAC CTCCACATGT ACCAGTAATT ATAGATGAAG CTTATTTTGA ATTTGTGACA	1200
	GCAGAGGACT ACCCGGATAC ACTTGCTTTG CAACAAAAAT ATGACAATGC TTTCTTATTA	1260
	CGTACATTTT CAAAGGCGTA TGGATTAGCG GGTTTACGTG TAGGATATGT GGTAGCAAGT	1320
30	GAACATGCGA TTGAAAAATG GAACATCATT AGACCACCAT TTAATGTGAC ACGTATATCT	1380
	GAATACGCAG CAGTTGCAGC ACTTGAAGAT CAACAATATT TAAAAGAGGT AACACATAAA	1440
	AATAGTGTTG AACGCGAAAG ATTTTATCAA TTACCTCAA GTGAGTATTT CTTGCCAAGT	1500
35	CAAACGAATT TTATATTTGT AAAAACmAG CGGGTAAATG AACTTTATGA AGCACTTTTA	1560
	AATGTAGGGT GTATTACGCG ACCATTTCCA ACTGGTGTTA GAATTACAAT TGGTTTTTAA	1620
40	GAACAAAATG ATAAAATGTT AGAAGTTTIA TCAAACTTTA AATACGAATA GTAAGTGGGG	1680
	AGTGGGACAG AAATGATATT TTCGCAAAAT TTATTTGcC GTCCCACCCC AACTTGcATT	1740
	GTCTGTAGAA ATTGGGAATC CAATTTcCT TTGTTGGGGC CCCGCCGGCA AGGTTGACTA	1800
45	GAATTGAAAA AAGCTTGTTA CAAGCGCATT TTCGTTcAGT CAACTACTGC CAATATAACT	1860
	TTGTAGAGCA TTGAACATTG ATTTATGTCT CAAGCTCAAT GCAGTGTGAA TGATGAGGTG	1920
	AGAGTATTCA GTGTAAAAAG CAACAATAGA TGATATTGTT TTGTATCAAT TGCTTTTTTG	1980
50	CTATACTGAA TCAATACTGA TATTTTCAGG AGAAGATTAA AATGACCCGT AAATCAATCG	2040
	CGATTGATAT GGATGAAGTA TTGGCAGATA CATTAGGAGA AATCATTGAT GCTGTCAATT	2100

55

	TTCCTGAACA TGATGGATTA ATTACAGAAG TATTGAGAGA ACCAGGCTTC TTCAGACATC	2220
	TTAAAGTGAT GCCGTATGCA CAAGAAGTTG TGAAAAAATT AACTGAACAT TATGATGTAT	2280
5	ATATTGCTAC AGCAGCAATG GATGTACCAA CATCATTTAG TGATAAATAT GAATGGTTAC	2340
	TAGAGTTCTT TCCATTTTTTA GATCCTCAGC ATTTTGTGTTT TTGTGGTAGA AAAAACATCG	2400
10	TTAAAGCTGA TTATTTAATA GATGACAATC CTAGACAGCT TGAAATTTTT ACTGGTACAC	2460
	CGATTATGTT TACAGCAGTG CATAATATTA ATGATGATCG ATTTGAACGC GTAAATAGCT	2520
	GGAAAGATGT AGAACAGTAT TTTTGTAGATA ATATTGAGAA ATAAATATA TCACTTGAAA	2580
15	AATTTTCATGT AGAAAAGATG ATGGATAGGC TATAAAGTAA TTGTGACTGA GATGAACTTT	2640
	TATGTCTTAG ACACTACAAC ACTATATTGG CAGTAGTTGA CTGCGGGGCC CCAACATAGA	2700
	GAAATTGGAT TCCCAATTTT TACAGACAAT GCAAGTTGGG GTGGsCCCCA ACATAAGAA	2760
20	ATACTTTTTT TTTAGAAATT AGTATTTCTT ATGCATGAGT GTAACTCATG CATTCAATTT	2820
	TTTAAGTACA CATTAGCTGT GACTAATGAT AAAGAATCGC TACATAATCA ATCATTAGTC	2880
	GTTCTTTATC ATTTCCGTCC CGCTCTCAAT AAATGTTAGT CTATCTTATT ATTATAAATC	2940
25	GGATGAATGT GTTAATCTAT GGCAGATTAC ACGTCATCCG ATTTTTTATA GAATTTGAAA	3000
	AAGACGCATA AACCACATG ATTTAAATA CAACATCAAT CATTTTAGTG GCATGCGCCA	3060
	AAATTATATG TCTGTTTTTG AAACAGGGTA ATAGCTTAAA GCTAATAAAA ACGAATATAA	3120
30	GGTGCGTTGA ATCTTATGAT TACACTCCAA ACCTAATATA ATATCGGGTT AAGATCATTC	3180
	CGGATGCTTA CAAATCATTG ACAGTAAGTA ACTGAATGGC ATTTGGTATA ACCTCAATAT	3240
35	CAATAGGTGT TTCTAATGAA ATTTGCCAT CAATATCAAC TTTCAATGCT GGATCTGTTG	3300
	TAAGTGAAAT CTTTTTACCA GGTATATGCT CAATACCTTG AGTAATTTCA TTCCaATTCA	3360
	TGCTATCACG CTTTTTAAAA ATATCATTTA AAATACTGAA ACTTTGTTCA TTAAAAATGA	3420
40	AAGTGTTTCAG TTCACCATCT TGAGGAGACA AATCAGTCaA TGGTATACGA CTACCACCAA	3480
	TGAATGGACC ATTTGCTGTT AGTATCATGG TCGTTTCGCC AGAATATGTC TTATCATCTA	3540
	TTGATAATTG ATAATTAAAT TGTGTTGGAT TTAGCAGTGT TTTGACAGTT GATCCAATAT	3600
45	AACTCAATTT ACCAAATATA TCTTTTGAAC CATCTGTAC GTTTTCAGCG TTTTGAACAA	3660
	TGAGACCTAA GCCAACAAAG TTGAGTGCAT ATTGATTATT TATTTTAATT ACATCGTATG	3720
	TACCAACTTG TGCAGAAATC ATTTGTTTAC TAGCTTGTTT ATGATTAGGT GCTATATTTA	3780
50	GCGTTTTTGT AAAATCATTA AAAGTACCGC CTGGTAAAAT GCCAATAGGG AGTTGAAGGT	3840
	CATGTGTCAT AACACCGTTT ATAAGTTCGT TAACCGTGCC ATCACCGCCA AGAATAAATA	3900

55

	CACCTTCGTT TTCACTCAAT TGAATAGAAA GATGCTTACA AATTGAACTT AATGCTGTTG	4020
	TAACCTCCCC AATACCTTGA TTAATATTTT TTAATCCACT GTGTTTCATGG TAAAAGAGGA	4080
5	CACCATGTGT ATATTTATTT TCCATAGTTT AGCCTACTTT CTAAAAATTG GTTCATTAAA	4140
	TATATATACC CACTTTTAAT TGTTAATACC AAAAATATGT TTTTAAATAG AGAAAAATGGT	4200
	AATAAATGAA ATTGATTTCT ATAGAGTGGG ACGAGAAAAT ATAGTTATAG CTGTCTATAA	4260
10	TGAGCATATT AAGTTTTTAT TTATACTGAT ATCTTGAATT TAATTAATAG AAACCTATAA	4320
	AAAAACAGTA AGCCATTTAA ATGACTTACT GTTTTTTGAA TTAGGCCAAC AATATTAACG	4380
	TATACCTTTC ATCGCTTTGA TGATTAAAGG TGAGAATGCT AATACAATTG TTGTAACAAT	4440
15	AATTGCAACA ACACCTAGGA AAATAAAGTA ATTTGTTTGA CCTAGTGGTT CTATTAACCT	4500
	AACTAAAGTA CCATTGATTG CTTGTGCAGA AGCGTTAGTT AAGTACCAA TACTCATCAT	4560
20	TTGGGCATTA AATGCTTTAG GTGCTAAGTT AACAGCAGCA CTATTACCCG TTGGTGATAA	4620
	GCATAGCTCA CCGATAACAC AAATAATGTA CGATAAAATA ACCCAGTTAA CTGAAAAGTT	4680
	TGATGAACCT GATGCATAAC CTACAATACC AATTAGTATG TATGACGCAC CTGCTAAGAA	4740
25	CGTACCAATT GCAAATTTTA CTGGCAGGCT AGGTTGTTTA GTTCCAAGCT TTTGCCATAA	4800
	AAGTGAAATA ATTGGAGCTA GTAATAAAT AAATAATGGG TTAATTGATT GGAAGATCGC	4860
	TTACCAAAG TTTGTTTTCC AACCAAATAA GTTTAATTTT ATATCTGAAT GTTCAATTCC	4920
30	ATATATGTTT AATACATTAG ACCCTTGTTT TTGAATAGCC CAGAACACCA TTCCAAGAAT	4980
	AAATAATGGA ATAAATGCTT TAACACGAGA ACGTTCAGTA TCAGTGACAT CTTTACTTCT	5040
	AATAATTAAA GTGAAGTAAA TGANTGGTAA TGCAATACCT AATACTAAA CAGTATTACT	5100
35	AACTAAGTTA AATGATAATG AGTTAGTTAA TGCACCAATA ACGATAATTA ATACAATTGC	5160
	TAAACAACA CTTCCGATAA TAAGACCATA CTTTTTCTTT TCAGCTGGTG TCAATGGGTT	5220
40	AGTAGGTTTC ATACCAACGC TACCTAAGTT TTGCGGTTG AAAAGTACAT ACCATACTAA	5280
	ACCTAATGCC ATACCAACTG CTGCAATCAA GAATCCGCCG TGAAGTTTT TAACATTAAAC	5340
	AAAGTGTGCA AAAATAATAG GTGATAATAA TGCACCCATA TTAAGTACA TATAGAAAAT	5400
45	AACAAAACCT GCATCCATAC GTCTATCATT TTCAGGATAT AAACGGCCAA CGATATTTGA	5460
	AATGTTTGGC TTCATTAAAC CTGAACCAAT AATGATGAAG AACATTGATG TGAATAAGCC	5520
	GATTAATGCA AATGGTAAGC TTAAACAAAT ATGTCCGATA ATAATAAAGA CTGCACCTAA	5580
50	TAAAGTAGCG CCTCTAGTGC CTGTAATTCT GTCAGCAATC CATCCGCCTG GTATTGATGT	5640
	CATATAGATT AATGAACCAT AACTGACAT AATTGACATA GCTGTTGTTT TATCAATTCC	5700

55

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14078 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

	TGGACTATTA ACGGCGaAGA AGATTTAACG AAATACTTAC AAACCAATGT TGATGGTATT	60
	ATCACAGATG ACCCAGCATT AGCTGATCAG ATTAAAGAAG AAAAGAAAAGA CGAAACATAC	120
5	TTGATCGTTC CTATAAGAAT TTTGTTTGAA TAATATAAAC AAAGACCTCT AAAGTTATCA	180
	AGATGATACC TTCAGAGGTC TTTTAAATGT TGCCATCTAT GGGATAGGCA ATCGTTTCAT	240
10	TCGTTTATAT TCATATGACA AGTATTTGTA TGGCAATTTG GCGTCACAAA CACTTACATG	300
	ATTTATTGGT GAATTATTAA TTGTTTTGTG AATGCAAAGG GTTAGAAATT GAATTGTAAA	360
	TACTTTCTAA TCTTTGTTTC GCTTTAGTCA TTTGATCCAA ATTTTATAGTG CGTATAGCGG	420
25	ATTTTGCAAT ATAGTGCACA CTAAATATC GCGTTTTTGA AACGCATCTA AATTAGGTA	480
	CGATAATTTA TTTAAGTCAG TGTTTGCTAT TAATTCATGT AATTGATCTA CAAGCGCTTG	540
	ATGTTGATAC GTATGTGATG TAGTTTCAGA TTTGCTTGCT AATTAAATAC CAGTCGTATC	600
30	AAGGAGCGCC GCTTTAATAC CAGCAACTAA ATATGTTTTG ATTTTCATTT GTGTTGTCAT	660
	GCTTTGTTAC TCCTTTGATG TACATTAATC AAAAAAATTA TACACTATTG TATATTGCAA	720
	AGCTAATTAA CTATAACAAA AAGATAGTTA ATGCTTTGTT TATTCTAGTT AATATATAGT	780
35	TAATGTCTTT TAATATTTTG TTTCTTTAAT GTAGATTGGG CAATTACATT TTGGAGGAAT	840
	TAAAAAATTA TGAAAAGCA AATAATTTTCG CTAGGCGCAT TAGCAGTTGC ATCTAGCTTA	900
	TTTACATGGG ATAACAAAGC AGATGCGATA GTAACAAAGG ATTATAGTGG GAAATCACAA	960
40	GTTAATGCTG GGAGTAAAAA TGGGACATTA ATAGATAGCA GATATTTAAA TTCAGCTCTA	1020
	TATTATTTGG AAGACTATAT AATTATGCT ATAGGATTAA CTAATAAATA TGAATATGGA	1080
45	GATAATATTT ATAAAGAAGC TAAAGATAGG TTGTTGGAAA AGGTATTAAG GGAAGATCAA	1140
	TATCTTTTGG AGAGAAAGAA ATCTCAATAT GAAGATTATA AACAATGGTA TGCAAATTAT	1200
	AAAAAAGAAA ATCCTCGTAC AGATTTAAAA ATGGCTAATT TTCATAAATA TAATTAGAA	1260
50	GAACTTTCGA TGAAAGAATA CAATGAAC TAAGATGCAT TAAAGAGAGC ACTGGATGAT	1320
	TTTCACAGAG AAGTTAAAGA TATTAAGGAT AAGAATTCAG ACTTGAAAAC TTTTAATGCA	1380

EP 0 786 519 A2

	GTTGTATCAT ATTATGGTGA TAAGGATTAT GGGGAGCACG CGAAAGAGTT ACGAGCAAAA	1500
	CTGGACTTAA TCCTTGGAGA TACAGACAAT CCACATAAAA TTACAAATGA ACGTATTAAA	1560
5	AAAGAAATGA TTGATGACTT AAATTCAATT ATTGATGATT TCTTTATGGA AACTAAACAA	1620
	AATAGACCGA AATCTATAAC GAAATATAAT CCTACAACAC ATAACTATAA AACAAATAGT	1680
	GATAATAAAC CTAATTTTGA TAAATTAGTT GAAGAAACGA AAAAAGCAGT TAAAGAAGCA	1740
10	GATGATTCTT GGAAAAAGAA AACTGTCAAA AAATACGGAG AACTGAAAC AAAATCGCCA	1800
	GTAGTAAAAG AAGAGAAGAA AGTTGAAGAA CCTCAAGCAC CTAAAGTTGA TAACCAACAA	1860
	GAGGTTAAAA CTACGGCTGG TAAAGCTGAA GAAACAACAC AACCAGTTGC ACAACCATTA	1920
15	GTTAAAATTC CACAGGGCAC AATTACAGGT GAAATTGTAA AAGGTCCGGA ATATCCAACG	1980
	ATGGAAAATA AAACGGTACA AGGTGAAATC GTTCAAGGTC CCGATTTTCT AACAAATGGAA	2040
20	CAAAGCGGCC CATCATTAAAG CAATAATTAT ACAAACCCAC CGTTAACGAA CCCTATTTTA	2100
	GAAGGTCTTG AAGGTAGCTC ATCTAAACTT GAAATAAAAC CACAAGGTAC TGAaTCAACG	2160
	TTAAAAGGTA CTCAAGGAGA ATCAAGTGAT ATTGAAGTTA AACCTCAAGC AACTGAAACA	2220
25	ACAGAAGCTT CTCAATATGG TCCGAGACCG CAATTTAACA AAACACCTAA ATATGTTAAA	2280
	TATAGAGATG CTGGTACAGG TATCCGTGAA TACAACGATG GAACATTTGG ATATGAAGCG	2340
	AGACCAAGAT TCAATAAGCC ATCAGAAACA AATGCATATA ACGTAACAAC ACATGCAAAT	2400
30	GGTCAAGTAT CATACGGAGC TCGTCCGACA TACAAGAAGC CAAGCGAAAC GAATGCATAC	2460
	AATGTAACAA CACATGCAA CCGCCAAGTA TCATACGGAG CTCGTCCGAC ACAAACAAG	2520
	CCAAGCAAAA CAAACGCATA TAACGTAACA ACACATGGAA ACGGCCAAGT ATCATATGGC	2580
35	GCTCGCCCAA CACAAAACAA GCCAAGCAAA ACAAATGCAT ACAACGTAAC AACACATGCA	2640
	AACGGTCAAG TGTACATACG AGCTCGCCCG ACATACAAGA AGCCAAGTAA AACAAATGCA	2700
	TACAATGTAA CAACACATGC AGATGGTACT GCGACATATG GGCCTAGAGT AACAAAATAA	2760
40	GTTTGTAACCT CTATCCAAAG ACATACAGTC AATACAAAAC ATTACGTATC TTTACAACAG	2820
	TAATCATGCA TTCTATGATG CTTCTAACTG AATTAAAGCA TCGAACAATC GGAAGCATAT	2880
45	TTCTAAATTA TTTATTTCATT ATAGTCTTAA ACATAACATG ACCTAATATA TTACTAACCT	2940
	ATTAAAATAA ACCACGCACA TCTAAGTGAT ATACGACAAT CACAGCAATA ATAATTGCTT	3000
	TAGAAAGTCG TGCCGAAGTG GAACTTACAA GTCTAGTTCG AACACACACT GATGTGAGTG	3060
50	GTTTTCTTTA TTTTAAACAT GAACAATCAG ATAAGTTACT AGCATTAGCA AATATTATTA	3120
	AATCAAAGGG CTTCGATTCA TAAAATTTAA AACAAATGATT AAAATTAGAC GTGTAAATGT	3180

55

	TATTTACACAC AGCTTCATTA ATAAAACGAA ATTGCTTCAA CCCGCTTCAA CTTCAACTGG	3300
	CTTCAACTTC AGCCTACTTC ATTCAATAAC AAAACGAATC CGCTTCATCC AAAATCAACC	3360
5	ATTCTAACGC ACATATTCAA ATATAGCAGC TGCACCCATG CCGACACCAA TACACATCGT	3420
	AACCATGCCG TAACGGCTAT CGGGACGTCT ACCCATTTCA TTAAGTAAAC GCGCGGTTAA	3480
	CATTGCGCCT GTAGCACCTA ATGGATGACC TAAAGCAATA GCGCCACCAT TCACATTCGT	3540
10	ACGTGATATA TCTAGACCTA CTTCTTTAAT AGATGCAATC GTTTGAGAAG CAAATGCTTC	3600
	GTTCAATTCTG ATCAAATCAA TGTCTTCAAC AGATAGATTG CTGAGTGACA ATACTTCAGG	3660
	AATCGCATAT GCAGGCCCAA TACCCATAAT TTTGCGGTCA ACGCCTACTG CCTTAAAACC	3720
15	AACGAATCGT GCAATAGGTG TCACGCCGAG TTCTTTCACT TTATCTCCAG ACATTAAAAC	3780
	TACAAATCCT GCACCATCAG AAAGTGGGGC AGATGTTCTT GCAGTCATAG TGCCGTCAGC	3840
20	TTTAAATACT GTACGTAATT TGGCTAATGC CTCCATCGTG GTGTCAGGGC GTATAAATTC	3900
	ATCTTGGTCA AAGATATTTG TGTGTACTTT TGGTCCTGCG TTTGTATATT CAACTGAGTT	3960
	TACTTGTATT GGAATAATTT CATCTTTGAA CCGACCATCA CGTTGTGCGT CATAGGCACG	4020
25	TTGATGACTT CTGACAGCAT AAGCATCTTG ATCTTCGCGT GATACGTCAA ATTGGGATGC	4080
	TACATTTTCA GCAGTTAAAC CCATAGGATA TGACGCACCT ATATCATCAT ATTGTAAGGT	4140
	TGGATTGTTT GTGGGCTCGT TGCCACCCAT TGGTACGGCA CTCATCAATT CAACGCCACC	4200
30	AGCTACAAGT ATATCTCCTT GACCAGCCAT AATTTGATTG GCTGCAATCG CGATGGTTTG	4260
	TAATCCTGAT GAGCAGTAGC GATTCACTGT TTGACCCGGT ACCGTGTCAG ATAATCCCGC	4320
	ACGCAATGCA ATCGTTTCGT CAATGTTTTG GCCTTGTAAT CTTTCTGGAA AAGCCGTACC	4380
35	AACAATGACA TCTTCAATCA TATTCTTATT GAATTTTCCG TCAATACGTT TCAATACGCC	4440
	TTGTAATACT TTGGCTGCGA CATCATCAGG TCTTTCGTGG AATAATGCGC CTTGCTTTGC	4500
	TTTCGCTGCG GCTGAACGCC CATAAGCTAC AATGTATGCT TCTTGCATGG TTATCATCCT	4560
40	CTCTTAATGA CTATCTTTTA ATTACGTAAT GGCTTACCAG TTTTAAACAT ATGTGCAATT	4620
	CTTTCATATG ATTTTPTAGA TTTTAGTAAG TCAATAAAGC CAATTTTCTC CAACGATTGA	4680
45	ATGTAACGTT GATTGATAAA TGTATTTCTT GGTAAATCAC CACCCGCTAA AATTGTGGCG	4740
	ATATTTAAGG CAATATGATA ATCATGGTCG CTAATAAAAT GACCCCGTCT TTGCGCATCT	4800
	AATGTGCTT GGATCAATGC TTTGAAGTCT TCACCTAAAG CGATATATTG ATGTCTAGGA	4860
50	TTCGGAATAT AGTTTGTTC TGCTTCATAT TTCGCACGTT TGAGCGCAAC TTCGACACGT	4920
	TGTGCTGTAT TGAAAATAAT CGTATCTGTA TCACGTAAAT AACCATAACG ACGTGCCTCA	4980

55

EP 0 786 519 A2

	TGTTTGTTCAT CAAACTTATG CGATGTGCGT AATATGCGAT CAGCCATTTC TGCAAGGCCA	5100
	CCGCCACTCG GTAATAAGCC AACACCTGCT TCAACAAGAC CGATATATGT TTCACTTGCA	5160
5	GCGACAACAA TAGGTGAGTA AAGTACAAGC TCACAGCCAC CGCCTAAGGC ACGACCTTGA	5220
	ACAGCTGTGA CTAAGTGGTTT CAAACTATAC TTCAAACGAT TAAAGCTATA ATGTAATTTA	5280
	TCAATTGATT GTGCAACGAC ATCATCTACA AGACCGTCTT CATGCGCCTT TTTCAATTAAG	5340
10	AAAAGGTTAG CACCCACACT GAAATTGTTA CCATCTGCGT AAATAACCAT ACTTGTGTAA	5400
	TGGTCATTTT CCAGTAAATC AATCGCATCA ACTAACGCAT CGTTGAATTC ATCGGTAATG	5460
	ACATTATTTT TACTTTGTAA TTTTCAAGTA AGTTGATCAT CATGAGTTAC GGAAAGTTTG	5520
15	GCAATCACCTT TATCCCAAAG TTCATCTTTT ACGAAGTGAG AAATAGGTGT TGCATATTCA	5580
	ATGGTCTCAT CTGTTTATA AAAGCCACCA TCTAAATCAC TAATCCATTG TGGTAAGTCT	5640
20	CCAAGTTCGT CTTCCATACG TGTTTTAACA CGTTCGTATC CCATTGCATC CCATAATTGG	5700
	AATGGACCAA GTTTCCAGTT GAACCCCCAG ACAAGCGCAC GGTCTATGTC TCGGAAATCA	5760
	TCGGTAGCTT TAGGTACATT GATAGCAGAG TAATAGAAAT TATTACGTAA TGTCTCCAT	5820
25	AAAAATAGTC CCGCTTCGTC TTGCGCATTG AATATGGTAT CAAGGTTATG CACTAAGTCT	5880
	TTATTAAAT CATTTAAAT TGGTAATTGT GGTGCGATA CAGGTACATA ATCTTGTTTT	5940
	TCAACATCGT AAACAAGTCG AGCTTTAGTT TCTTTATCCT TTTTGTAATA TCCTTGTTTC	6000
30	GTTTTACGTC CGAGTGCGCC ATTGTCAAAC AACGTATTTA CAATTTTGAC ATCATGAAAA	6060
	TAAGGTGTTT CTTCAAGTAC TTGTTGCATG CCTTTAATTA CAGACACTGC AATATCTAAA	6120
	CCGACTAGGT CAGATAGCGC ATATGTACCT GTTTTAGGAC GACCAATCGC TTGCCAGTT	6180
35	AAAGCATCCA CATCTACAAT GCTTATCTTG TGTGCTCGG CGCGATACAT AATATCATTC	6240
	ATTGTTTGCG TGCCGACTCT ATTTGCGACA AAGCCAGGCA CATCATTGAC GACAATGACA	6300
	CCTTTACCTA ACACATTTTG CGCGAAATTT TTTACATCTA ATATAATAGA TTCCTTCGTG	6360
40	TGTGACGTAG GTATTAACTC CACTAATTTT ATAATACGTG GTGGGTAAA GAAATGTAGA	6420
	CCAAAGAATC GTTCTTGATC CTTCTCGTTA AATGCTTGAG CAATCGCATT AATTGGAATA	6480
	CCTGATGTAT TTGTAGCGAA TAAAGCATCT TCTTTAGCAT GTTGTAGAAC TTGTTGCCAA	6540
45	ACAGCATGCT TAATTTCAAT ATCTTCTTTG ACTGCTTCGA TATATAAATC AGCATCATCA	6600
	TTTACCAAGT CATCATCAAA ATTACCATAT GTTAAATGAC TCGCTAGATT TAAGTCGAAT	6660
50	AGTAGCGGCC GTTCTTATC TGTAATTTA TCGTAAGATT TTTTCGCAAT GAGATTTGGA	6720
	TCGTTTTTGT CCACTACAAT ATCTAATAGT TTTACTTTAA GTCCAGCATT CACAAAAAGT	6780

55

EP 0 786 519 A2

	GTGATTCTC CAATTTAGTT GAGGATAAGA TAACCATTAA GATAATTGGA ATAACGTTGC	6900
	TATTTTATAA AATTAATTAA GATCTTTTGA CAGTCATCTT AGCCTCTTAT TTAAGGAAAA	6960
5	AGCTTTATGC TTAAAATAAG TCTTTTTTGT TGAAATTAAT GCATCTCATA TAATTATTGT	7020
	CTATTTATAC GAAAGCAGAA TCTCCAGTCA AAGCGCGTCC AATTACTAAG GCATTAATTT	7080
	CATGTGTACC TTCGTACGTG TAAATCGCTT CTGCATCAGA GAAGAAACGT GCAATATCAT	7140
10	AATCGTCAGC TAGTATGCCA TTACCACCTG TAATACCGCG GCCCATAGCT ACTGTCTCAC	7200
	GCAAACGTAA GGCATTCATC ATCTTCGCCG TTGAAGTTGC AACCTCGTCA TATTCACCAT	7260
	GTGCTTGCAT ATTAGCTAAT TGAGCACATG TTGCCATTGC TTGAGCTAAA TTACCTTGCA	7320
15	TCATTGCTAG CTTTTCTTGT ATTAAGTGAT ATTTACTAAT TGGTTTGCCG AATTGCTTAC	7380
	GCTCAGTGAC ATAATCTAAT GTGGCACGTA AAGCGCCAGC CATACCACCT GTAGCCATAT	7440
	AAGCAACGCC TGCTCTCGTT GAATAAAGAA TTTTGGCAAT ATCTTTAAAG CTTGTTATGT	7500
20	TTTGTAAGCG ATCCGCTTCA TCTACTTTGA CATTAGTTAA TTTAATTAGG GCGTTAGGAA	7560
	CAATGCGAAG TGCGATTTTA TTATCAATGA CTTCAATATC GACGCCATCT TGTTCGGTC	7620
25	TGACTACAAA GCAATGGGGT TTGCCAGTTT CTTTATTTAC TGCGAATACT GGAATGACAT	7680
	CAGATACATG TGCACCAECA ATCCATTTCT TTTACCAATT GATAACCCAA GTATCGCCTT	7740
	GGCGTTCAGC GACTGTTTCA AGACCTCCCG CAACGTCCGA ACCGTGTTCT GGTTCAGTTA	7800
30	AAGCAAAGCA TGTACGCAGT TCATGTGACT GTAATTTAGG TACATATTTT GCAATTTGTT	7860
	CTTTGCTACC TCCGAAATAG AAAGTGTAT GCCCTAAACC TTGGTGAACA CCGAGTAGGG	7920
	TAGCTAAGGA AATATCAAAT CGCGCGAGTA GGTAAGACAT GAAAACTGA AATAGTTGAC	7980
35	TAGGCATTTT GGCGTTTGGA CGATCCTTGT AAAGTAATGG ATTGTTAAAA TAATTAAATT	8040
	CTCCAGATC TTTAAAATAG TCCTCGGGTA CAGTAGCGTC TATCCAATGT TGATTAAATAT	8100
	TTTACGGTA CTTACTTTCT AGCAATGAAT CTACTTGTG TAAAAATTCG ACTTCACCGT	8160
40	CTGTTAAACC TTTAGCAATA CTAAGTACAT CTTCAGGAAA TAATGTTTTT AAGACCGTTT	8220
	CTTTTTCAAA TGTATATAA ATTCCTCCTA AAAATAATAT GAATACTAAT GTGAAATGCA	8280
45	TTTAATTCAA AAACAACACG CTTTATTTGT AAACGCTTAC ACTAAATGTC AAAAAATTTT	8340
	ATCACCTTTA AAGTGTTCG GAGACTTTGT CATTATCAT TTGTCGAATC GCAAGTTTAT	8400
	CTGGTTTCTG CGTACTGTTT AACGGCATAT GTGTCACTGG TACATACATT CTTGGGACTT	8460
50	TATAACCTGC TAAACGACTT CGCATATGTT GATTTAAAAT TTCAGCGTAA TGAGGTTTAT	8520
	CTTCGCGAAG TATAATGGCT GCAGCAATTG ATTCACCATA TTTTGGATGA TCATAGCCAA	8580

55

	AGACATTTTC GCCACCAGTT ATGATTAATT CTTTTTTGCG GTCAATAATA AATATATCGC	8700
	CATCGTTGTC CATCTTCGCT AAGTCACCAG TTAATAAATA TCGACCATGA AATGCTTTGG	8760
5	CAGTCTCTGC TGGTTTATTC CAATATCCTG GCGTGACATT TTTAGCCTTA ATTGCAAGTT	8820
	CGCCAATCTC ACCAGTAGGT ACTTCCTCAC CGTTATCATC AAGGATACGT GCATCAACGA	8880
	ACATGACTGC TTTACCAATA CTCATTGGCT TACGTTTTGA ATTTTCCGGT GTATTAAACA	8940
10	GTACAAGAGG TGCTTCAGTT AAACCATAGC CGTTAATAAT GTTTATGCCA TATTGTTTAA	9000
	AAGCTGCTTG GATACTTGGT AATGGTTGTG AACCACCTTG GATGATATAA TCCATAGCTC	9060
	TAAAATTTTC AGGATTAAAA TTACTAGCAC GTAGCGTACT ATAATACATT GTCGGAATCA	9120
15	TGATAATAAA TGTAGGGTGA TATTGTGCAA TCATGTCATT CAATTCTTCG CCGTTAAAGT	9180
	AACGTTGAAG AATAAGTGTG CCACCTGACA TTAATACTGG TAATACAGTA TCGTTAAACC	9240
20	CTAAACATG GAACATTGGT GTTGATACAA TCGTAATATA GTTTGAATTG AACTTATACG	9300
	TCAGCTCTAA GTTTGCACCG TTATGAACAA ATGATTCATA TGAGAACATC ACACCTTTAG	9360
	GTGATCCGGT TGTACCACCT GTATAAATTA ATGCTGCAAG ATCTTGTGGT TCAACAGGTG	9420
25	TTGCTTGAAA AGGTTGGTGA TAATCTGGAT TTACGATTC ATCATATTGC GCCACATCAA	9480
	TATCCATATG CAATAAGTTT TGGTCAATAT CGGTGAGTGA ACTTAAATGT TTTTCAGCAT	9540
	AGAAGAGCAG TTTTAATTGT GCATCTTCCA CAATGGCTGC AATTTCTTTT GGGTTAAGCC	9600
30	GCCAATTCAA TGGTAAAAAA ACCGCACCTG TTTTAAACA AGCAAACAAT AAATCTAATA	9660
	TTGCAATATC ATTTGGCGCA AAAATACCGA TAACATCGCC TTTTTTAACA CCTTGAGATG	9720
	TTAAATAATG TGCCATATTA TCAGCGCGTG CATTGAGTTG TTGGTATGTC CAAGATGTTT	9780
35	GTTTTGCGTG ATCAATAACG GCAGGCTTGT CATCATCGAA GTCTGAACGC GTTTTTATCC	9840
	AATCGAAATT CATTAGTATA CCCCCTTTAG CTTCACTTTC ATACTTTATG AATTGATTGT	9900
	TTAAGTTGTC CCCATTTTTC TTTGTAAATG CTGGTATCAA TTAATTTTAA ATGATCAGCA	9960
40	ATAATTGGTT TAAAAGCCAT TTGATTCAAA ATATCTTTAT GCAAATCAAG ACCTGGTGCA	10020
	ATTTCAATTA GTTTCAAGCC TTGATTGGTG AGTTCGAATA CTGCACGATC AGTAACAAAA	10080
45	TAGATTTCTT GCTCGAGTGA TTGTGAATAT TGTGCATTAA AGTCGATATG GCTCACATCT	10140
	GATACAAATT TCTGGTTTTG TCCTTCAGTT TCAATGTTTA ATCGTTGATT ATGGCATGAG	10200
	ACATGACTGC CAGCTACAAA AGTACCTGAA AAGATAATTT TATTTACAGA TTGCGTAATG	10260
50	TCTATAAAGC CACCACATCC ATTTAGTCGG TCATTGAAGT AAGACACGTT GACATTGCCG	10320
	TATTGATCAA CCTCAGCAA GCTAAGATAG GCAACTGATA CACCATTGTT ATAAATAAAA	10380

55

EP 0 786 519 A2

	CGACTCCCAA CGAATCCACC GAAAAATGCCA ACATCTAAAA TCGGTTGCAC ATCATGTTCA	10500
	ACACATTCTT CATGCAATAA ATTAGAGAGT TCATTATTGA TGCCATAACC GATGCTAATT	10560
5	GTATCGCCAT AAGTTAAAAA CTGAGCAGCA CGTCGGAGAA TCAATTTGCG ACTATTAAAA	10620
	GGTAATGCGG GTTCAGGTAT TCCATCAATT CGTTCTTCTC CAGACAAGGC TGGTAAATAA	10680
	TGACTTTGAA TTACTTGGCG GTGATTCTTT TCATCTTCTG TGACGTATAC ATAATCGACA	10740
10	AGATTTCTTG GGATAACAAC TTCATTTCGGT TTTAGTTGAT AGTCGTCAAC TAAAGCTTTA	10800
	ACTTGTACAA TAACTTTCCC ATGATTGGCT TTCGCGTTTA ATGCGACATG ATAACACTCG	10860
15	CTCAAGTACG CTTCTTGAGT TAAATAAATG TTACCTTGTT GATCTGCGTA TGTTCTCTCTC	10920
	AGTAGTGCCA CATCAACGCT AGGGAATGTG TAATGTAAGT ATGTTTCATC GTTGATGGTT	10980
	ACTAATGAAA CTAAATCATC CGTTGTTCGT GTATTACTT TACCGCCACC GTATCTAGGA	11040
20	TCAACAGCTG TGTTTAATCC GATTTTAGTA ATAACTCCAG GTAATAATTG ATTACTCTGA	11100
	CGATAATGAG TTGCAATGAT ACCTTGTTGGT AAAAAATAAG CTTCAATGTC ATTATTTTTC	11160
	ATTGCTTG TG CCGTTTTGGA AGAAGCCGTT AAAATACTCA TAATGACACG TTTAATCATG	11220
25	CGACGTTCTA TAAATCATC TAAATCCGGT GCGGCACCTA AACTATGAAT ATCATTTCGT	11280
	AATATAAACG TTAAATCATT GGGCGTATGA TATGTGTCAT GTTGCGCTAA CACAGCACGT	11340
	AGAACTTCGG CGGGTAAGTT GGCTACAGCT AATGCTGGTA AACCAATCAC ATCACCATCT	11400
30	TTAATGATAT GTTGTAAGTC GTGCCATGTG ATTTGTTTCA AGCAAGTCAC CTCCATCACA	11460
	TTTGATAAAA TATAGCGTTT TTACACTTTG TGTAACCCT TaCAAGAAAT ATAACATAAC	11520
	GACGTTTAAA ATCAATTAGA AATATCTTTT TATTCTGATA ATAGACACAG TATAGACACA	11580
35	TTTTGATGGT CGATAACAAT TGTAATATCA AGGGTTTGTA ATGAATTGAA TATCATTAAA	11640
	ATACATATAT AAAAAATTG TTCGGAATAT AAAAAGTTAA ATAGGTTTTG ATTTTTAAAT	11700
	ATGAAATACA AAGTGCCCAA TCGAACAAAG TATTTATATT AAAATATGGA AAATCCATCA	11760
40	ATATTAAATT AAAATAGTTT TATTATGAAA AGTGAAAGTA GGTAAGTCTA TGGAAGGTCT	11820
	TAATCATCGA AGAAATACAG AAAAAGAAGA GACAACACAA ACGCAATCaG TTGCACCTAA	11880
45	TACAGGTGAA GAGGGGATGT CATCAGCAAG TACACAATCA ACTAAGACGT CCGACATACA	11940
	TAATGAATCT ATCGATAAAC AAATGGAAGC TAAAGCGCAT GAAACAGCGC AAAATACAGA	12000
	TTTAAAAAAC GAAGCAAGAA GTTTATTTGA TAATGCAACC AAATCAATCG GTAGACTAGC	12060
50	GGGCAATGAT GAAAGCTTAA ATCTTAATTT AAAAGATATG CTTTCTGAAG TATTTAAGCC	12120
	GCATACTAAA AACGAAGCAG ATGAAATATT TATAGCGGGT ACTGCTAAAA CTACGCCAGC	12180

55

	TTTCACAGTA ACATTTATTG GATTATGGGT CATGGCAGCA ATTTTAAATA AACTAACGC	12300
	GATTCCGGGT CTCATTTTTA TAGGGGCTTT AACAGTACCA TTATCGGGT TGTTCCTTCTT	12360
5	TTATGAATCA AATGCGTTTA AAAATATTAG CATTTTGTAA GTTATTATCA TGTTCCTTAT	12420
	TGGCGGCGTA TTTTCATTAC TAAGTACGAT GGTATTATAT AGATTGTGTCG TTTTLAGTGA	12480
	TCAATTCGAA AGGTTTGGTT CTTTAACATT TTTCGATGCA TTTTLAGTAG GATTAGTTGA	12540
10	AGAAACTGGA AAAGCACTCA TTATTGTTTA TTTCGTCAAT AAATTGAAAA CAAATAAGAT	12600
	TTTGAATGGA TTATTAATCG GTGCTGCTAT TGGTGCAGG TTCGCAGTTT TTGAATCAGC	12660
	AGGTTATATT TTGAATTCG CTTTAGGAGA AAATGTCCCA TTATTAGATA TTGTCTTCAC	12720
15	ACGTGCGTGG ACTGCGATTG GTGGTCATTT AGTTTGGTCA KCGATTGTTG GTGCTGCAAT	12780
	AGTTATTGCG AAAGAACAGC ATGGCTTTGA ATTCAAAGAT ATTTTGTATA AACGCTTTTT	12840
20	AATATTCTTT TTATCAGCCG TTGTTTTACA TGGCATTGCG GATACATCTT TAACTGTACT	12900
	TGGCAGTGAT ACGTTGAAAA TATTTATTTT AATCGTTATT GTGTGGATAC TTGTATTCaT	12960
	TTTAATGGGG GCAGGTTTAA AACAAGTGAA TTTACTGCAG AAAGAATTTA AAGAACAACA	13020
25	GAAAAAAGTA GACGAATAAT AATTAAAGCT TATGTTGCTC ATATGTTTGT GACATAAGCT	13080
	ATTTTATAA TTTGTCTTTA AAAGAGTGGA ATAGGAATAC TTTTGGAGT TAAAAAGTG	13140
	TTTCACGTTA AACAAATAGT GACAATTAGA TTTATATAAA ATGAACATGA TTCACTGAAA	13200
30	GTATGTAATA ATCATTTTTAT TGAAATTCAT CAAACAGAAA TTAATACAAT CATATAAGCA	13260
	AATTAAACCA CGCCATAATC ATATTGGATG ACTTCGGCGT GGTTTTTATA GTTGAAGCAG	13320
	GGCTGAGACA TAAATCAATG TCCCACACTC CCTTATCGTT CAATCGTTGT TCGATAATCG	13380
35	ATTAAATAGA TACCTTCAGG TGTTACTTTA TAATTTTAA CCTTAGAGTT AGCAGCGACT	13440
	ATTGATCGT TGTAAGCAAT ATAAGTGTG GGTACATCTC GACTTGATAA TTTAATAATA	13500
	TCAATTAGAAA TATTGTGACG TTCCTTAACA TCTACAGTAT GATTCAATTG ATTAATTAAA	13560
40	TCATCGACGT TGCTATTATT GTAGTCTCCT TTATTAATAG CACCATCTTT TTTATATGCT	13620
	TGATTAAAGA AATAACCTGT ATCTCCACGA GGAATTGTTT CGAAACTATA CATCGTTGCA	13680
45	TCCCATGCAG AACGGTCTTT TAAGTAACCT TCTATGTCAT CAACACTTTT AATGTCGATT	13740
	TCAATATTTG CTTTTTTAGC ATCTGATTGT AATACCTGCG CAATTTTCGA TAGCTCTGGA	13800
	CGACCGTCAT ACGTAATTAA CTTAATTTTT AAAGGTGTT CTTTTGTATA ACCATCTTTA	13860
50	GCTAATAACA TTTTGTCTG TTCGATATTT TGTGTTGTTA ACTTAGGTTT TTTAATATAT	13920
	GGAATTTTAT CATTAAATGG ACTCGTTGCA GGTTCGCAT AACCTTGATA AATATGATCT	13980

55

TTATTAGTAT GATTATACAT AAGTaAGAAG TTCTAAAn

14078

(2) INFORMATION FOR SEQ ID NO: 192:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 486 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

TGAAAATAAA AGTGTCTCTA ATGCGTGACT AAAATTAGTA ATAATTAAGT TCTCATGATA	60
ATAGGTATTT TTGAAAAATG GAGGAGTCTA TAAATGGGTA AAAAAATGGG TCTAGGTTTA	120
TCTATTGCAT TGGTTGTTAT TGGTATTGCC GTTGTATGTT TAATGATTTT TTCTAGTCAA	180
AAAACGACTT ATTTTGGTTA TATGAATAGT AATACAAATG CAGAAAAAGT TGTCAGTGAA	240
AAAGATGGAT TAGTCAAACA TAATATCAAA GTAGAACCAT CTAATGATTT CAAGCCGAAA	300
AAAGGAGACT TTGTAAAATT AGTTTCTAAA GATGATGGGA AGACATTTTA TAAACAAGAG	360
ATTGTTAAAC ATGATGACGT CCCACACGGT TTAATGATGA AAATTCACGA CATGCATATG	420
AATTAATAAAA AAAGCATCTA TAACGTAATT TTGAAGAAGT AGAGTTATCT TCTTATGCGT	480
TTTAGA	486

(2) INFORMATION FOR SEQ ID NO: 193:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1626 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:

GAGGTCTATA TACAATTATG GTTGTTCAG TTAACGAAC TGATGGCTTT ATTACTAAGT	60
TTAATAGATT AATTGAAAGA CGATTATTAC GTCATTTTCA TAAAAAAGGT TATATCACAT	120
GGGAGGAAAA TTGATTGTCT GACATTTTAA AATGTATCGG TTGTGGTGCG CCACTTCAAT	180
CTGAAGATAA AAATAAACCT GGTTTTGTAC CAGAGCATAA TATGTTTCGT GATGACGTGA	240
TTTGACGACG TTGTTTCCGC TTGAAAAATT ATAACGAATT CAAGATGTAG GATTAGAAAG	300
TGAAGACTTT TTAATAATTAT TATCAGGACT TCGGATAAAA AAGGGTATTG TCGTCAATGT	360
CGTGGATGTA TTTGACTTTG AAGGATCATT TATTAATGCA GTTAAACGTA TTGTCGGAAA	420

TCGAGTTAAA GAATGGTTAA AACGAACAGC AAGAAAATAT GGTTTGGAAG CTGACGATGT 540
 CGTATTAATT TCAGCTGAAA AAGGCTGGGG CATAGACGAC TTATTATCAT CAATTGCGAA 600
 5 TATTCGAGAA AATGAAGATG TGTATATTGT AGGGACAACG AATGTTGGGA AATCTACATT 660
 GATTAATAAA CTGATTGAAG CTAGTGTGG TGAAAAAGAT GTAGTAACAA CTTCAAGATT 720
 CCCTGGAACA ACTTTAGATA TGATAGATAT TCCTTTAGAT GAAACATCAT TTATGTATGA 780
 10 TACACCAGGT ATTATTCAAG ATCACCAAAT GACGCATTTA GTTAGTGAAA AAGAATTGAA 840
 AATTATTATG CCTAAGAAAG AAATAAAACA ACGCGTATAT CAATTAAATG AGGCGCAGAC 900
 ATTATTCTTC GCGGTCTAG CGCGCATAGA TTATGTATCA GGTGGTAAAC GTCCGTTAGT 960
 15 TTGTTTCTTT TCTAATGACT TGAATATACA TCGTACTAAM ACGGAGAAGG CTAATGATTT 1020
 ATGGCGTAAT CAACTTGGCG ATTTATTAAC GCCACCTGGA AATCCACAAA ATTTTGATCT 1080
 20 TAATGAGGTA AAGGCTGTTA GACTTGAAAC AGGCAAAGAG AAACGCGATG TTATGATCTC 1140
 TGGTCTAGGC TTTATAACTA TAGGACCAGG GGCTAAAGTA ATCGTTCGTG TTCCTAAAAA 1200
 TGTGAKGTT/ GTATTAAGAA ATTCTATTTT ATAAGGTGaT TAAAAAATG AAATTGcAG 1260
 25 TTATAGGAAA TCCTATTTCA CATTCTTGT CGCCCGTTAT GCATAGAGCA AATTTTAATT 1320
 CTTTAGGATT AGATGATACT TATGAAGCTT TAAATATTnC CAATTGAAGA TTTTCATTTA 1380
 ATTAAAGAAA TTATTTGAA AAAAGAATTa GAAGGCTTTA ATATCACAAT TCCTCATAAA 1440
 30 GAACGTATCA TACCGTATTT AGATTATGTT GATGAACAAG CGATTAATGC AGGTGCAGTT 1500
 AACACTGTTT tGATAAAAGA TGGCAAGTGG ATAGGTATA ATACAGATGG TATTGGTTAT 1560
 GTTAAAGGAT TGCACAGCGT TTAnCCAGAT TTAGAAAATG CATACATTTT AATTTTGGGC 1620
 35 GCAGGT 1626

(2) INFORMATION FOR SEQ ID NO: 194:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 635 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

AGGGTTAATT GTCGGTTTAA TTGCAATGAA TAAGTTCCAT GTATTAGCTG GCTATAGAGC 60
 50 GAAATTCATC TTAATGGTGA TTTTAACTAT GATGGTCTTC GTACTTATTA ATACGTATTT 120
 ACTAAGACAG GTAAAATCTA TCGGTATGTT CTTAATGATT GCTGCATTGG GTCTATACTT 180

GTCTTATATC GATAACATGT TCTTCAATTA TTTAAATGCA GAGCATCCTA TAGGCTTGGT 300
 GCTAGTAATA TTAACAGTAC TTGTGATTAT TGGCTTTGTA CTGAACATGT TTATAAAACA 360
 5 CTTTAAGAAA GAGAGATTAA TCTAATGTTG ATGAATAGCG TGATTGCTTT AACTTTTTTA 420
 ACAGCATCTA GCAATAATGG CGGACTTAAT ATTGATGTGC AACAAGAAGA GGAAAAGCGA 480
 ATCAATAATG ATTTAAATCA ATATGATACA ACGCTATTTA ATAAAGACAG CAAAGCGGTT 540
 10 AATGATGCCA TTGCTAAGCA GAAAAAGAA CGACAACAAC AAATAAAAAA TGATATGTTT 600
 CAAAATCAAG CGAGTCACTC GACTCGCTTG AATGA 635

(2) INFORMATION FOR SEQ ID NO: 195:

15 (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 13715 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

25 CTGAAATGGG TATTATTGT CTTCTTCATC ATAAAGTAAT AAAGATTGTT CATCATTGCG 60
 ACGTTGCCAA TTTTCATTTG GCGATCTTGG AACGTATAAT AATTGCCTAG TAGCATACGC 120
 TTTGATTGAA ACATACAAGT CATTCCCTGA ACTTGCACCT CAATTTCCAC ATTTGAATTT 180
 30 TCTTTTATAT TAATAATTTT ATCCAAATTC AGCTCACGTG CTAAGACAGC TCTTGATGCG 240
 CCTCTTTTAC CCCAGTAATT ACATTGAAAA TGATTAGTTA CTAACGTCTC TGCATTCCAA 300
 TGAAGTGGTA TTGGATTTTC TTGCGCCTTC ACATACATTA CTAAGTCTGG ATCCCCGAAA 360
 35 ATAATTCTGT CAACTCGTAT TTCATGTAAA AAATTAATAT AATCTTCTAC AGCATCTAAA 420
 TGATAATTAT GAAATAATCC ATTCACTGCC GCATATACTT TTTTATCGTT TTTGTGAGCT 480
 AATGCGACAG CCTCTGTCAT TTGTTGTCTA TTGAATTCCC CTGGAAGTCT TAAACCAAAC 540
 40 TTTTGCTCGC CAATTACAAA AGCATCTGCA CCTAAATCAA TAAGTGTTTC CATATGGCTT 600
 AATGACTTGG GTGTGACAAG TAATTCTGTC ATAGTCATTC TCCTTTAATT GAAATCGCTA 660
 45 ATCCATCGTC TATATTTAAA AAATTCGTTG TATATCCTGG TTGCTTTATT AACCACTCAT 720
 TATAATCTTG AACCTTTTTA ACCATTTGTC TTACATTTCT CGATCTAACA ATCCCAATAT 780
 CCGATACAAA ACCGTGATAT AAAACATTAT CTGTAATTAC GAGACCTTGG TGCTTTAAAA 840
 50 GTGGTGATA TATTTCAAAA AATTTCTTTG ATTGCGCTTT TGCTGCATCA ATAAATATCA 900
 TATCATAAAC TTTGTCATTT ACATTTTCAA ATTGCTCTAA AGCATTACCT TCAATAATTC 960

55

EP 0 786 519 A2

	CATTACGCTC TATCGTTGTG ACATGAATGT CATCAGATAT AGAAGCGAAT TGCATAGAAC	1080
	TATAGCCGAT TGCTGTACCA ATTTCTAAAA TATTTTAAAC ATTATTCATA CGAATTAATT	1140
5	GCTTAATTAA ATCTAATGTT AAACGATCTA CAATTGGCAC TTCATTTACC TCGGCAAATT	1200
	CACGCAAAC TTCGATTGAA CTATTTTGAT GTTGATGTAA ATCTATTAA TATTTTTTAT	1260
	TTAGGTCATC CATGTTTTAA ACTTCCTTTA TGTAAATAA GTCAATATGA TTATGACAAT	1320
10	AAAATAAATC AGCCTTCACA ATTGATTATA ATTTTGCCAA CCAATTAAAT GACTGATTTT	1380
	GTGTTAGACG CAAAGCTATT TTATTTATAG AAGCGAATCA TTCATATAAA ATTTAACTTT	1440
	AGATATTTTA CCATATTTTC AATAAAATTA TAAGCGTTAA TTATTTATAC ATTGCTTGAC	1500
15	TTAAAAATA CTCTTGCCCTC CCCATCTTTA AGGTTAGCAA GAGTAAAATC TTTTAAATTA	1560
	TTCTTCCATT TCAGTATTTA CAACTTCTTC AATCATGTCC CATTCTTCAT CAGTTTCGAT	1620
	TGGTACTAAC TTACCACCGT CACCTGACTC ATCTGGTTCA TTGATCATTG GTACAAGCTC	1680
20	AATCATATCG TCTTCATCTG ATTGAGCACC TTCTTCAGCT AAGATAACAT ACTCTTTTTT	1740
	GAATTCAGGA TGATAAAATT CTAAACTTT TCGGTATAAA ACTTCATTTT CCTCTTCATC	1800
25	GAATAAAGTT AATAATTCTT CTTGCTTATT AATTTCTAGT TGTGAATCAT GATTATGTTC	1860
	AGTCATAGTA AAATCTCCTT TTAATGTAGT GAATCTAAAT AGCCTTGTA AATAAATACC	1920
	GCTGCCATTT TATCAATCAC TTGTTTTCTT TTTTGTCTTG AAACATCTGC TTCTAATAAT	1980
30	GATCGTTCAG CAGCCATTGT GCTTAATCTT TCATCCCACA TCACAATCTC AATAGAAGGA	2040
	TAAGCTTCTA ATAATTTTTT TTTATATGTT AACGAAGCTT CGCCTCGAAA TCCTATTGAA	2100
	TTATTCATGT TTTTAGGTAG TCCTATTACG ACTGTACCCA CATTATGTTT TTTAATAATG	2160
35	TCTACTAATT GGTCAATACC TAATTCATTA TTTTCTTCAT TGATTGCGAG TGTGTCTAAT	2220
	CCTTGTGCCG TCCAACCCAT TATATCACTA ATTGCAATTC CTACCGTTCT ACTACCGACA	2280
	TCGAGTCCTA AAATTTTATG TTGTAACATA AATTATTTAT TTTGCTCTTT TAAATAGTAA	2340
40	GAAACAAGCT CTTCCATAAT AACATCTCTA TCAATATGAC GAATTTGATT TCTTGCTTCA	2400
	TTTTGGCGTG GAATATACGC AGGGTCACCT GATAATAAAT AACCTACAAT TTGGTTTACG	2460
	GCATTATATC CTCGTTTCATC TAATGTTCTGA TAAACATTAT TTAAAACATC TCTTACATCT	2520
45	TGCGTTGGAA GTTCTTCATA GTCGAATTTT ATTGTTTTAT CAAAGTTTTT CATTTGCGAC	2580
	ACTCCTTTAA TTACAAATAT AACTCACTAT CATCATACAA TATTATGGCT TTAAATTATA	2640
50	GATTTTTAAT GTAATCTTTA ATAAAGCTTA ATGATTTTGA GATATTTTCA GGTGTGTAC	2700
	CGCCACCTTG AGCCATATCT GGACGACCGC CACCTTTACC ACCAACGATT GGTGCCATTT	2760

55

	TCGATACTTT ATCATCAACA TTA	CTTGCAA GAATGATAAT TGTATCTTGT AGTTTAGATT	2880
	TAAAATCGTC CATTGTCGAG CGAATTGCTT	TCGCATTGG TACATCCACT TCAGTAACCA	2940
5	ATACTTTATA GCCATTGATT TCTTCAACTT	GATCTTCAAT ATTACCCATT TTAAGTGATG	3000
	TGATTTCTTT GTCACGTTGC TCTAATTGTT	TTAATAATGC TTTTCTTCA TCTTGTAATT	3060
	GTGTTAACTT ATCGACTACT TGATCATCAG	ATTTCACTTT CAGCTGTGAT TTCATCGTAT	3120
10	TAAATTTCTC TTGAATATCT TCTAAATATA	AGAAAGCTGC TTTACCTGTT AATGCTTCAA	3180
	TACGACGCAC ACCAGCTCCT GTACCTGACT	CACTTACTAT TTTGAATAAG CCAATTTTCA	3240
	AAGTATTGCG GACATGAATA CCACCACATA	ATTCAATTGA AAATGGTGCC ATATTTACTA	3300
15	CACGCACAAC ATCACCATAT TTTTCACCGA	ATAATGCCAT TGCGCCCAT TCTTTAGCTG	3360
	AAGCAATATC CATTTCTTGA ATGTTAACGT	CAATACCTTT CCAAATTTCT TCATTTACTA	3420
	AGCGTTCAAC TTGATCAATT TCATCATTAG	TCATTGGACC AAAATGAGAG AAATCAAAAC	3480
20	GTAAACGATC TGCTTCTACT AGTGAACCAG	CTTGGTTAAC ATGATCACCC AGTACTGATT	3540
	TCAACGCTGC ATGTAATAAA TGTGTTGCAC	TATGGTTCTT TTGAATGTCA CGTCGATCAT	3600
25	TTTGGTTCAC TTCAGCAGAC ACTGTAGCGC	CAACATTAC TTGGCCAAAT TGTACTACTC	3660
	CTTTATGCAA GTTTTGACCA TTTGGTGCTT	TGGTTACTTC ACTAACAGCA ATTTCAAAT	3720
	TGTCATTATA AACAAATACCT GTATCCGCAA	CTTGTCCACC ACTGATTGCA TAAATGGTG	3780
30	TTTCCGTTAA CATGAAGTAT ACTGTTTCAC	CCGCTTCAAC TTGTGAAACT TCTTCACCAT	3840
	TGTATATCAA GTGTGTTAGT GTTGTGTTGAG	CTGTCGCAGT ATCATAACCA ACAAAGTAC	3900
	TGTCAGATGT AATATTTTTT AATACTTCAC	TTTGAACCTG CATTGATTGA GAATTTTGAC	3960
35	GTGCTTGACG TGCACGATCA CGTTGTTGTT	GCATTTCTGA CTCGAATGTT GTCATATCAA	4020
	CTTTCAATCC TGCTTGCACT GCTATTTCTT	CAGTTAATTC AATTGGGAAC CCATACGTAT	4080
	CATACAATTT AAATGCATCT TTCCCATTA	TTTCATTGT TGTGCTTTA GCTTTTTTAA	4140
40	TTAATTCATT TAAAATCGCT AAACCATCTT	CTAATGTTTC ATGGAATCGT TCTTCTTCAG	4200
	ACTTTATAAC ACGCTTAATG AAATCTGCTT	TTTCCTTAAC ATTTGGATAA TATGGTTCCA	4260
45	TAATGTCTGC AACAAATATCA ACAAGTTTGT	ACATAAATGG CTCATTGATT CCTAACGTTT	4320
	GACTAAAACG AACGGCACGA CGTAACAATC	GACGTAATAC ATACCCTCTA CTTTCATTGG	4380
	CAGGTAATGC ACCATCAGAA ATTGCAAATG	CAATCGTACG AATGTGGTCA GCAATTACTT	4440
50	TAAATGCCAC ATCTTGTTG TTGTTTACTA	AATATTGTTT ACCTGATACT TTTTCGATTT	4500
	CATTCATTAT AGGCATAAAT AAATCTGTTT	CATAGTTAGT ACGTACATTT TGAGAACTG	4560

55

	TATGATCTTT ATTATGATTG AATTCATAA ATACTAAGTT CCATACTTCA AGATAGCGTT	4680
	CATTTTCTCC ACCTGGATAC ATTTCTTCTG CCGGATCGTC TTGTCCATAT GCTTCTCCGC	4740
5	GATCATAGAA AATCTCAGTG TTCGGTCCTG AAGGCCCTTC ACCAATATCC CAGAAGTTAC	4800
	CTTCAATGCG AATAATACGA CTTTCTTCAA GCCCAATATC TTTATGCCAA ATGTTGTATG	4860
	CTTCCATATC TTCCGGATGA ATCGTAACGT ACAATTTATC TGGCTCCATA CCCATCCATT	4920
10	TATCACTCGT TAAAAATTCC CAAGCAAAT CAATCGCTTC TTGTTTAAAA TAATCACCAA	4980
	TTGAGAAGTT ACCTAACATT TCAAAGAATG TATGGTGACG CGCTGTGAAA CCAACATTTT	5040
	CAATATCAAT TGTACGAATA GCTTTTTGAG AGTTTACAAT TCTTGGCTTT TTAGGTGTTT	5100
15	CACGTCCATC AAAATATTTT TTTAATGTTG CTACACCTGA ATTAATCCAT AATAATGTAT	5160
	CATCATCAAT TGGCACTAAT GGTGCAGAAG GTTCAACCAT ATGTCCTTTT TCAACAAAGA	5220
20	AATCTAGATA TTTTGTCTA ATTTCACTCG CTTTAACTT TTTTCATCATT TACACATCCT	5280
	ATTTACTGTT TTTAAATTAC CATTCCATAA AAATTGATGA CACAGATAGT CGATTTGCAA	5340
	AACTAGTATA AATCAATATC ATTTTTTATT ATTAAAAAT AAAAAACGCC CATCCTCAAA	5400
25	AGGGACGAAC GTTATCGCGG TACCACCCTA GTTATAAATG CAATTCAACA CATTATCAC	5460
	TTTAATTCGA CTATACAGTT GTGCATAAAG TAGCGTTCAC TAATGTTTGT TGTACTTTTC	5520
	ACCAACCACT ACATCTCTGA TAAACAAATC ATTAACACT CATCTTTATA CGAATTTAAT	5580
30	TCTATTTTAG TTACATTTAC GCTTGTGTG AACGTTCTAT AAAGTCATAC GCGGTGATTT	5640
	CTCCCATATT AATCATTGGG TCAATTTTAA ACATTGTAGC TTCCGTTAAT ACATTGTAT	5700
	CTGTTTTTGT TGAATCAGAC ATAACTTCTT CACTATCATT CGATGACATT GGCGTTCTA	5760
35	CTTGATCATC TATTGTCGTT TGTGAAGCTC CTGTATCATT AGTTGCTGTG TTTCCAGCA	5820
	TTTCTTCATC TTCTGAATTA AAATAATTTT TCAACAATGT ACATAATTGT GTTAAACGCG	5880
	CTTGACCATT TGTTTTCAAT CCAATATCAA ATGCTTCCGG ATCACCAAGT AAAACTAAAC	5940
40	TCGTTTTCGC TCTAGTTAAA CCAGTATATA ATATCGGTCT TTGTAACATT CTAAAATACT	6000
	GTTTAACAAT AGGCATGATA ACAATAGGAA ATTCTGAACC TTGTGATTTA TGGATTGATG	6060
45	TACAATAAGC ATGTGTTAAT TCCATCATAT CTGTTTCGT AAATGTAATT TCATTACCTT	6120
	CAAAATCCAC AACAGTACA TCTTTATTAA GGGCATTTC TTTGCCCCAA AAAATACCAA	6180
	CAATAACTCC TATGTCACCA TTGAATATGT TATCATTTGG CCTATTACA AGTTGTAATA	6240
50	CTTTGTCACC TTTTCTAAAG ACTACATCAC CAACTCAAT TTCTCGTGTG TCTTTCTTTT	6300
	TAGGGTTTAA AATATCTTGT AAACTTGAT TTAAACGTTT AATACCGCA TTTCTTTTAT	6360

55

	CTACCTTCTC AACAACTGTT GGTATTTGGT TTGCCTGACA GTTAATAAAA CTTCTATCAT	6480
	GAAAACGCTG TGTAATATCA ATTTTCTGAC CCAACTTCAT TCGATGTGCT AATTCTATAA	6540
5	TGCTTGAACC ATCTTGTGTA CGATATACTT CAGTCAGATT TACTCGTGGT ATAGCTTTTCG	6600
	ATTCAATTAA ATCTTTAAAT ACTTGACCAG GACCTACAGA AGGCAATTGG TCCTCATCAC	6660
	CTACAAATAT CAATTGTGCA TCTAAAGGAA CTGCACTTAA AAATTGGTGG AACAAACCAAG	6720
10	TATCTACCAT AGACATCTCA TCAATGATTA TGAGTCGTGC GTTTATTTCA TTTTCTAATA	6780
	TATCCTCTGG CTTTGTGTCT TGATTCCAAC CTATTAAACG ATGAATCGTC ATTGCTTCTA	6840
	ATCCAGTTGA CTCTTGTAGT CTCTTAGACG CTCTTCCTGT TGGCGCTGCT AATACAACTG	6900
15	GATAATCATC ATTGACATAA TCATCATAAT CTAATGATAA GCCATGAATC TCAGCATATA	6960
	ATTCAACAAT ACCTTTAATT ACTGTCGTTT TTCTGTTC CCGTCCACCG GTTAATAGCA	7020
20	TCACCTTAGA ATTGATAGCC GTTTGCAAAG CTTCTTTTTG TGAAGCTGCA TAGTTCACCTT	7080
	GATTGCGATC TTCTATTICA CCAATATGCA TTTGTAAATC TGACTGTTCA ATTTCTGTAA	7140
	GTTCATTTGT ATGCGTCTTT ATTCTGAATA AGTTTTGAAC ACTTTTGATT TCaGAATAAT	7200
25	ACAACTTGG AATTGCAACT TGTTCaTTGT CAATAATTAG TCGTTTTTCC TCATTTAAGT	7260
	ATTGCAACAT TTCGTCTAAT TTTTCAGGTT CGATGACCTC TTCATCTTga TAATTTAATA	7320
	CATCAACCGT TAAATCTATA ACAACATTGA TAGGCAAATA TGTATGTCCC TGTTTAATAC	7380
30	ATTCTTCTTC TAACGTATAG AGCAACGCAG CTTTTAATCG TTCATTATCG TTATAAGCGA	7440
	TACCAATATT TCTAGCAAGT TGATCTGCTT TATTAAAACC AATACCTTTA ATATCATAAA	7500
	TCAATTGATA TGGATTTGCA TCTAAAATAG TCAGTGTATC GCCGAGATAA AACTGATAAA	7560
35	TTGCCATTGA AAGTTTAGGA CCAAACCCTA AATCATGTAA ACGAATCATT ATTTTTTCAG	7620
	ATTCTTGATT TGCTGAAATT TGTCTGCAA TTTGTTTCTG TTTCTTTTGA GATAATCCCG	7680
	AAACTTTTTTC TAGCACTGAA TGGTCATCTA ATATATCATT TATCGCATTG TCACCTAATG	7740
40	TATTAACAAT ATTTTGAGCT GTCTTTTTAC CTACACCTTT AAACAAATCA CTAGATAAAT	7800
	AACCTATAAT TGCTTCTTTC GTTTGTGGCA TTTCTTTTTC AAAAGTCTCT GCTTTTAATT	7860
45	GTTTACCATA ACGTGGATGA TCAACAACTT GCCCTTTAAA TGTGTAGACA TCGCCTTCAA	7920
	CAATATTCGG AAGAAACCCT ACAACAGTTG GCATTGTATC AAAGTCTTCA TTTGTTTCAA	7980
	TAGTATCTAC TTTAAGCACT GTATAAAAAT TATCACTGTT TTGAAACAAT ATCGCTTCAA	8040
50	CAGTACCTTT GATCATTGAA TAATCAAATA GTGTAGGGTC TGACATGTTA CTCCTCCTCT	8100
	TTCATTTTAG TGAATGTTTT CAGCGCATGC TGACTTAATA AGTGTTTAGG GTCGATAGTC	8160

55

	AAGCCCAAAT TGTATCTTGC ATCAACATGA TTTTATCAA TCGTTAATAC ATGTTTAAGT	8280
	TGAGTTATGG CTTCAATAAA CATTTCTAAT TGACATAATA CAAGACCATA TTGAAATTGA	8340
5	ACTTCTGCAT CTTTGTCTTT ATCTAGTTCC GCAGCAGTCA TTAAATACGG CAATGCCAAG	8400
	CTTAAATGAT TCTAACTGAT TAAACGCCAT ACCGATCATA TAATTACAAT CAACTTGTTT	8460
	AATCTCTGTT TGTAATGCTT GTTGATATAA TTTAATAGCT TCTTGATAAC GTTGCTGATT	8520
10	ATAATATACA TTTGCTAGAT TAAAAAATAC GACGCCATTG TTCGGATCTA TTGTAAAAGC	8580
	TTTTTGGAAA AAACGCTCTG CCTTTCAy CyCATTCgCA TCAGCAAGTA CGATmCCaGC	8640
	ATTAATATAA TTTTCAATAA TTGTAGGATT TTCTTCGATA TTTCCGAACA ATGCTTGTA	8700
15	CGCTTCTTCT ATTTTCCAT TTTGTATGTA TTGATAAATT GTTGTTGAT CTATCATTTA	8760
	CGAACCTCAT TTCTCATCAA TTATAACATC TTGATAAATT GTATGTCTCG AATCACTTAA	8820
20	CAACGAATAA AATATAATCT AATATCATCT TCATTCATGA AAAAGCGGGA ATGGAATAGA	8880
	AATGCTTAAG AACCATTAAAC GGTATTATTAT GTAATGGTTC TTCCACATTA GCCACCACTA	8940
	TTATGTACTT AAAAATAAGA ATACATAATT AGATTCATGC ATAGGGAGTG GGACAGAAAT	9000
25	GATATTTTAA CAAAATTAAA TTCGTTATCC CCAACTGGCA TTGCCTGTAG AATTTCTTTA	9060
	CGAAATTCTC TATGTTGTGG TCCCGCCAAT ATAACATTGT AGAGCCTAGG ACATTGTGAT	9120
	GTCCAGACT CTATCCTCAT GAATTATTCT CATCAAAAAC TGTCTTTCGT CATTTTCAAC	9180
30	GTTGAACTT CAAATAAGTA ATTTATTGTT GCCATTGTTT ATACAACATA ATTTAATTGA	9240
	CCTTCATTTT TGAACACATC GTCAATTGTT GCACCACCAA GACACACATC ACCTTGATAA	9300
	AAAACAACTG CTTGTCCAGG TGTGATTGCT CTTACTGGCT CAGCAAAAGT AACACGTAGg	9360
35	CATGGtCGTT TTCACGTTTC ACAAAAACCT TCGTATCTTT TTGGCGATAT CTAAATTTAG	9420
	CTGtACATTC AAAACCTTGA TCTAAGTCAT TATCTTCTGG ATTTACAAAT GAATAGTCTG	9480
	AAGCAATTAA GTAATCACTG TATAATGCAT CGTGATGGAA TCCTTGTTCT ACATATAAAA	9540
40	CATTATCTTT TAGGTTTTTA CCGACAACAA ACCAAGGATC GCCATCTCCA CCTATACCTA	9600
	ATCCATGTCT TTGTCCTATT GTGTAATACA TCAAACCACT ATGTTTACCC ATTTTCTTAC	9660
45	CATCAAGTGT TATCATATCA CCCGGTTGTG CAGGTAAATA TTGTGATAAA AATGTTTTAA	9720
	AGTTTTTTTC GCCGATAAAA CAAATGCCTG TAGAATCTTT TTTCTTAGCA GTAACAAGTC	9780
	CTTGTTCTTC AGCAATTCGA CGCACTTCAC TCTTTTCGAT GTCGCCAATT GGGaACATCA	9840
50	CTTTTGAAAG TTGTTGTTGA GATAATTGAT TCAAGAAGTA TGTTTGATCT TTATTATTAT	9900
	CTACACCACG TAACATTTCA ACATGACCAT CTTCATGACG ATGTATGCGT GCGTAATGTC	9960

55

	TTTCTTTATT ACACATAACG TCTGGATTTG GAGTACGACC TTTTTTGTAT TCATCTAAGA	10080
	AATACGTAAA GACTTTATCC CAATATTCTT TTTCAAAT AACAGCGTAA TACGGAATGC	10140
5	CAATTGATT ACACACTTCA ATAACATCGT TGTAATCTTC AGTTGCAGTA CATACGCCAT	10200
	TTTCGTCAGT GTCATCCAG TTTTTCATAA ATATGCCAAT GACATCATAA CCTTGTTCTT	10260
	TTAAGACGTG GGCTGTTACA GAACTATCTA CACCGCCTGA CATAACCAAC ACAACACGTA	10320
10	TATCTTTATT TGACAATTAT GACTCCTCCT TAAATTTAA ATATATTTTA TGAATTCAG	10380
	CTACAATTGC ATTAATTTCA TTTTCAGTAG TCAATTCGTT AAAACTAAAT CGAATCGAAT	10440
	GATTTGATCG CTCCTCATCT TCGAACATTG CATCTAAAAC ATGCGACGGT TGTGTAGAGC	10500
15	CTGCTGTACA TGCAGATCCA GACGACACAT AGATTTGTGC CATATCCAAC AATGTTAACA	10560
	TCGTTTCAAC TTCAACAAAC GGAAAATATA GATTTACAAT ATGGCCTGTA GCATCCGTCA	10620
	TTGAACCATT TAATTCAAAT GGAATCGCTC TTCTTGTA TTTAACTAAA AATTGTTCTT	10680
20	TTAAATTCAT TAAATGAATA TTGTTATCGT CTCGATTCTT TTCTGCTAAT TGTAATGCTT	10740
	TAGCCATCCC AACAATTTGC GCAAGATTTT CAGTGCCTGC ACGGCGTTTC AATTCTTGTT	10800
25	CACCGCCAAG TTGAGGATAA TCTAGTGTA CATGGTCTTT AACTAGTAAT GCACCGACAC	10860
	CTTTTGGTCC GCCAACTTA TGAGCAGTAA TACTCATTGC GTCGATCTCA AATTCGTCAA	10920
	ACTTAACATC AAGATGTCCA ATTGCTTGAA CCGCATCAAC ATGGAAATAT GCATTTGTCT	10980
30	CAGCAATAAT ATCTTGAATA TCATAAATTT GTtGCACTGT GCCAaCTTCA TTATTTACAA	11040
	ACATraTAGa TACTAAAATC GTCTTATCTG tAATTGTTTC TTCAAGTTGA TCTAAATCAA	11100
	TAGCACCTGT ATCATCAACA TCTAGATATG TTACATCAA ACCTTCTCGC TCTAATTGTT	11160
35	CAAAAACATG TAACACAGAA TGATGTTCAA TCTTCGATGT GATAATGTGA TTACCCAATT	11220
	GTTCAATTTGC TTTTACTATG CCTTTAATTG CCGTATTATT CGATTCTGTT GCGCCACTCG	11280
	TAAATATAAT TTCATGTGTA TCTGCACCAA GTAATTGTGC AATTGACGT CTTGACTCAT	11340
40	CTAAATATTT ACGCGCATCT CTCCCTTAG CATGTATTGA TGATGGATTA CCATAATGCG	11400
	AATGTAAAT CGTCATCATC GCATCTACTA CTTCAGGTTT TACTGGTGTG GTCGCAGCAT	11460
	AATCTGCATA AATTTCCATG TTGGACACT CCTCACAATT TTATCAATGT TCCAATAATA	11520
45	GCACCTTACA TACTATTTTT CTACTTTTCT GTTTAACTTT ATTTATAATG TTTTAAATTA	11580
	TATTTTACCA TTTTCTACAC ATGCTTTTCG ATAGGCTTTT TTAAGTTTAT CGCTTTATTC	11640
50	TTGTCTTTTT TATAAATTTT AGTATTTGCA GATATTTTTT TATTTGTAAA ATGTAACGTA	11700
	CTATTATTTT GGTATGAGC AATTTAATAT TTATCTGGTT ATTCGATTGG TATACTTCTT	11760

55

EP 0 786 519 A2

	GTCCCTATTC GAGAAGGTGA AGATGAACAA ACAGCAATTA ATAATATGGT TAATCTCGCA	11880
	CAACATTTAG ACGAATTATC ATATGAAAGA TATTGGATTG CTGAACACCA TAACGCTCCC	11940
5	AACCTAGTAA GTTCAGCAAC TGCTTTATTA ATTCAACATA CGTTAGAACA TACGAAACAC	12000
	ATACGTGTAG GTTCTGGAGG CATCATGTTA CCTAATCATG CTCCATTAAT CGTTGCGGAA	12060
	CAATTTGGCA CGATGGCAAC ATTATTTCCA AATCGTGTCTG ATTTAGGATT AGGACGTGCA	12120
10	CCTGGAACAG ATATGATGAC CGCAAGTGCA TTAAGACGAG ATCAACATGA TGGTGTTTAT	12180
	AAATTTCCAG AAGAGGTTTC ATTATTACAA CAATATTTCTG GCCCTGCTCA CCAACAAGCA	12240
	TATGTTCTGT CTTATCCAGC AGTAGGTAAA AATGTGCCTT TATACATTCT TGGTCTTCA	12300
15	ACAGATTCTG CACATTTAGC TGCTCGCAA GGGCTTCCAT ATGTGTTCTG TGGACATTTT	12360
	GCACCTCAAC AAATGAAAGA AGCTATCGAA ATTTACAAAA CGTTATTTGA ACCTTCTGAT	12420
20	GTATTAGACG AACCTTATGT TATTGTATGT TTAAATACAA TCGTTGCTGA AAATGATGAC	12480
	GAAGCACAAT ATTTAGCTTC ATCTATGGCA CAAGTAATGG TTAGTATCAC TCGTGGCAGA	12540
	ATGCAGCCCG TTCAACCGCC AACACATGAA CTACAAAATA TATTAACGCC GAGAGAATAC	12600
25	GCGATGGCTA TGGAAAGACA GAAAATATCA TTAATAGGTT CAGAAAATAC TGTTCACAA	12660
	AAAATCAAG ATTTTATGGA AACTTATGGT GAAGTCAACG AAATTATGGC AATAAGTTAT	12720
	ATTTATGATA AAGATATGCA ATTAGACTCT TATCGTCGGT TCAAGAATGT TATAAATCAG	12780
30	ATAAATGAAA AAAACACTTT ATAATGTGAT AAATAAACTA AGTGAAAGTA TGTATCCATA	12840
	ATATTAATAA AAATATACAG TAACAGCATT TTGAATGAAA GATGTCTTTA TTGTTCAATC	12900
	ATTTATTTTA GTAATGATTC AAATTCACCT AAAATyCTAA tGCAAATATG AAAGCGCCCC	12960
35	TTCAcTTTAC ACTGTGTAAG TGTTTATTTG ATGGGGCGCT TTCAAAATAT TGAAAAGCAT	13020
	ATCCéAAAT TAAAGAAAT TATTTCTCTT TATCTTCATT TTCTTTTTTC TCTTCGTTAT	13080
	TCGATCCTGT ATATTCAATT ATCTTATCTT TTACATTTTT AACTTGTTCA TTATCGCTAT	13140
40	TTTTAAATTT TTCTACGCGT CTTTAGCTTT ATCCATAAAA CTCATATTAA TCGCTCCTCT	13200
	TATATTTGAT TAGTTTAATT GAACTTATTT TTTAAGTTTA TCAATTGCAT CAGTTATTTT	13260
	GTTTTTAGCA TTTTCAACAA CTTCTTTTGC TTTaCCAGTC GCTTTATCTT GCTGACCTTC	13320
45	TTTTTCTAAT TCTTTGTTAT CAGTAACGTT ACCTACTGTT TCTTTAACAT TTCCTTTAA	13380
	TTGATCGAAC TtACTTTCGT CTGCCATAGT GAAACCTCCT TGGATGTATA TATTTATATA	13440
50	CCACTAAGGA GGTTCGCTmm mCAyymyAAT ATGAAGTTTT TATGTTATAG TATAGTATTT	13500
	ATACGATTAA ATATAAAACA TGTATCCGTC TAAATCTTCA CTTGTATCTA CATATTCGCG	13560

55

TAGTTGTTTT TGC GCAGGTG GTTCTGATTC AATACTTTCA ACAAATGTAA TTGGACCTTC 13680
 TAACAGTCTT ATAATATCCC CTGCTGAGAT TTCTT 13715

5 (2) INFORMATION FOR SEQ ID NO: 196:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 873 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

AAATCCATAA TGT CATGATA ATCTGCATAT GCTTCATATA ATTCAATCAT TGTGAATTCA 60
 GGGTTATGTC TAGTTGATAC ACCTTCATTA CGGAATACTC TACCAATTTT ATATACTTTT 120
 20 TCAAGTCCAC CGACAATTAm ACGTTTTTAAA TGCAACYCAA TAGCAATACG CATGTATAaC 180
 GTTGCATCTA ATGCATTATG ATGTGTTACA AATGGTCTAG CAGCTGCTCC ACCAGCAATT 240
 TGGTGCATCA TAGGTGTTTC TACTTCCAAG AAACCTTIAT TATTAAATA ATTACGCATT 300
 25 TCTTGAATGA TTTTACTACG ATTAATAAAT GTACGAGTGC TATCTTCGTT CGTAATTAAA 360
 TCTAAATATC TTTGACGATA tCTCTGTTCA ATATCCTGTA AACCGTGGAA TTTATCCGGT 420
 AATGGTCGCA ATGATTTAGT TAGTAGCGTG AATTTCTTCG CTTTAACCGA TAATTCGCCA 480
 30 GTATTTGTTT TGAACATTAC ACCTTCAACA CCAACGATAT CGCCTAAATC AGCATTTTTTT 540
 CATAAATCAA ATTCGTCATC GCCAACTTGA TCTTTACGAA CGTAAATTTG AATTGTCCA 600
 GCTAAGTCCT GAACGTGTGC AAATCCTGCT TTACCTTTAC CACGCTTAGT CATTAATCGT 660
 35 CCAGCTATAG CGACATGACT ATCCGCTTCT TTTTCTACCA ATTCTTCTTT AGAATACTGG 720
 TCCGACTCTT CTTTCAAATC ACTAGATAAA CCTGAACGGT CAAATTTAGA ACCAAACGGG 780
 TCTATACCAA GATCATATAA TTCTTGTAAT TtTTGACGTC GAACCAACAT TTGGTCATTC 840
 40 ATTTCTTCTG ACATAACTtT CTCTCCTTIA ACT 873

(2) INFORMATION FOR SEQ ID NO: 197:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 452 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:

55

ACCATAATAT GAATGGCTTC AGGATCAAAA TAAAGACCAA CTTCACTGCC TACTTCAGCT 120
 TTTTtagTCG TTTGTATTAC CCATTCATAA CCTTTATTGT CTATACAACA TATTTTCATAG 180
 5 TGGACCCCTC TAAATAACAT AGAATCAACA GTTGCTTTAA ATAATCCTTC TTCAGCTTTG 240
 ATTAATGATA TATCTTCTGG TCGAATAACG ACTTCTACTT TTTTATTTTC AGGAATACCC 300
 ATATCGACAC ATTGGAATC TTGCCCATAA ATATTACGA CATAATCTCT AACCATGCGC 360
 10 CCTTCAACAA TATTAGATT CCAATAAAA TCAGCTACAA ATCGATTAC TGGTTCGTCA 420
 TaTATATCTG TTGGTGTGCC AAATTGTTGA AT 452

(2) INFORMATION FOR SEQ ID NO: 198:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:

25 TAGGTTGGGT TCTAACATAC GATAAAGCTC AACAAATCAA CACAGCTTTC TTTGTAAAAT 60
 TGTTTAATAC TGCATTAGCA GAACGTGATT ATTATTTTAA TATAGATGGA ACAAATGCTT 120
 TTAGATTATT TAATGCTGAA GGTGATGGTG TTGGGGGATT AACATCGAC AATTACGATG 180
 30 GTCATTTGTT GATTCAATGG TACTCAAAAG GTATTTATAA ATTTAAATAT GCCATTCTTG 240
 AAGCGGTTAG AAAAGTATTT GATTATAAAT CTATTTACGA AAAAGTAAGA TTTAAAGACA 300
 GCGAATATAG TGGTGGTTTT GTTGAAGGAG ATGCACCTGa GTTCCCAATT GTTATCGAAG 360
 35 AAAACTTCAC ATTTTATAAT GTAGACCTTG AAGATGGTTT GATGACAGGT ATCTTTTTAG 420
 ATCAAAAAGA AGTGCACAAG AaATTAAGGG ATCAATATGC CAAAGAACGC CATGTTTTAA 480
 ACTTATTTAG TTATACAGGT GCTTTTTCTG CAATAGCAGC AAGTGAGGCA TCTTCAACAA 540
 40 CAAGTGTAGA TTGGCTAAT CGTTCTCGTA GTTTAACTGA AGAAAATTTT GGATTAAATG 600
 CTATTGATCC TAAATCCCAA TATATTTATG TCATGGACAC TTTTGATTTC TATAAATATG 660
 CTGCACGACA TGGACATAGT TATGACACGA TCGTGATTGA TCCACCTAGC TTTGCGCGTA 720
 45 ACAAAAAACG TACATTTTCA GTGCAAAAAG ATTATGACAA ATTAATTAAT GGCGCCTTAA 780
 ATATCTTATC ATCTGAAGGA ACATTATTGT TATGTACAAA CGCAAGTGTA TATCCATTAA 840
 50 AGCAATTTAA AAATACTATT AAAAAGACGC TTGAAGAGAG TGGCGTTGAT TATGAATTAA 900
 CTGAAGTTAT GGGATTACCA AAAGATTTTA AAACGCATCC ACATTATAAG CCATCTAAT 960

55

TATTGAGAAA AAGAAGGGTG ATAATATTAT GGGATTCAAA AACAAATTTAA CATCAAATTT 1080
 AACAAATAAA ATCGGTAATT CAGTCTTTAA AATAGAAAAT GTTGACGGAA AAGGTGCAAT 1140
 5 GCCAACGACG ATTCAAGAAT TGAGAGAAAG ACGACAACGT GCTGAAGCAA TTGTAAAGAG 1200
 AAAGTCTTTA ATGTCATCAA CAATGAGCGT TGTCCAATT CCGGGTTTAG ATTTTGGTGT 1260
 TGATTTAAAA TTAATGAAAG ATATTATCGA AGATGTTAAT AAAATTTATG GTTTAGATCA 1320
 10 TAAGCAAGTT AATAGCCTTG GGGATGATGT GAAAGAAAGA ATTATGTCTG CAGCAGCAAT 1380
 TCAAGGTAGT CAATTTATTG GTAAAGAAT TTCAAATGCA TTTTAAAAA TTGTAATTAG 1440
 AGATGTAGCT AAACGTACTG CTGCAAAACA AACAAATGG TTTCTGTG TAGGACAAGC 1500
 15 TGTGTCTGCA TCTATTAGTT ACTATTTTAT GAATAAAATT GGAAAAGATC ACATTCAAAA 1560
 ATGCGAAAAAT GTTATTAAAA ATGTCATGTA GGTGCTATAA TAGTTTTGCA ATTTGCAAAT 1620
 20 TTTACTGAAA CCGGTTTTAA ACGAATTGAA TTAAAGCAT GGTTTTGGTA AAGTTAATGT 1680
 ATAAACTAA GTTAGyATTG TAATAATATk GAAGATTCTA ACTATACGAA GGAGAAATGT 1740
 AATTATGGAA CAAAATTCAT ATGTAATCAT CGACGAGAmT GGTATTCACG CTAGACCAGC 1800
 25 AACAAATGTTA GTACAAACAG CTTCAAATTT CGATTCTGAT ATTCAATTAG AATATAACGG 1860
 TAAGAAAGTA AACTTAAAAAT CAATCATGGG TGTATGAGC CTTGGTGTG GTAAAGATGC 1920
 TGAAATTACA ATTTATGCTG ACGGTAGTGA TGAATCTGAC GCCATTCAAG CAATCAGTGA 1980
 30 CGTCTTATCA AAAGAAGGAT TGAATAATA ATCATGTCTA AATTAATTAA AGGTATTGCC 2040
 GCATCTGATG GTGTCGCAAT TGCTAAAGCT TATTTATTAG TTGAGCCAGA CTTAACATTC 2100
 GACAAAAATG AAAAAGTCAC TGATGTTGAA GGAGAAGTTG CAAAGTTCAA TAGCGCTATC 2160
 35 GAAGCTTCTA AAGTTGAGTT AACTAAAATT AGAAATAATG CAGAGGTTCA ACTAGGTGCT 2220
 GATAAAGCTG CTATCTTTGA TGCacaTTGG GGGGTGGTAG ATGACCCTGA ATTAATTCAA 2280
 CCAATCCAAG ATAAGATTAA AAATGAAA 2308

(2) INFORMATION FOR SEQ ID NO: 199:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5559 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

AAGTAATAAA TCGTCTCATT TGGCAACTGA CGCATAATTT CTTTAGCTAC TGTCAAACCT 60

	TTTAATTTTT AGTTTATCAT AACTAAGCAT TGGATTTTAG TATTATGCAC TGTGTTTACC	180
	ATTTTTGTCA TTATAATATT TATTTTAAAT CAGCCCACTA TCATATTGTC ATGTAATCTG	240
5	CTTATTAAAA AAATCCCTTC CAAGTTATTG TGTATCTCCA TTCAATTTAA TTTTGAAAGG	300
	AACATAACWT TTTAACTCAA AAGGGATTAA TTTnTAnTCT ACTTCATGGT CTGAACCAAA	360
	GAATGATTTA AACATGTGGA ATGTTGTTTC TCTGTTCATT GCTGCAATGG ATGTTGTTAA	420
10	TGGAATACCT TTAGGGCAAG CATTAAACACA GTTTTGTGAA TTACCACACT GCTGTAAGCC	480
	ACCAGTACCC ATTAATGCAT TTAAACGTTT ATCTTTAGTC ATAGATCCTG TTGGGTGCAA	540
	ATTAAACAAA CGAAGTTGCG AGATTGCTTG TGCACCAACG AaTTTATTAT TTTcAGTAAC	600
15	ATTAGGACAA ACCTCTAAAC ATACACCACA TGTcATACAT TTAGATAATT CATAAGCTGT	660
	TTGACGTTTT TTCTCTGGCA TACGTGGTCC CGGACCTAAA TCATACGTTT CATCAATTGG	720
20	GATCCATGCT TTCATACGTT TTAAGTTATC GAACATTCTA GAACGATCAA CTTGTAAGTC	780
	ACGGATAACT GGGAAAGTAT TCATTGGCTC TAAACGAATA GGTGTTTCTA ATTGATCAAC	840
	AATCGCAGAA CAAGATTGTC TTGCACGACC ATTGATAACC ATAGAACATG CTCCACATAC	900
25	TTCTTCTAAG CAGTTCATAT CCCAGACAAC AGGTGTTGTT TTTTCACCTT TAATATTAAC	960
	TGGGTTACGT CTAATTTCCA TTAAACAAGC AATGACGTTT AAATTTTCAC GATATGGAAT	1020
	TTCAAATGTT TCTTCATAAG GCTTAGAATC ACTTGATCTT TGTCGTTTAA TAATTAATTT	1080
30	TACTGTTTTT TGTTTCGGTT TAGATTGTGT TTCATGTTGT GGAGTGTTTT TCACTGATTG	1140
	TTcAGTCATT ATTTTTTACC CCCTTTAGAC TTACTTGTGT AATCACGTTT ACGAGGTGGT	1200
	ATTAAACTCA CATCGACGTC ATCATAAGTA AACTGCGGTT TTTCAAATGC GCCTTGGAAAT	1260
35	GAGGCCATTG TCGTTTTTAA CCACTCTTCA TCATTACGCT CTGGGAATTC TGGTTTATAA	1320
	TGGGCACCGC GTGATTCGTT ACGGTTATAT GCACCAATCG TAATAACACG TGCAAGTACT	1380
40	AACATGTTCC ATAGTTGACG GGTAAAGAAT ACCGCTTGGT TACTCCAAGT TTGAGTATCT	1440
	TCCATATCAA TATCTTCATA ACGTTTCATC AATTCAACAA TCTTTTATC TGTTTCTAAC	1500
	AGTTTTTTCAT TTTcACGAAC AACAGTTACA TTTGCTGTCA TAATTTcACC AAGTTcACGG	1560
45	TGTAATTTAT ATGCATTTTC TGTACCGCGC ATAGCTAATA ATTTATCAAA ACGTTCTTGC	1620
	TCTTCAGCTT TACGCTTTTC AAAAATACTT TCGTCCATAT CAGTATATGA TCGATCAATA	1680
	TTTGAAATAT AATCAATCGC GTTTGGACCT GCTACTGTAC CACCATAAAT CGCTGATAAC	1740
50	AATGAATTGG CACCTAAGCG GTTACCACCA TGTGAGAGA AGTCACATTC TCCAGCTGCA	1800
	AATAACCCTT TAATATTTGT CATTTGATCA TAATCTACAT ATAGACCACC CATTGAATAG	1860

55

EP 0 786 519 A2

	TAAATCTCAA	TGATACCACC	TAGTTTTACA	TCTAACTCAT	GTGGATCTTT	ATGTGACAAA	1980
	TCAAGATATA	CCATGTTTTTC	GCCATTTATA	CCTAATTTTT	GGTTAATACA	TACATCGAAA	2040
5	ATTTACGCG	TTGCGATATC	ACGAGGTACT	AAGTTACCAT	AATCAGGATA	TTTCTCTTCT	2100
	AAGAAGTACC	AAGGCTTACC	ATCTTTATAT	GTCCAAATTC	GTCCACCTTC	ACCACGTGCT	2160
	GATTCACTCA	TTAGTCGCAG	TTTATCATCA	CCAGGGATTG	CAGTAGGATG	AATTTGAATG	2220
10	AACTCACCAT	TAGCATAAAT	AGCGCCTTGT	TGGTAAACAA	TGGAAGCCGC	TGATCCTGTA	2280
	TTAATCATTG	AGTTTGTGT	TTTACCGAAA	ATAATACCAG	GGCCACCCGT	TGCCATAATA	2340
	ACTGCATCTG	AACCAAATGT	TTCAATCTCA	GCAGTTGTCA	TATTTTGTGC	AnCGATACCT	2400
15	CTTGCACTAT	CATCGTCACC	TTTAACTATG	CCAAGGAATT	CCCATCCTTC	ATACTTCGTA	2460
	ACTAATCCAT	CTACTTCATA	TGCACGAACT	TGTTTCATCCA	ATGCATATAA	TAATTGTTGT	2520
20	CCAGTTGTTG	CCCCTGCATA	TGCTGTTCTG	TGATGTAATG	TACCACCGAA	ACGTCTAAAA	2580
	TCTAATAGAC	CTTCATTGT	TCTATTGAAC	ATTACGCCCA	TACGGTCTAA	TAAATGAATA	2640
	ATTTTAGGTG	CTGCCTCTGT	CATCGCTTTA	ACAGGTGGTT	GGTTTGCAAG	GAAATCGCCA	2700
25	CCATACACTG	TATCATCAAA	GTGAATCCAA	GGAGAATCGC	CTTCCCCTTT	AGTATTGACC	2760
	GCACCATTAA	TGCCACCTTG	GGCACAAACA	GAGTGCGAAC	GCTTTACTGG	TACAACTGAG	2820
	AACAAATCTA	CATGTGCACC	TTTTTCTGCC	GCTTTAATTG	TTGACATTAA	GCCCGCTAGG	2880
30	CCACCTCCGA	CAACAATAAG	ATGTTTCTCT	GCCATAAAAA	TTTCACTCCC	CTAAATTTTC	2940
	AATCTATATT	TGTTAAATGC	GATGTATTAC	ATAAAGGCAA	TAATTGCAGT	AACACCAATA	3000
	TACGAAATAA	CTAAAAATAC	GATTAATGAA	ACCATGTAA	ATACTCGTTG	TGATTTTGGA	3060
35	GATTGAAGTC	CACCCCAAGT	AACTAAGAAT	GACCATAAGC	CATTTGCAAA	GTGGAACACA	3120
	ACAGCAATAA	TACAAATAAT	ATAAAATATT	GCCCATCCAG	GATGTTGCAA	TGTTTCGTGC	3180
	ATTAAATCGT	AATTCACCTC	TTTGCCGTAA	AATGCTTTTT	GTAAACGTGT	TGCCATAAAA	3240
40	TGGATACCAA	TAAAGATAAA	TGTTAAGATA	CCACTCACTC	TTTGAAGAA	GAACATCCAG	3300
	TTTCTAAAAA	TCGAGTAATG	TCCAACATTT	TCTTTTGCTG	TAAATGCAAT	GTGTATACCA	3360
	AACAAACCGT	GATATAACAA	CGGAATGTAT	ATAAATAAAA	ATTCTACAAT	AATTAGAAAT	3420
45	GGTAATGATT	CCATAAAGTT	AGATGCCCTTA	TTAAACGCTT	CAGCACCTTG	TGTTGCTTGG	3480
	TGATTCACTA	ATAAATGAAC	GACCAAAAAT	GCCTTATTG	GGATAATACC	TAATAACGAG	3540
50	TGAATACGTC	TTAGATAAAA	TTCATTTTTT	GATTGAGCCA	AAAGGAGTCC	CCCCTGTGAA	3600
	CGAATATTTA	ATTTATTGAG	CTATTTATAT	TAAACGTACG	CTTAACCCCC	TAAAGTGATA	3660

55

EP 0 786 519 A2

	CGATCACCAA ACTGCATGTC GAACAATGTA ACATTTGGAT TCGATATTTA AAATTGCTTG	3780
	TGATGATAAA CTTTCTCATT TAGAAAACGC TTCCACGTAC ATTCAAAAAA ATAACTTTGT	3840
5	TAACCATATT GTAACATTAT TTCATATATT TTGGGGCATG AGAATGATTC TCACGCCCAG	3900
	TAATTTATTT ATGCAATTGT TCATGTAGGT TCTTTGCGAC GTTTTTCAGGA ATACCTATAT	3960
	TTTTAAATC TTCAAGTGTA GCTTCCTTCA TTTTCTTGAT TGAACCGAAT GAACGCAATA	4020
10	ATAATGTTTT ACGTTTGTGA CCGATACCAT CTATATCATC AAGTATTGAT TTCAAGCCTG	4080
	TCTTTTGACG TGTTTGTCTA TGAAATGTGA TTGCGAATCT GTGAACCTCA TCTTGGATAC	4140
	GGTGCAACAA ATAAAATGCC TGGCTATTTT TCTTCAGTGG TACAATTTCT GCACTAGCGC	4200
15	CATATAATAA TTCAGATGTT TGGTGTTTAT CATTTTTCTG CAAACCTGCA ACAGGGATAT	4260
	CAAGACCTAA TTCGTTTTGT AGCACATCAA TAACCCCGTT CATATGTCCT TTACCACCAT	4320
20	CTACTATTAT TAAATCAGGT AATGGTAATC CTTCGTTTAA AACGCGAGAA TATCGTCGTC	4380
	TTACTACTTC TCTCATTGAT TTGTAATCAT CTGGACCTTT AACCGTTTTG ATTTTATACT	4440
	TTCTATAATT TTTCTTATCT GGTTTACCGT CGACAAATGT AACCAATGCT GACACTGGAT	4500
25	CCACACCTTG AATATTAGAA TTATCGAATG CTTCAATTCT AATTGGTGTT TGAATTCCCA	4560
	TTTGTGTTCC AAGTTCTTCA ATAGCTTTAA TCGTTCTGGA CTCATCACGT GATATTAATT	4620
	CAAATTTATT ATTTAAGGAT ACTTTAGCGT TATGTGCAGC TAGGTCAACC ATATCTTTTT	4680
30	TGGGACCTCG CGCGGGTTGA ACGATTTTAG TGTCCACAAC AGATTGAATC ATTTCTTTAT	4740
	CCAAATTACG TGGTACATGA ACTTCCTTAG GTAAATATG TTGGTTTAAAG CTATAAAATT	4800
	GTCCAATAAA TGTATAAAAT TCTTCTTCTT CTGTTTGCTG TAATGGAATC ATCGTTGTAT	4860
35	CTCGCTTTAT CATATTACCT TGTCGTATAA AGAAACTTG GATACACATC CATCCTTTAT	4920
	CAACACTATA ACCAAAGACA TCACGAATCG TTTTATCTGA TGACATAATT TTTTGTTTGT	4980
	TTGTGAGATT TTGAATATGT TGAATTAAAT CTCTATATTC TTTAGCCCGT TCAAAATCAA	5040
40	GTGATTCACT TGCAGTTAAC ATTCGCTCTT CTAAACTTTT TAAAATTGTT TTGTCTTCCC	5100
	CATTCAGAAA ATCAGTAATT TCCTTCGTCA TTTGTGCGTA TTTACTCAA TCAACGTCAT	5160
	ATACACATGG TCCTAAACAT TGTCCAATAT GGTAATAAAG ACATAATTTA TCTGGCATCT	5220
45	TATCACATTT GCGATATGGA TATATTCTGT CTAATAACTT TTTAGTTTCT TGAGCAGAAT	5280
	ATGCATTCCG ATACGGTCCG AAATATTTGC CAGTACCTTG TTTTACAGTT CTCGTCACTA	5340
50	GTAGTCTAGG ATATTTCTCC TTCGTAATTT TAATAAATGG ATAACTTTAA TCATCCTTTA	5400
	ATAATATATT ATATCTTGGT TGATATTGTT TAATCAGATT CAATTCCAGT AAAAGTGATT	5460

55

TTTTAGCATC ATGAGCACCC GTAAAATATG ATCGCAATC

5559

(2) INFORMATION FOR SEQ ID NO: 200:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4594 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

15	AAATCAATCG AGTGGCATGT CAAGGTCATA TCAATATTTT AGAATCTGCG ACTATGAGAG	60
	AGGAAATAAA TGAAATTGCG CGACGTATCA TCGTTGATAT TCGTGATAAG CAATTACGAT	120
	ATCAAGATAT TGCTATTTTA TATCGTGATG AATCTTATGC TTATTTATTT GATTCCATAT	180
20	TACCGCTTTA TAATATTCCT TATAATATTG ATACAAAGCG TTCGATGACA CATCATCCGG	240
	TCATGGAAAT GATTTCGTTCA TTGATTGAAG TTATTCAATC TAATTGGCAA GTGAATCCAA	300
	TGCTACGCTT ATTGAAGACT GATGTGTTAA CGGCATCATA TCTAAAAAGT GCATACCTAG	360
25	TTGATTTACT TGAAAATTTT GTACTTGAAC GTGGTATATA CGGTAAACGT TGGTTAGATG	420
	ATGAGCTATT TAATGTCGAA CATTTTAGCA AAATGGGGCG TAAAGCGCAT AAAGTACCG	480
	AAGATGAACG TAACACATTT GAACAAGTCG TTAAGTTAAA GAAAGATGTC ATTGATAAAA	540
30	TTTTACATTT TGAAAAGCAA ATGTCACAAG CGGAAACTGT AAAAGATTTT GCAACTGCTT	600
	TTTATGAAAG TATGGAATAT TTCGAACTGC CAAATCAATT GATGACAGAG CGAGATGAAC	660
	TTGATTTAAA TGTAATCAT GAAAAGGCGG AGGAAATTGA TCAAATATGG AATGGCTTAA	720
35	TTCAAATCCT TGATGACTTA GTTCTAGTAT TTGGAGATGA ACCAATGTCG ATGGAACGTT	780
	TCTTAGAAGT ATTTGATATT GGTTTAGAAC AATTAGAATT TGTTATGATT CCGCAAACAT	840
	TGGACCAAGT AAGTATTGGT ACGATGGATT TGGCTAAAGT CGATAATAAG CAACATGTTT	900
40	ACTTAGTAGG TATGAATGAT GGAACGATGC CACAACCAAGT AATGCGTCAA GCTTGATTAC	960
	AGATGAAGAA AAGAAATACT TTGAACAGCA GGCTAATGTC GAGTTAAGTC CAACATCAGA	1020
45	TATTTTACAG ATGGATGAAG CATTGTGTTG TTATGTTGCT ATGACTAGAG CTAAGGGAGA	1080
	TGTTACATTT TCTTACAGTC TAATGGGATC AAGTGGTGAT GATAAGGAGA TCAGCCCATT	1140
	TTTAAATCAA ATTCAATCAT TGTTCACCA ATTGGAAATT ACTAACATTC CTCAATACCA	1200
50	TGAAGTTAAC CCATTGTCAC TAATGCAACA TGCTAAGCAA ACCAAAATTA CATTATTTGA	1260
	AGCATTGCGT GCTTGGTTAT ATGATGAAAT TGTGGCTGAT AGTTGGTTAG ATGCTTATCA	1320

55

EP 0 786 519 A2

	GTTTGACAAT	GAAACTGTAA	AATTAGGTGA	AACGTTGTCT	AAAGATTTAT	ATGGTAAGGA	1440
	AATCAATGCC	AGTGTATCCC	GTTTTGAAGG	TTATCAACAA	TGCCCATTTA	AACACTATGC	1500
5	GTCACATGGT	CTGAAACTAA	ATGAGCGAAC	GAAGTATGAA	CTTCAAAAC	TTGATTTAGG	1560
	TGATATTTTC	CATTCTGTTT	TAAAATATAT	ATCTGAACGT	ATTAATGGCG	ATTTTAAACA	1620
	ATTAGACCTG	AAAAAAATAA	GACAATTAAC	GAATGAAGCA	TTGGAAGAAA	TTTTACCTAA	1680
10	AGTTCAGTTT	AATTTATTAA	ATTCTTCAGC	TTACTATCGT	TATTTATCAA	GACGCATTGG	1740
	CGCTATTGTA	GAAACAACAC	TAAGCGCATT	AAAATATCAA	GGCACGTATT	CAAAGTTTAT	1800
	GCCAAAACAT	TTTGAGACAA	GTTTTAGAAG	GAAACCAAGA	ACAAATGACG	AATTAATTGC	1860
15	ACAAACATTA	ACGACAAC	AAGGTATTCC	AATTAATATT	AGAGGGCAAA	TTGACCGTAT	1920
	CGATACGTAT	ACAAAGAATG	ATACAAGTTT	TGTTAATATC	ATTGACTATA	AATCCTCTGA	1980
20	AGGTAGTGCG	ACACTTGATT	TAACGAAAGT	ATATTATGGT	ATGCAAATGC	AAATGATGAC	2040
	ATACATGGAT	ATCGTTTTAC	AAAATAAACA	ACGCCCTTGA	TTAACAGATA	TTGTGAACCA	2100
	GGTGGATTAT	TATACTTCCA	TGTACATGAA	CCTAGAATTA	AATTTAAATC	ATGGTCTGAT	2160
25	ATTGATGAAG	ATAAACTAGA	ACAAGATTTA	ATTAAAAAGT	TTAAGTTGAG	TGGTTTAGTT	2220
	AATGCAGACC	AACTGTTAT	TGATGCATTG	GATATTCGTT	TAGAACCTAA	ATTCACCTCA	2280
	GATATTGTAC	CAGTTGGTTT	GAATAAAGAT	GGCTCTTTGA	GTAAACGAGG	CAGCCAAGTG	2340
30	GCAGATGAAG	CAACGATTTA	TAAATTCATC	CAACATAACA	AAGAGAATTT	TATAGAAACA	2400
	GCTTCAAATA	TTATGGATGG	ACATACTGAA	GTTGCACCAT	TAAAGTACAA	ACAAAAATTG	2460
	CCATGTGCTT	TTGTAGTTA	TCAATCGGTA	TGTCATGTAG	ATGGCATGAT	TGATAGTAAG	2520
35	CGATATCGAA	CTGTAGATGA	AACAATAAAT	CCAATTGAAG	CAATTCAAAA	TATTAACATT	2580
	AATGATGAAT	TTGGGGGTGA	GCAATAGATG	ACAATTCCAG	AGAAACCACA	AGGCGTGATT	2640
	TGGACTGACG	CGCAATGGCA	AAGTATTTAC	GCAACTGGAC	AAGATGTACT	TGTTGCAGCC	2700
40	GCGGCAGGTT	CAGGTAAAC	AGCTGTACTA	GTTGAGCGTA	TTATCCAAAA	GATTTTACGT	2760
	GATGGCATTG	ATGTCGATCG	ACTTTTAGTC	GTAACGTTTA	CAAACCTAAG	CGCACGTGAA	2820
45	ATGAAGCATC	GTGTAGACCA	ACGTATTCAA	GAGGCATCGA	TTGCTGATCC	TGCAAATGCA	2880
	CACTTGAAAA	ACCAACGCAT	CAAAATTCAT	CAAGCACAAA	TATCTACACT	CCATAGTTTT	2940
	TGCTTGAAAT	TAATTCAACA	GCATTATGAT	GTATTAAATA	TTGACCCGAA	CTTTAGAACA	3000
50	AGCAGTGAAG	CTGAAAATAT	TTTATTATTA	GAACAAACGA	TAGATGAGGT	CATAGAACAA	3060
	CATTACGATA	TCCTTGATCC	TGCTTTTATT	GAATTAACAG	AGCAATTGTC	TTCAGATAGA	3120

55

AATCCTACAA ATTGGTTGGA TCAATTGGTG ACACCATACG AAGAAGAAGC ACAACAAGCG 3240
 CAACTTATTC AACTACTAAC AGACTTATCT AAAGTATTTA TCACAGCTGC TTATGATGCT 3300
 5 TTAATAAGG CGTATGATTT GTTAGTATG ATGGATAGCG TCGATAAACA TTTAGCTGTT 3360
 ATAGAAGATG AACGACGTTT AATGGGGCGT GTTTTAGAAG GTGGCTTTAT TGATATACCT 3420
 TATTTAACTG GTCACGAATT TGGCGCGCGT TTGCCTAATG TAACAGCGAA AATTAAAGAA 3480
 10 GCAAATGAAA TGATGGTCGA TGCCTTAGAA GATGCTAAAC TTCAGTATAA AAAATATAAA 3540
 TCATTAATTG ATAAAGTGAA GAGTGATTAC TTTTCAAGAG AAGCTGATGA TTTGAAAGCT 3600
 GATATGCAAC AATTGGCGCC ACGAGTAAAG TACCTTGCGC GTATTGTGAA AGATGTTATG 3660
 15 TCAGAATTCA ATCGAAAAA GCGTAGCAAA AATATTTTGG ATTTTCTGA TTATGAACAT 3720
 TTTGCATTAC AAATTTTAAC TAATGAGGAT GGTTCGCCTT CAGAAATTGC CGAATCATAC 3780
 CGTCAACACT TCCAAGAAAT ATTGGTCGAT GAGTATCAAG ATACGAACCG AGTTCAGAG 3840
 20 AAAATACTAT CTTGCATCAA AACGGGTGAT GAACATAATG GTAATTTATT TATGGTTGGA 3900
 GATGTTAAGC AATCCATTTA TAAATTTAGA CAAGCTGATC CAAGTTTATT TATTGAAAAG 3960
 25 TATCAACGCT TTACTATAGA TGGAGATGGC ACTGGACGTC GAATTGATTT GTCGCAAAAC 4020
 TTCCGTTCTC GAAAAGAAGT ACTGTCAACG ACTAACTATA TATTCAAACA TATGATGGAT 4080
 GAACAAGTCG GTGAAGTAAA ATATGATGAA GCGGCACAGT TGTATTATGG TGCACCATAT 4140
 30 GATGAATCGG ACCATCCaGT AAACTTAAAA GTCCTTGTTG AAGCGGATCA AGAACATAGT 4200
 GATTTAACTG GTAGTGAACA AGAAGCGCAT TTTATAGTAG AACAAGTTAA AGATATCTTA 4260
 GAACATCAAA AAGTTTATGA TATGAAAACA GGAAGCTATA GAAGTGGAC ATACAAGGAT 4320
 35 ATCGTTATTC TAGAACGCAG CTTTGGACAA GCTCGCAATT TACAACAAGC CTTTAAAAAT 4380
 GAAGATATTC CATTCCATGT GAATAGTCGT GAAGGTTACT TTGAACAAAC AGAAGTCCGC 4440
 TTAGTATTAT CATTTTTAAG AGCGATAGAT AATCCATTAC AAGATATTIA TTTAGTTGGG 4500
 40 TTAATGCGCT CCGTTATATA TCAGTTCAAA GAAGACGAAT TAGCTCAAAT TAGAATATTG 4560
 AGTCCAAATG ATGACTACTT CTATCAATCG ATTG 4594

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6313 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

GGTTTTChTG GAAAGATAGT GAAAATCTCG TGTTTTTTGG TTTTgAGGTG TTGTTTGTAT 60
TTTaTAAAT GGCTTACATA TATGAAGCGT TGATTAAGTA TGAATTGTT AATTAATTGA 120
5 ACCTATTTAG CTTTAAGAAG GCATAACAAG ATGACCTTAT TTTATGCTAT AATATTTCTA 180
TTATGCCGAAG ATTAAGGTGA GTAGTAAATT GGATAAAAAA GTAAGTATTC AAACAAAGCA 240
AGTGTGAAA CAGCACACG AAAAAGAAAA ATTTGAATTT ACTACTGAAG GAACTTGGCA 300
10 ACAAGGCAA TCTAACTTTA TTCGGTATGT AGAACAAATT GAGGATGCAA CAGTTAATGT 360
TACAATAAAA GTGGATGATG ATAGCGTTAA GTTGATTCTG AAAGGCGACA TTAATATGAA 420
TTTGCATTTT GTTGAAGGAC AAACGACAAC AACTTTTTTAC GATATATCGG CTGGACGAAT 480
15 TCCACTAGAA GTTAAACAT TACGCATTTT ACATTTCTGTA AGTGGAGACG GTGGCAAGCT 540
AAAGATTCAT TATGAATTAT ATCAAGATAA TGAAAAATG GGTCTTTATC AATATGAAAT 600
20 TAACTATAAG GAGATAGGCG AATGAATATT ATTGATCAAG TGAAACAAAC ATTAGTAGAA 660
GAAATTGCAG CAAGTATTAA CAAAGCAGGA TTAGCAGATG AGATTCCTGA TATTAAATTT 720
GAAGTCTTA AAGATACAAA AAATGGAGAT TATGCTACTA ATATTGCGAT GGTACTGACT 780
25 AAGATTGCAA AGCGTAATCC TCGTGAAATT GCTCAAGCGA TTGTTGATAA CTTAGATACT 840
GAAAAAGCAC ATGTAAACA AATTGACATT GCTGGTCCAG GATTCATTAA TTTTACTTA 900
GATAATCAGT ATTTACAGC AATTATTCCT GAAGCAATTG AAAAAGGTGA TCAATTTGGA 960
30 CATGTAAATG AATCAAAAGG TCAAAATGTA TTGCTTGAGT ATGTTTCAGC TAACCTACA 1020
GGAGATTTAC ATATTGGTCA TGCTAGAAAT GCAGCAGTTG GTGATGCTTT AgcTAAaATT 1080
TTAACTGCAG CTGGCTATAA TGTAACACGT GAATATTATA TTAATGATGC TGGTAATCAA 1140
35 ATTACTAACT TAGCGCGTTC GATTGAAACA CGTTTCTTTG AAGCTTTAGG TGACAATAGT 1200
TATTCAATGC CAGAAGATGG CTATAATGGA AAAGATATTA TTGAAATAGG TAAAGATTTA 1260
GCAGAGAAAC ACCCTGAAAT TAAAGATTAT TCTGAAGAAG CACGTTTGAA AGAATTTAGA 1320
40 AAATTAGGCG TAGAATACGA AATGGCTAAA TTGAAAAATG ATTTAGCAGA GTTCAATACG 1380
CATTTTGATA ATTGGTTTAG TGAAaCATCT TTATATGAAA AAGGAGAAAT TCTTGAAGTT 1440
45 TTAGCAAAAA TGAAAGAATT AGGTTATACG TATGAAGCTG ATGGCGCTAC ATGGTTACGT 1500
ACAACTGATT TTAAAGACGA CAAAGACAGA GTATTAATTA AAAATGACGG TACATATACG 1560
TATTTCTTAC CAGATATTGC GTACCACTTC GATAAAGTAA AACGTGGTAA TGACATTTTA 1620
50 ATCGATTTAT TTGGTGCTGA TCATCATGGT TATATTAATC GTTTGAAAGC ATCTCTTGAA 1680
ACGTTTGGTG TAGATAGTAA TCGTTTAGAA ATTCAAATCA TGCAAATGGT TCGTTTAATG 1740

55

	ATTATGGACG AaGTTGGCGT TGACGCTGCA CGTTATTTCT TAACTATGCG TAGTCCTGAT	1860
	AGTCACTTTG ATTTTGATAT GGAATTAGCG AAAGAGCAAT CTCAAGACAA TCCAGTTTAC	1920
5	TATGCTCAAT ATGCACATGC GCGTATTTGT TCAATTTTAA AACAAGCGAA AGAGCAAGGT	1980
	ATTGAAGTGA CTGCTGCGAA TGATTTTACA ACGATTACTA ATGAAAAAGC GATTGAATTG	2040
	TTGAAAAAAG TAGCTGATTT CGAACCTACA ATTGAAAGTG CTGCTGAGCA TAGATCGGCA	2100
10	CATAGAATTA CTAATTATAT TCAAGATTTA GCTTCTCATT TCCATAAATT CTATAATGCT	2160
	GAAAAAGTGT TAACAGATGA TATTGAAAAA ACAAAGCAC ATGTTGCTAT GATTGAAGCG	2220
	GTCAGAATTA CATTGAAAAA TGCATTGGCA ATGGTCGGTG TAAGCGCACC TGAATCAATG	2280
15	TAAGAACATT TATATACACT CCAACGTAGA GTTTCTCGAA AGATACTTTG TGTGAGTG	2340
	TTTTTTTTAG GTATGTGACA TATTGGGGAA TGCTTAGTAT GTGAATAAGG TTAAGAGGAA	2400
	CACAGTTGGA TGCTCTGCAC AACTGCATAA GAGAGCCTGA GACATAAATC AATGTTCTAT	2460
20	GCTCTACAAA GTTATAATGG CAGTAGTTGA CTGAACGAAA ATTCGCTTGT AACAGCTTT	2520
	TTTCAATTCT AGTCAACCTT GCCGGCGGGG CCCCACAAA GAGAAATTGG ATTCCCAATT	2580
25	TCTACAGACA ATGCAAGTTG GGGTGGGACG ACGAAATAAA TTTTACGATA ATATCATTTT	2640
	TGTCCCACTC CCTCTAAAAT GGAGGGTGTA AATGTTAGGA ACTGATGAAT TATATAAAGT	2700
	TTTATATGAA CATCTCGGAC CACAATTTTG GTGGCCTGCT GATAATGACA TTGAAATGAT	2760
30	GTTAGGTGCA ATTTTAGTTC AAAATACTAG ATGGCGAAAT GCAGAAATTG CATTGAATCA	2820
	GATTAAAGAA CATACGCATT TTAATCCAAA TCATATATTA GAACTACCTA TTGAAACGTT	2880
	ACAATCATTG ATACATTCAA GTGGCTTTTA TAAAAGTAAA TCACTGACGA TTAAAACATT	2940
35	ATTAACATGG TTAGCACGAC ATCATTTCOA TTATCAAGAG ATTAATGAGC GATATAAAGG	3000
	TGGATTAAGA AAAGAATTAT TATCTTTGAA AGGTATTGGA AGTGAAACAG CAGATGTCTT	3060
	ACTTGTTTAT ATATTCGGAC GTATTGAATT TATTCCAGAT AGCTATACAA GAAAAATATA	3120
40	TGATAAATTA GGATATGAAA AACTAAAAA TTATGATCAA TTAAAAAAG TAGTCAATT	3180
	ACCAAATCAT TTTACAAATC AAGATGCTAA TGAATTCAT GCTCTGTTAG ATGTATTTGG	3240
45	TAAACATTAC TTTAGAGACA AAGATATAAA GAATTATGAT TTTTLAGAAC CTTACTTTAA	3300
	AAAGTAAACG CTGTGAAGTT AGATAGATGA GTTTATATGA AATATAAAAA ATAATTTACT	3360
	ATTTTCTTTT AGTATGTGGA CTTATATAAT AAATAGAAGC ATATAAAGAA AAAACAGTT	3420
50	GTTTGTGTTT GCAGCAACTG CATAAGAGCC CCTAATCGCT AAAGCTCAAG GGGAGTAAAG	3480
	GAATACAGTT GTTTGTGCAG CAACTGCATA AAAGCCTCTA ATCACTAAAG GTGAAGAGGA	3540

55

	AACGCAGTTG GATgCTACCG CACAACCTGCA TAAATCCCTC TaATCgcTAA AGCGAAAAGT	3660
	GGGATTAAAA AGGAGATGTG ATAGTGTGAA GAAATCGTTA ATTGCTTTTA TTTTGATTTT	3720
5	TATGCTTGTC CTGAGTGGCT GTGGTATGAA AGATAATGAT AAACAAGGTA GCAATGATAA	3780
	TGGCTCGTCT AAATCGCCGT ACCATAGAAT TGTTCGTTA ATGCCTAGTA ATACTGAAAT	3840
	TTTATATGAA TTAGGATTAG GTAAATACAT AGTTGGTGTT TCAACGGTTG ATGATTATCC	3900
10	AAAAGATGTG AAAAAGGGTA AGAAACAATT TGATGCTTTG AATCTAAATA AAGAGGAACT	3960
	TTTAAAGGCA AAGCCAGATC TAATTCTTGC GCATGAGTCG CAAAAGGCAA CTGCTAATAA	4020
	AGTATTGTCA TCATTAGAGA AACAAGGCAT CAAAGTAGTG TATGTTAAAG ATGCACAATC	4080
15	AATTGATGAA ACTTACAACA CATTTAAGCA AATTGGGAAA TTAACGCATC ATGATAAGCA	4140
	GGCTGAACAA CTTGTTGAGG AAACATAAGA TAATATCGAT AAAGTCATAG ATTCAATTCC	4200
20	TGCTCATCAT AAAAAATCAA AAGTATTTAT TGAGGTTTCA TCAAAGCCTG AATATATAC	4260
	AGCAGGGAAG CATACTTTT TTAATGATAT GTTAGAAAAA TTAGAAGCCC AAAATGTGTA	4320
	TAGTGACATT AATGGTTGGA ACCCTGTAAC GAAGGAAAGT ATTATTAAAA AGAACCCAGA	4380
25	TATATTAATT TCGACGGAAG CTAAGACAAG ATCAGATTAT ATGGATATCA TCAAAAAAAG	4440
	AGGTGGATTG AATAAAATTA ATGCTGTCAA GAATACACGT ATTGAAGTTG TAAATGGTGA	4500
	TGAAGTATCA AGACCAGGTC CACGTATTGA TGAAGGATTA AAAGAATTAA GAGATGCAAT	4560
30	TTATAGAAAA TAAACCATTC TAATTATGCC CCTTATTGCT ACATGTAAAA AATACATGTT	4620
	TGAGATAAGG GGTTTTTaAA ATATATTTAG TGAATGATAG CAACGCGAGT ATGTGATTGC	4680
	TATAATGAAT GTAATTATCG ATGAACaAAA GAGAATGCTA TGACATTTAA TAAAGTATTA	4740
35	TTGAGCTGGa TAGTCmTATT GATTATAACA ACTAGCATAT ATCTATTTTG GCAGTTGGGC	4800
	GATATCAATG ATGTATTTAA CCAGTCTATT TTAATCAATG TTAGATTACC GAGATTATTA	4860
	GAAGCATTGT TGACAGGTAT GATATTAAC TTTGACAGCC TTATATTTCA AACAGTTTTA	4920
40	AATAATGCAT TGGCAGATAG CTTTACATTA GGATTGGCAA GCGGCGCTAC ATTTGGTTCA	4980
	GGATTAGCAT TATTTTtagg TTTAACAACG TTATGGATTG CTGTATTTTC AATAACATTT	5040
45	AGTTTGATAA CATTAAATAAC TGTATTAGTC ATTACGTCGG TATTGAGCCA AGGCTATCCA	5100
	GTTAGAACTT TAATATTAAG TGGTTTAATG ATTGGTGCCT TATTCAATTC ACTTCTATAT	5160
	TTTTTGATTT TATTAAAACC TCGCAAATTA AATACAATTG CCAATTATCT GTTTGGTGGT	5220
50	TTTGGTGATG CAGAATACTC AAATGTATCT ATAATAGCAA TCACATTTAT CATTGCATTG	5280
	TTTGGTATAT TTATCATTCT TAATCAACTA AAGTTATTGC AATTAGGAGA ACTAAAAAGT	5340

55

5 ATAACGGCGA TAAATGTGCG ATATGTTGGC ATCATTGGAT TCATTGGTAT GGTGATACCG 5460
 CAACTCATT A GAAAATGGCA GTGGAAACAA TCATTAGGAA GACAATTGGC TTTAAATATT 5520
 10 GTAAGTGGAG GACAAATAAT GGTTATGGCA GATTTTATTG GTAGCCATAT ATTGTCACCA 5580
 GTACAAATAC CGGCAAGTAT TATCATTGCA TTAATTGGTA TACCAGTGTT AtTTTACaTG 5640
 CkAAwatCtC aGTCgAAAcG GTTACaCTAG CACACGACaT TTGCTAAaAT AAAAATAACT 5700
 15 ATAAACATAA AGAGGGCATA AGCGATGGAT TTGAATCAAA TTAAAGCAGT TGTATTTGAT 5760
 TTAGAAGGTA CGTTGTTGGA CAGAGTTAAA TCTCGAGAGA AATTATCGA AGAGCAATAT 5820
 GAACGATTTC ATGACTACTT AATTCATGTT CAACTGGCAG ATTTTAAAAA AgCATTATT 5880
 20 GAGCTAGATG ACGATGAAGA TAATGATAAA CCTGATTTAT ATAAAGAAAT CATTAAACGT 5940
 TTCCATGTAG ATAGGTTAAC TTGGAAAGAC TTATTTAATG ATTTTGAAAT GCATTTTTAT 6000
 CGTTATGTAT TTCCTTATTA CGATACTTTG TATACACTAG AAAAgCTATC GCAAAAAGGC 6060
 25 TTTCAAATTG GTGTTATCGC AAATGGTAAA TCTAAGATTA AACAAATTCG ATTACATTCA 6120
 CTTGGTTTGA TGCATGTTAT TAATTATTTA TCAACATCAG AACAGTTGG TTTTCGTAAA 6180
 CCACATCCTA AAATTTTGA AGATATGATT GATCAACTAG GGGTATTACC TGAGCAAATT 6240
 30 ATGTATGTTG GCGATGATGC GTTAAATGAT GTAGCTCCAG CACGAGCTAT GGGCATGGTT 6300
 AGTGTATGGT ATA 6313

(2) INFORMATION FOR SEQ ID NO: 202:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2174 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

40 CCGTAAACAC ATCAACAAAA GAAGGCTATA TTACAAAAGA AGACTTGGAC TTATGCTGCA 60
 CGTCGCTCTA ATTCAGCTGG AATGCAAGTC ACCGGACGAC TGGCTTACAT TGAACCTTAT 120
 45 GGGGCAACAA GTCGCACAAA ATAAACGCGC GAGAAGCaAG AATAGGAAGT GATATCTATG 180
 AAATGGTTAT CACGAATATT AACAGTAATA GTGACCATGT CtATGGcGTG TGGTGcATTG 240
 ATATTTAATC GTAGACATCA GCTAAAGGCG AAAACGCTGA ACTTCAATCA TAAAGCATT A 300
 50 ACAATTATTA TTCCGGCTAG AAACGAAGAA AAAAGAATAG GTCATTTACT ACATTCGATA 360
 ATACAACAGC AAGTTCCAGT AGATGTCATT GTTATGAATG ACGGATCGAC AGATGAAACA 420

	AAATGGTATG GGAAATCACA TGCTTGTTAT CAAGGTGTGA CGCATGCATG TACGAATCGC	540
	ATTGCCTTTG TAGATGCTGA TGTAACCTTC TTAAGGAAAG ATGCTGTTGA AACGTTGATT	600
5	AATCAGTATC AATTACAAGG TGA AAAAGGA TTGTTAAGCG TACAGCCTTA TCATATAACA	660
	AAGCGTTTCT ACGAAGGGTT TTCAGCGATA TTTAATTTAA TGACAGTCGT TGGTATGAAT	720
	GTATTTTCTA CCTTAGACGA CGGTCGGACT AACCAGCATG CATTGGACC GGTGACATTA	780
10	ACAAATAAAG AAGATTATTA TGCAACTGGA GGTCAATAAA GTGCAAACCG TCATATTATT	840
	GAAGGATTTG CTTTAGGAAG TGCATATACT TCACAATCAT TGCCCGTAAC AGTTTATGAA	900
	GGGTTTCCAT TTGTTGCATT TCGCATGTAT CAAGAAGGAT TTCAGTCATT ACAAGAAGGA	960
15	TGGACAAAGC ATTTGTCAAC TGGGGCAGGT GGCACAAAGC CTAAGATCAT GACAGCAATT	1020
	GTGTTGTGGT TGTGTTGGTC TATAGCGAGT ATTTTAGGGC TATGTCTTAG TTTAAATAT	1080
20	CGCCAAATGT CTGTAAGAAA AATGGTAGCA CTTTACTTGA GCTATACTAC ACAATTTATT	1140
	TATCTGCATC GAAGGGTCGG CCAATTTTCT AATTTATTAA TGGTATGTCA TCCATTGTTA	1200
	TTTATGTTTT TTAATAAAAT TTTTCATCAA TCTTGAAAC AAACGCATCG TTATGGTGTA	1260
25	GTTGAATGGA AAGGTCGTCA ATATTCTATA TCTAAAGAAC AATAAATCAA GGTAAATGGCA	1320
	TTTCAATATA GGAGGACTAG TATGACAATG ATGGATATGA ATTTTAAATA TTGTCATAAA	1380
	ATCATGAAGA AACATTCAA AAGCTTTTCT TACGCTTTTG ACTTGTTACC AGAAGATCAA	1440
30	AGAAAAGCGG TTTGGGCAAT TTATGCTGTG TGTCGTAAAA TTGATGACAG TATAGATGTT	1500
	TATGGCGATA TTCAATTTTT AAATCAAATA AAAGAAGATA TACAATCTAT TGAAAAATAC	1560
	CCATATGAAC ATCATCACTT TCAAAGTGAT CGTAGAATCA TGATGGCGCT TCAGCATGTT	1620
35	GCACAACATA AAAATATCGC CTTTCAATCT TTTTATAATC TCATTGATAC TGTATATAAA	1680
	GATCAACATT TTACAATGTT TGAAACGGAC GCTGAATTAT TCGGATATTG TTATGGTGTT	1740
	GCTGGTACAG TAGGTGAAGT ATTGACGCCG ATTTTAAGTG ATCATGAAAC ACATCAGACA	1800
40	TACGATGTG CAAGAAGACT TGGTGAATCG TTGCAATTGA TTAATATATT AAGAGATGTC	1860
	GGTGAAGATT TTGACAATGA ACGGATATAT TTTAGTAAGC AACGATTAAA GCAATATGAA	1920
45	GTGATATTG CTGAAGTGTA CCAAATGGT GTTAATAATC ATTATATTGA CTTATGGGAA	1980
	TATTATGCAG CTATCGCAGA AAAAGATTTT CAAGATGTTA TGGATCAAAT CAAAGTATTT	2040
	AGTATTGAAG CACAACCAAT CATAGAATTA GCAGCACGTA TATATATTGA AATACTGGAC	2100
50	GAaGTGAGaC AGGCTAACTA TACATTACAT GAACGTGTTT TTGTGGaTAA GAGGAAAAAG	2160
	GCAAAGTTGT TTCA	2174

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4715 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

10	GAAnCAGnTA GACAAATTAT GGaAAmCGGT GTGAATCaAG GATTCTTTGG TGTAGCTGGT	60
	TTTGACCTAC TCGTCGATGA GGATGATAAC GTTTATGCCA TTGATTTAAA CTTTAGACAA	120
	AATGGTTCaA CGAGCATGTT ATTACTTGCT AACGAGTTGA ATTCAGGATA TCAAAAGTTT	180
15	TATAGTTATC ATTCAAAAGG TGATAACACA CATTTCTTCA ATACGATTTT GAAATATGTC	240
	AAAGAAGGTA GTTTATACCC GTTATCTTAT TATGATGGTG ATTGGTACGG TGAAGATAAA	300
	GTTAAATCAA GGTTCGGCTG TATTTGGCAT GGTGATTCAA AAGAAACAGT ACTGGAGAAT	360
20	GAACGCGCAT TTTTAGCTGA ACTTGAACAC TATTAGAGTT CGGAACATAA GGCGCTACAA	420
	TGTTGTGTTG CCAGTAGTTG ACTGAATATG CGTTTGTAAC AAGCTTTTTT CGATTCTAGT	480
	CAACAGTAAT TAAATTTATG ATATGGCAAT ACTTTGTAAT ACTAATATTA AATGGCGACT	540
	TTTATTTTAC TATGTTATAA GAGTTGCCAT TTTGTTGATA AAGGTATACT AAAGGTTATC	600
	GTTTTGAAAT TTTTAGTAAC TAGATATGTT TCGTGTTATA GACCGAATTT GTGTATACGT	660
30	AAAATTTAAT GCTATTGAAT TTTTAAATG AAAAACATGA CATTAATTG AATTCATAAT	720
	ATGTCTAATT GACTAACTTG TTGGAGTCAT TTAATTTTTT ATGTATGACA TATTTTAAAA	780
	AGTGAGGGTC AAGCATGTCT TATAAAGCAT ATCCATTCTT TAGAGATATA TTAATAAATG	840
35	AATGTATTTA TTTCGCCTCT AAAAATAAAA AACTAGTACG CCTAAATTAT AAAAGTGAAG	900
	CGnATGTAGG CGTTTGGACA GAAGAAAGTG TGGCCGTATC ATTTTAAACA AGTCGTGATA	960
	TTCCATTTGA TAAAGTTGTA AAAATGGACG TTGATCGTTT TGCTACTTAT GAATTAGATG	1020
40	AATTGTTTGA TGAACAAGAC CATATTATTA TGAATCAAAC AATGGAAGAw GAAGGGCATC	1080
	TACTAAACGT TGTAGCTGTT ACACAAGAAG TGATGACGGA ATTAGATAAA ATTAGAATCA	1140
	AAGAATTTGT CCAAGATGTA GCGAAATATG ATGAAGTATA CGGCTTAACT AAAAAAGGTA	1200
45	GTAAGCAGTT TATTCTCATT AGTGAAAATG ATAGCGACGA AAAAAAGCCG CATATTATGC	1260
	CTGTATGGAG TATTA AAAAC AGAGCGTTAA AAGTTGAGA TGAAGATTTT GAAGAGTGTG	1320
	ATTTAATTAC GATTGAAGGT TCTGTTTTCG GAGAATGGCT AGATGAACTT AGAGATGATC	1380
50	ATAAAGCCGT TGCGATAGAT TTA AAAACTG GCGTGGTTGG TACAATTGTT TCAGCGCAAA	1440

	ATGGAACAAT ACGTATTCAA AACACTTAGA CCATAAAATA AAAGGCCATT TATATAGCGT	1560
	TTATTTAAAA CAACGCGCAT ATAAATGGTC TTTTCTATT TTTCTAAATA TAATGCACCA	1620
5	ATAGCACCTG AAAAATGCGC CGTTTTCAAC ATAGTACGGT TTGCAACCGC GTAACACAGT	1680
	ATAATCTTCC ACAACTTTGC GTAATAAAGC GTTATTATGA AATGAAGAAC CGATATAAAC	1740
	GATATTTTCA GTTTTAAATT CACGTGCAAC AGTAATGGCC ATTGTCGTAA CAACTTCGCC	1800
10	AACGACACCA ATAACGGCTG CTAATTTATT GCTAGGTGTA AAATCAGCAT CTAAATGATG	1860
	TAGTACATGA CCAAATTAG CTGCTGTAA ATCACC GGGA ATGGGTGGTT CGGTATCTTT	1920
	ATAAATATGT CTAACCTTTA AATCGATAGT GTTACGATCA CCGTGTGTG CCATGTCAGT	1980
15	TAAGTGTGTA TAATCAGTGA TTTGACTTAG TAAATAACCG AGTCCTTGAA TCATGCCTCC	2040
	ACCTGTACCG ATACCGCCTA CACGACGTTG TGATTGGCCG TCGAAATAAT GTAGTGACGT	2100
20	ACCGGTACCA ACATTTGCAA AAATATAATC TGCTAAGTCA TGGCCTTGCT CTTTAAACAA	2160
	AATACCTAGT CCTTGAGATG CAGCATCAAA CTCTACAAAA ATTTGTGCAG GAATGTTGAT	2220
	GTTTTAGCA ATGACACCTG CATTACCTCC AGTTAAGCAT AATTTTTCAA TTTGCTGTTG	2280
25	GTTTAACCAT TCCACAACCT GATCAATATT TTTAGTTAAT TCAGTTTTAA AAGTACGTTG	2340
	GTTATCTTGC TCTTGAACGA TTTTAATTAG TGTACCGCCA GCGTCAATGC CAACTTTCAT	2400
	AAGATTCCCA CCTCATTATT AATGTCTATC CTAAATAAT AGTATAGTAA AATGACTAAA	2460
30	AAACAAGTAA TAATAGTAAT TATTAACAAA TTTGATGCCA TTGCATTTCA ACATTGTAAG	2520
	CGTATCGCAA TTAAGTTTT ACAACGTGG ACGTTAAGT ATATATATTA TTTCTAGGA	2580
	ATTTTGAAGT TGTATAGGAT TGTTAGTTAG TGACGCAATA TTAAGTAG TTCGTACGCA	2640
35	GTGTATTTGT AAGTCTCTGA TTAAGATGAT AAGTAATGAG GAATAGTACA TTAATTTTGA	2700
	AATTTAAAA ATATAAATA GTAATTTATT TAACTTAGAG CAAATAATGG TATCGTAGTG	2760
	AAATAATAGG TAAATAATA TGGGGATTCA TGCTTCATAT ATAAAAAGAT AGGGGTAA	2820
40	TATATGGCTA AAGAACTTTG TTTGAAGGT ATCACTTTAA AAGCATTGTA TGAACAATAT	2880
	CGTTCAGCAA TTAATGATTT TGAATGAAT GAAAGACAAC AAATATATTC ATCTTTACCT	2940
45	AAAGAAGTTA TTGATGATGC AATTAATGAT GCTGATAGGA TTGCTAACGT AGCAwTAaMc	3000
	GATAAAATG AAGTGGTGGG CTTTTTTGTA TTACATCGTT ACTATCAGCA TGAAGGTTAT	3060
	GATACACCTG AAAATGTCGT TTATATTCGT TCATTATCGA TTAATGAAAA ATATCAAGGT	3120
50	TTTGATATG GCACGAAAAT AATGATGTCA TTGCCGCAAT ATGTTCAAGG TGTATTTCTT	3180
	GATTTTAATC ATCTATATCT AGTAGTAGAT GCGGAAAATG ACAATGCTTG GAACCTATAC	3240

55

CTATATTACT TGGACTTAGA TTCAAAACAT GTTTCATCAT TAAAGCTTGA AGAAGAAAGT 3360
 CGTTCAGAAG TGACCAATGT ACATATCATT AATTTAATGA TTGATGGCCA AAAGGTTGGC 3420
 5 TTTATCGCAT TGGAGCAGAT TGGTGAACGC ATGAACATTG CTGCTATTGA AGTGGATAAA 3480
 TCATATCGCT TTAATGGTAT TGGTTCAAGT GCTCTGCGAC AATTGCCAAC TTAETTAAGA 3540
 AAAAATATG ACAACCTTAA TGTGATTACG ATGATTCTGT TTGGAGAGAA TAATGATTTT 3600
 10 AAACCATTAT GTTTAAATAG TAATTCGTT GAAATCGAAC AAATGATGA TTATGTCGTT 3660
 TTCGAAAAAT ATTTAAATTA CTAACAGTGA TTGCGAAATA TGATATTGTC ATTTATAATT 3720
 TAGTTTTGTT ACTATATATA AATGAATTCA GACGTATAAA TTTAGATTAT ATCCTTCGAA 3780
 15 AGGAAGTATT GGGCAATGAA AATTCAAGAT TATACAAAAC AAATGGTTGA TGAAAAATCA 3840
 TTTATTGATA TGGCTTATAC ATTATTGAAT GATAAAGGCG AAACAATGAm mTTATATGAT 3900
 20 ATYATCGATG AATTTAGAGC GTTAGGTGAT TATGAGTACG AAGAAATTGA AAATCGTGTT 3960
 GTACAATTTT ACACGGATTT AAACACAGAT GGTCGTTTTT TAAATGTTGG AGAAAATTTA 4020
 TGGGGATTAC GTGATTGGTA TTCGTTAGAT GATATTGAAG AGAAAATCGC ACCAACTATT 4080
 25 CAAAAATTCG ATATTCTGGA TGCAGATGAT GAAGAAGATC AAAACTTAAA ATTATTGGGC 4140
 GAAGATGAAA TGGATGACGA CGATGATATT CCAGCTCAA CAGATGATCA AGAAGAACTA 4200
 AATGATCCAG AAGATGAGCA GGTGAAGAA GAAATCAATC ATTCGGATAT AGTCATTGAA 4260
 30 GAAGATGAAG ATGAAGTAGA CGAAGACGAA GAAGTGTTTG AAGACGAAGA AGACTTCAAC 4320
 GATTAATTTT TTGTTTGACT TTTAGTTGAA AGATGATAAA ATTTTATTCTG GGCTCCTTTA 4380
 AATAGGACAC GTGTATAAAA TTTATACGCT CCCCTTACAG AATTTGTGAG AGGGAGCGTT 4440
 35 TTTTtATTTA ATTGAGTAAA TCAAGAAATG ATAACGCAA AATCAAAGTT GTAAATGATA 4500
 TACATAGTGA CATAGCAGTA TGGAAACGGT AAGTAAACAG AATTTAATT TGTCGAtTCG 4560
 ACAAtAAaCA aCTtGAaTGA GCTTGCTTTA ATGTTATGTh nTACGTAATT TTTACAATTG 4620
 40 ATGAGGAAGC ATTCCTTTA ATAATTAGGA GGTCAAGACA TGACAAAATT TATTTTTGTA 4680
 ACAGGTGGCG TAGTTTCATC CATTAGGGGA AGGGT 4715

45 (2) INFORMATION FOR SEQ ID NO: 204:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 918 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATAATAACTG AAATTAAAAT TGCTAAATmG TGTtaAgCTA TCGCmACAAT GAAAATwCCG 60
 ATTTTGC GTTAAAAATA TCTTTCCAAA CCAAGAATCG ATAATGGCAA TAAATATAAT 120
 AAATTTCCAT AAAATGACCA AGTAAAATTA AAGTATATAA CGACAGTTGA CATGCCGTAT 180
 AAAATCGTAG CGATCATATT TGCTGAGCGT TTAAAGTGTA ATATTTTAAA TAAGTAGAAG 240
 GTCACGACAA ATGTTATGAT AGCTCGTATC ATAGCCATAA TAAGTTGGTT TGTCGGCCAA 300
 AAATGTATTG TCGTCGGATT AAATATACCA ACCGTTTCTC CTATTTTAAAT GAAKAGAAAA 360
 TTTAGCCACA TTAAAGGTGA CAGCGAATAA TAATmTGATA GTCCTTTTCAT ATAATCGCCA 420
 CCTAmTCCAA ACGATGCATC ATtTAAACTA GAAnAACTAC GTAGATGTTT ATACAnATAC 480
 ATTTGAAATG GCATCATTG ACGGAATCCA TCTCCAGCCC CGCTAAAAAC AGTACCATTG 540
 ACAATATAAT CATAGATATG AGTAGAAAAT AAAATAAGCG TTAATATTAC ACTAATGAAA 600
 GTTATAACAA AGAATTGTTT GACGTTTGAA TTTAGCCACT TTTTAAACAC AACATTATCC 660
 TCAACTTTCA AATTTAAAAT TAAGTTTAAAC TGAAACTAAA GTTAATGAGG TTCTTGATAG 720
 GTAAAGACGA AGATGACTGT GGAACAGATA CCTTATCATA GTTACTTAAA CTTTGGATCA 780
 TTTTCAGTTT ATCATTAAC AAATATATTG AATAATAAAA aTGTCACTACT GATAAAGATG 840
 AATGTCACCT AATAAGTAAC TTAGaTTTAA CAAATGATGA TTTTAAATTG TAGAAAACTT 900
 GAAATAATCA CkTATACC 918

(2) INFORMATION FOR SEQ ID NO: 205:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16397 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

40
 45
 50
 55

TCGCCChATA ATCAATTAT TTTTCATGTG CCACTCCTAT ACAAGCTnAC AATGCTTCTT 60
 CAGTTAAGGC AATATCTTTT AATTTTGTTT GATATTTTGT TTCAAAGTCA TATTGTAAC 120
 GAACAATTTC TGGCAAACCA ATATGCCAAT CCGCCAATTT TTTTTTAyCT TtGAAGAGCT 180
 CTTTTGGTGA TGkTTGcGAC ACTATACTAC CTTCTTTTCAT AACGATGACT TCATCTGCAT 240
 AACGCGCGAC TTCATTGATA TCATGTGAAA TTAGGATAAT TGCCTTATTT TCATCTGTTT 300
 GTAGTGACTT TAGTAATCTC ATTACTTGTC GTTTACTTTG TGGATCAAGT CCTGCTGTAG 360
 GTTCATCAAC CACGATAATA TCAGGATTCA TTGCCAATAT CGATACAATC GCTATTTTAC 420

	AATCCATCAA CAGACGATGG GCATAGTTTT TGGCTTCATC TAAATTCATT TTAAAGTTTT	540
	TAGGTCCAAA TATCATTCTA CGCTCTACTG TGTCCTCAAA TAATTGAGAT TCGGGAAATT	600
5	GAAATACCAT TCCAATTCTT TTTCTTACAG GTCTAATATA TTTATCTTTG GTCTTATGTG	660
	TAATAGTAAT GTCATCAACT GTAACGTGCC CAGTAGTCGG CTFTAACAGC GCATTAATAT	720
	TTTGTATCAA CGTTGATTTA CCACTACCCG TTTGTCCAAC GATGGCGTAA TATTTACCTT	780
10	GTTCAAATTC TGTATTAACA TCATGAATAG CTTGATGCTG ATATGGTGTC CCTTTTTGAT	840
	AGGTATAACT TACATTGTCA AACCGTATAG TCATAGTTGA TCCACCAGCC CTTCATAAGT	900
15	TAAGAATGAT GTTTGGTGTC CCAGCATTTG ATTTATTTTG ATTGGGAATG GCAAATCTAG	960
	ACCTATTCTT GTTAACTCTT CTGCATTGTC GAAAATTCA GTCGCTGTGC CTCTTTTATA	1020
	GACAGTCCCT TTATTCATAA CGATAACATG ATCTGCTTCC ATCGCCTCAG ATAAATCATG	1080
20	CGTAATAGAA ATGATTGTAA TATTATGTTT TGATTTAACT TTTCTCACTA AATCCAATAA	1140
	ATTTTGACGT GCATCAGGAT CTAACATAGA AGTCGCCTCA TCTAATATAA TGACAGAGGG	1200
	GTTAAGTGCT AATACACTTG CTATAGCCAC ACGCTGCTTC TGTCCCCCG ATAATGCATT	1260
25	AGGTTCAATA TCTGCACGTT CTAACATATC AACTTGTTTA AGTGCTTCGC TGACTCTTCT	1320
	ATGCATTTCG TCATATGGAA CCGCATGATT TTCGAGTCCA AATGCCACAT CGTATTTTAC	1380
	AATTGAACCA ACAAATTGAT TATCCGGATT CTGAAATACA ATTCCTATGT CTTTTCTTAA	1440
30	CTTTTCAAAA TTATCATCAG TTATAGCTTG ATTATTTATA AAAATTTCTC CAGATTTAAC	1500
	TTTCTCTATG CCAATCATTG ACTTGGCAAT TGTAGATTTT CCAGAACCGT TATGACCAAC	1560
	AATAGATGTC CACTGACCTT TAGGTATATT AAAAGAAACA TCTTTCAATG TGAAGGATGC	1620
35	ATCCTTTGA TATTGAAATG AAACATTTTT AAATACAATA ACTGAATTCT TATCCTCCAC	1680
	TTGTCTCTCT CCTTTACGAT TCGTGTATCT ATCATATTTT ACAATATTTA TAAATCGCTG	1740
	TATATGACAT TGACTGGGTT CTCTATATAT TACTAGTATT TTCTGACTCA TTTCTAGTCT	1800
40	TTAAAGTGTT GTTTAACAAC TAATGATAAG GACTTTTATT CCTCTCTAAC AATTATGTAT	1860
	AAACGTAAAT AAAATAAATG ATTTACTAAT ATAGGGGTGG TCGCGTTTGA TTCAACGATA	1920
45	ATACTTTCAC TTCATTCACT TCTAGTGAAA TTGATCAAAC TAGCTTCATC ATATTTTTAG	1980
	ATTCGCACTC AAAAAAGTAA ATATAAAGAA ATCGGACTTA AAAACATTTT TGTTCATAAG	2040
	TCCGATATTT TATTCAATAA AAAAGCGCGC ACCCCATCAT AAGTTTGTTG AGTTCACGCT	2100
50	TTAAATCTTT ATTTAGTTGA TGGGGTACTC TGAGCTAGAC AATATTTGTA TGTGGCAAAC	2160
	ATTATCGTTG CACTCATTTG CTTTATATAA AAGTAGTTAG TGTATTTATA TAAATCTTA	2220

55

EP 0 786 519 A2

	ACGAGTGTA	CCACCTTGAC	GTTCTGTGTA	AcGCTCTGCG	ATTTCACCAA	ATAATTTTTG	2340
	AAGTGCAGTT	TGTGTAGTTT	CATCTTCGTT	TAAGATTTCA	ACATTACGTA	AAGTTTTAGC	2400
5	TGCATTACGA	CGAGAAGCTA	AATCTCCTTT	TTTACCTAAA	GTGATTAATT	TCTCAACAAC	2460
	ACTGCGAACT	TCTTTTGCAC	GAGCTTCTGT	AGTTTCAATA	CGTTCACTAA	TAATAAGTGA	2520
	TGTAGCTAAG	TCACGTAACA	TAGCTTTACG	TTGATCAGAA	GTACGACCTA	ATTTTCTGTA	2580
10	ACCCATGAGT	TAACCTCCTT	TATCAATCTT	CTTTTCTTAA	TCCTAATCCT	AAATCTTCTA	2640
	ATTGTATTTT	AACCTCTTCT	AAAGATTTAC	GACCTAAATT	ACGCACTTTC	ATCATGTCAG	2700
	CTTCAGATTT	GTCAGCTAAC	TCTTGAACAG	AATTGATTCC	TGCGCGTTTT	AAGCAGTTAT	2760
15	ATGAACGTAC	AGATAAGTCT	AATCTTCAA	TAGACATTTT	TAATACTTTT	TCTTTTTTGAT	2820
	CTTCTTCTTT	TTCAATCATG	ATTTCAAGCT	TTTGCCTTC	ATCAGTAAGA	CCAACGAAGA	2880
20	TATTCAGTG	TTCAGTCATT	ATTTTGTCTG	CTAATGAAAC	TGATTCTTGT	GGTGTGATTG	2940
	AACCATTAGT	CCAAACATCC	AATGTTAATT	TATCAAAATC	ACTGCTTTGA	CCTACACGTG	3000
	TATTTTCAAC	AGTATAGTTC	ACACGTTCAA	CAGGTGAATA	CAATGAATCA	ACAGGGATTA	3060
25	CACCAATTGG	TAAATCACTA	GTATTATTTT	GTTCTGCTAA	TGCGTAACCT	CTACCCTTGT	3120
	TAGCAACTAG	ACGAATTTTT	AAGTGACCAC	CTTAGATAC	TGTTGCAATT	TTAAGCTCTG	3180
	GGTTTAAAAT	TTCAACATCA	CTATCATGTG	TAATGTCGCT	TGCTGTTACT	TCGCCTTCAT	3240
30	CACGTACATC	AATTTCTAAA	GTTTTATCTT	CTTCAGAGTA	AATTTTCAAT	GCTAATTGTT	3300
	TAATGTTTAT	AATAATTGTA	GAAACATCTT	CAACTACATT	GTCTACTGCT	GAGAATTCAT	3360
	GTAAAACTCC	CTCAATTTCA	ATATACTTAA	CGGCTGCACC	TGGTAATGAA	GATAGTAGGA	3420
35	TACGACGTAA	GGAGTTTCCT	AGTGTAGTAC	CGTAGCCACG	TTCTAGTGGT	TCAACAACGA	3480
	ACTTACCGAA	TTTAGCATCT	TCCTAATTTT	CAATTGTCTC	AATTCTAGGT	TTTTCGATTT	3540
	CTATCATTTA	AATATCCTCC	TTATATACGT	CGACTTAATT	TAAAATGTTT	GCTCAGTGAC	3600
40	CTGTAACAAT	ACCATCATAA	ATTATACACG	ACGACGTTTT	GGTGGACGAC	AACCGTTATG	3660
	AGGTACTGGA	GTAACGTCTC	TGATCGCAGT	TACTTCTAAA	CCTGCAGATT	GTAATGCACG	3720
	AATAGCTGAT	TCACGACCTG	GACCAGGTCC	TTTAACTGTT	ACTTCAACTG	TTTTTAAACC	3780
45	ATGCTCCATA	GCTGATTTAG	ATGCAGTTTC	AGAAGCCATT	TGTGCTGCAA	ATGGTGTTGA	3840
	TTTTTTAGAT	CCTTTGAATC	CTAATGCACC	AGCTGATGAC	CATGATAAAG	CATTACCGAA	3900
50	CTCATCAGTG	ATAGTTACAA	TAGTGTGTTT	GAATGTTGAA	CGGATGTGTG	CTACACCATT	3960
	TTCAATATTTC	TTTTTCACTC	TACGTTTACG	AGATACTTGT	TTACGTGCCA	TTTAAAATTT	4020

55

	CGCGCGTkgT TTTTCGTTTT TTGACCACGA ACTGGTAAAC CACGACGGTG ACGGATACCC	4140
	ACGGTATGAT GAAATThCCA TTAAACGTTT GATATTTAAG TTAGTTTCAC GACGTAAGTC	4200
5	ACCTTCGACT TTATAACCGT CTACAACCTC ACGGATGCGA CCTAATTCGT CATCAGTTAA	4260
	ATCTTTCACA CGAGTATCAG CTGATACGTT AGCTTCTTCA AGAATTTTTT GAGCAGTTGA	4320
	CGTACCGATA CCGTATATAT AAGTTAATGA GATAACTACG CGTTTTTCAC GTGGAATATC	4380
10	TACTCCTGCA ATACGTGCCA TATTAATTTA CACCTCTCTT TTATTAACCT TGTCTTTGTT	4440
	TGTGTTTTGG ATTTTCACAA ATTACCATTA CTTTACCTTT ACGTTTAATG ACTTTACATT	4500
15	TTTCGCAAAT AGGTTTTACT GATGGTCTTA CTTTCATTTT TATACCTCCC TATATTATGG	4560
	AGTGACGATT ATTTATAACG ATAAGTAATT CTCCGCGTG TTAAATCGTA CGGAGACATC	4620
	TCAACAGTTA CTTTGTGCGC AGGTAGAATA CGAATGTAAT TCATTCTGAT TTTACCACTT	4680
20	ACGTGAGCnA AAATCTCATG ACCATTTTCT AATTCTACTT TAAACATTGC GTTCGGTAAA	4740
	GSTATCTAATA CAGTACCTTC TAATTCAATT ACATCTTGTT TAGCCATTGA TTAACCTCCC	4800
	CCTTTTTGCA ATAGTAAGGT AATCGTCAAT AGACAACCTT ATTGTTACGA ATCTATCAGT	4860
25	GATTAATTTT ATAAGTTAAA CAAAAATTAC GGAATTAAT TATCGTTAAT TGCCACTCTC	4920
	ATCTATCTAA TATGATTAAA TCATGCCTCA CTTAAAATAG ACCGCTAAAA GTTGATCTAT	4980
	TACAAATGAT CTAAAATATC AATGACATCT TTGGTAACGT CGCTAATATC TTTTGAACCA	5040
30	TCAATATTTT TCAATACACC TTTTGTGATCA TAGAAATCTA AAATAGGCTT AGATTGTTTA	5100
	ATATTAACAC TCAAACGATT AGCTACCGTT TCAGGATTAT CATCTTCTCG TTGATACAAT	5160
35	TTACCACCAT CGATATCACA AATACCTTCG ACTTCGGAGG ATTAAATACA AGATGATACG	5220
	TTGTACCACA TGA CTCACAG ATTGACGAC CTGTAAGACG GTTCATTAAT TCTTCTCCG	5280
	GAAC ^T TCGAT ATTGATGACA GCATCAATGT TTCTGTCAAG CTCAGACATA ATATTATTTA	5340
40	ATGCCTCAGC TTGCTCGATT GTTCTTGGA AGCCATCTAA TAAAAAGCCT TTTTTGCAT	5400
	CGTCTTCAGA AATTCTTTCC TTAACGATAC CTACAGTCAC TTCATCAGGA ACTAATTCGC	5460
	CACGGTCCAT ATAAGACTTA GCTTCTTTAC CTAATTCAGT TTCTTCTTTT ATAGCTTTTC	5520
45	TGAACATGTC ACCAGTTGAA ATGTGGGGTA TTGGGAATTT CTTGaCAATT TCACTTGCTT	5580
	GAGTTCCTTT ACCTGCGCCA GGTAACCCA TCAAAATGAT ATTCATAAGT GCCCTCCTAA	5640
50	AATTATCTAC CACCAAAGCC TTTATATTCT TTTTGAGATA CTTGAGCTTC TAAAGATTTT	5700
	ATTGTTTCAA TCGCTACACC AATAACGATA AGTAAACTTG TACCACCAAT CTGAATTGAT	5760
	TGTGGTAATC CCATAAACTT AGTTGCTAAT ATCGGTAGAA TTGAAATAAC GGCTAAGAAG	5820

55

	CCAGGTCTAA TACCTGGAAC ATAGCTACCT TGTTCCTTAA GGTTATCAGC CATTTCCTCC	5940
	GGATTAACCT GTACAAATGC ATAGAAGTAT GTGAATAGTA TAATTAGTAC AATATATACA	6000
5	ACCATACCAA CATTACTTGA AGGATTGCA GCATTGCGAA TGTTCCTGTC CCATTCTTTA	6060
	TCTGGATAGA ACAACGTAA TGTTCCTAGC AGTAAGAAGA ACGCCATTGC AAAGATTACA	6120
	GGATAACAC CGGCTGAGTT CACTTTTAAA GGTAGATAAG TTGCCTGTGA ACCTAATCTT	6180
10	TGAGCAGTTT GTTCCTTAGC ATATTGAATC GGAATTTTAC GAACGGCTTC AAGTACATAA	6240
	ATAGCACCTA CAGTTAATAG TATCAGTGAC ACTAAAAGTC CTAATACTTT CAACCATGCT	6300
	AATGATGTAT CTTCCTGCCC AACGAACGCA TTTGTCCAA TTGAATTAGA CTGGCTGGCA	6360
15	ACGTGATAA AATACCCGCA AATATGATAA TAGAAATACC ATTACCAACA CCGAAGTGAG	6420
	TGATTTGATC ACCAAGCCAT ATTAAGAAAG CAGTTCCTGC TGTnCAAAAC TAGTGCTATT	6480
20	AATAAATAAC TCATAATTGA CTGATTGATA ATCAGCGCAC CTTTGAGATA ATTATTAAAT	6540
	TGGAATGCCA TACCTATAGA TTGGATAAAT GCTAAAGAAA TTGCTAAATA ACGAGTAACG	6600
	TTATTTAACT TTCTCTACC TACTTCACCT TGTTCCTGCCC ATTCTGAGAA TTTAGGGACA	6660
25	ATATCCATTT GTAATAATTG CATTACGATT GATGCACTGA TGTAGGGTAC AATACCCATT	6720
	GCAAAAATAG AAAATCGTTT CAGGCTCCG CCACCAAAG TATTTAATAA CTCAGTGGCA	6780
	CCTTGAGAAC CTGCGGATT ATCAAAGCT GCAGGATTTA CTCCTGGAGC TGGTATATAA	6840
30	GTCCCTATTT TAAAAATTAC TAACATTGCT AGTGTGAAGA AAATCTTGTT ACGAACCTCT	6900
	TTTGTCTAA AGAAGTTCAC AAGGGTTTGA ATCATTAGAT CACCTCGTGT GCTCCACCTT	6960
	TAGCATCAAT AGCTTCTGCT GCTGAAGCTG AGAATTTATG AGCTTTCCT GTCAATTTCT	7020
35	TATCAAGTGA ACCATTACCT AGTATTTTGA TACCAGATTT TTCATTCTTA ACAACACCAG	7080
	ATTCTACTAA TAAAGCTGGA GTTACTTCAG TACCATCTTC AAATTTATTA AGTTGGTCTA	7140
40	AGTTAACAAT AGCATATTCT TTACGATTTA TGTTAGTAAA ACCACGTTTT GGTAAACGAC	7200
	GGAATAATGG TAATTGACCA CTTTCAAATC CTGGTCTTAC ACCACCGCCT GAACGAGCTT	7260
	TTTGACCTTT GTGTCCGCGA CCACTTGTTT TACCGTTACC TGTCGCAACA CCACGTCCAA	7320
45	CACGATTGCG TTCTTTACGT GAACCTTCTG CCGGTTTTAA CTCATGTAAT TTCATTTCGG	7380
	CACCTCCTTG ATTATTTTTT TTCTACTGTT ACTAAGTGCT TAACCTTGTT GATTGCCCCA	7440
	CGAATAGCAG GGTTATCTTC AACAACTACT GAACTGTTAG TCTTTTAAAG ACCTAAAGCT	7500
50	TCAACAGTTT TACGTTGTGT TTCAGGACGA CCAATAACAC TACGAGTGAG GGTAATTTGT	7560
	AAATTTAGCCA TAACTAGTTT TCCCTCCTTA ATTGTATAAT TCTTCTACTG TTTTGCCACG	7620

55

	CATGTTGATT	GGTGTGTTTG	ATCCTAATGA	TTTACTTAAG	ATATCAGTGA	TACCTGCTAA	7740
	TTCAAGTACG	GCACGAACAG	GACCACCAGC	GATAACTCCT	GTACCAGGTG	CAGCCGGTTT	7800
5	CATAAATACG	CTTCCTGAAC	CGTAACGGCC	AGTAATTGTG	TGTGGAGTTG	TACCTTCAAC	7860
	ACGTGGAACA	ACTACTAAAT	CTTTTTTAGC	TGCTTCAACA	GCTTTTTTGA	TTGCTTCTGG	7920
	TACCTCTTGA	GCTTTACCAG	TACCGAAACC	TACACGACCA	TTTTTGTCTC	CAACTACAAC	7980
10	TAATGCAGTG	AAACGGAAAC	GACGACCACC	TTTTACAAC	TTTGCTACAC	GGTTGATTGT	8040
	AACAACGCGT	TCTTCAAAT	CTTTCGTCTC	TTCTTCTCTA	CGAGCCATGT	ATTTGTCCCT	8100
	CCTTTAAAT	AAAATTCTAA	TCCGCTTTCT	CTTGCTGCTT	CAGCTAATGC	TTTAAACGCT	8160
15	CCGTGATATA	AATATCCTCC	ACGGTCAAAT	ACGATTTCTT	TAATGCCTTT	GTCAGCAGCT	8220
	TTTTTAGCAA	TTGCTTCACC	GACTTTAGTT	GCTAATTCAA	CTTTAGTTGC	TGTAGTAGCA	8280
20	ATGTCGCTGT	CTTTTGAAGA	AGCTTGAGCT	AATGTTACGC	CTTTATTATC	ATCAATAATT	8340
	TGAGCGTAGA	TATGCTTGTT	TGAACGATAT	ACGTTTAAAC	GTGGCTTTTC	AGCTGTACCT	8400
	GATAAGTTAG	TACGAACACG	AGCATGTCTT	TTTAAACGCA	CTTTATTTTT	ATCAATTTTA	8460
25	CTGATCATTT	CAATACTCCT	TTCTTTAGAG	TTTATCTATT	ATTTACCAGT	TTTACCTTCT	8520
	TTACGGCGAA	CGTATTCACC	TTGGTAACGA	ATACCTTTAC	CTTTGTAAGG	CTCTGGAGGT	8580
	CTTACTGAAC	GGATGTTAGA	TGCTAATGCT	CCAACCTGTT	CTTTTGAAAT	ACCTTCAACT	8640
30	TTAACGACTG	TGTTTTTCTC	AACGTAGAAA	GTAATGTTTT	CTTCAGCTTT	AATTTCTACT	8700
	GGGTGAGAAT	AACCAACGTT	AAGGATTAAG	TCTTTACCTT	GCATTTGAGC	ACGGTAACCT	8760
	ACACCAACAA	GTTCAAGTAC	TTTTACGTAT	CCTTGAGAAA	CACCTTGATC	CATATTGTTT	8820
35	AATAAAGCAC	GAGTTGTACC	ATGGTTTGTT	CTATCTTCTT	TAGAATCAGA	TGGTCTTACA	8880
	ACTTCAATTG	TGTTTTCTTC	TTGTTTGAAT	GTCATTCTTT	CATTTAAAGT	TCTTGATAAT	8940
40	TCACCTTTAG	GACCTTTAAC	AGTTACATGA	TTTCCATCAA	AAGTTACTGT	TACGTCACCTA	9000
	GGGATGTCAA	TAATTTTCTT	ACCAACACGA	CTCATGTTAT	GGCACCTCCT	TATTTTTTAT	9060
	TACCAAACGT	ATGCGATAAT	TTCTCCACCA	ACATTACGTT	TTCTTGCTTC	TTTGTCAGTG	9120
45	ATTACACCTT	CAGAAGTTGA	TACTAATGCA	ATACCTAAAC	CATTTAATAC	TTTAGGCATT	9180
	TCGCTAGCTT	TTGCATAAAC	ACGTAAACCT	GGTTTTGAAA	TACGTTTTAA	TCCTGTGATA	9240
	ACACGCTCAT	CGTTTTGACC	ATATTTTAAG	AATAAACGAA	GTACACCTTG	TTTATCATCT	9300
50	TCTACGTATT	CAACATTTTT	AATGAAACCT	TCACTCTTTA	AGATTTACAGC	AATTTCTTTT	9360
	TTAATATTTG	ATGCAGGTAA	TTCTAACCTC	TCGTGACGCA	CCATGTTTGC	GTTTCTTACA	9420
55							

TCTTTTTTAT TACCAGCTAG CTTTACGAAC GCCAGGGATT TGGCCTTTGT AAGCTAATTC 9540
 ACGGAAACAA ATACGGCATA ATTTAAATTT ACGATATACA GAATGTGGAC GGCCACAACG 9600
 5 TTCACAACGA GTGTATTCAC GAACTGCATA TTTTGTGTTT TTTTGTGCT TAGCAACCAT 9660
 TGAAGTTTTA GCCACTTAAT TAGCCTCCTT TAAATAATTA TTTACGGAAT GGCATACCGA 9720
 10 AGTTAGCTAA CAATTCACGA GCTTCTTCAT CAGTGTTAGC AGTCGTTACG ATAACAATAT 9780
 CCATTCTCTT AACTTTACTT ACTTTATCAT AGTCGATTTT TGGGAAAATT AATTGTTCTT 9840
 TAACACCTAA AGTGTAGTTA CCGCGTCCGT CAAATGCTTT TTTAGAAACA CCTTGGAAGT 9900
 15 CACGTACACG TGGTaaTGAT ACTGAAATTA ATTTGTCTAA GAATTCATAC ATTCTTTCAC 9960
 CGCGAAGTGT TACTTTTCGA CCGATTGGCA TACCTTCACG TAAACGGAAA GTCGCGATTG 10020
 aTTTTTTAGC TTTAGTTACT AATGGTTTTT GACCAGTGAT CAATTCTAAT TCTTCAACAG 10080
 20 CATTGTCTAA TACTTTAGAA TTTTGTACTG CGTCACCTAC ACCCATGTTT ACAACGATTT 10140
 TATCTATTTT TGGTACTTCC ATTACTGAAC TATAATTGAA TTTTTCATT AAGTTTTTCAG 10200
 TAACTTCAGT GTTaaACTTT TCTTTTaAAC GGTTCaAAGT GGGATCCTCC TTTCaACTTG 10260
 25 TtATTAAATTA TTAGAKTTAA TTTCTTCGCC AGATTTTTTA GCGATACGAA CTTTTTTTACC 10320
 ATCAACAAAT TTGTAACCTA CACGAGTTGG TTCGTTTGTT TTAGGGTCCA ATAATTGTAC 10380
 ATTAGAAACA TGGATTGCTG CCTCTGTTTC TAAGATTCCA CCTTCAGGAT TTAATTGAGT 10440
 30 TGGTTTTTGG TGTTTTTTCA TAATGTTAAC ACCTTCACAA ACGACACGGT CTTTTTTTAGG 10500
 TAGAGTAGCA ATTACTTTAC CTTCTTTACC TTTGTCTTTA CCTGCGATAA CTTTAACGTT 10560
 35 GTCACCTTTT TTGATATGCA TGTGGGCACC TCCTTATTTG TATTGGTTGT TATTAATTAA 10620
 AGTACTTCTG GTGCTAATGA TACGATTTTC ATGAAGTTAC CTTACGTAA TTCACGACGA 10680
 ACAGGTCCGA AGATACGAGT ACCACGTGGG CCTTTGTCAT CACGGATGAT AACACATGCA 10740
 40 TTTTCATCAA ATTTGATGTA TGAACCGTCA TTACGACGAA CACCTGACTT AGTACGTACG 10800
 ATTACAGCTT TGACAACGTC ACCTTTTTTA ACAACGCCAC CTGGTGTTGC ATTTTTAACA 10860
 GTACATACGA TAACATCGCC GATGTTTGCT GTTTTACGAC CAGATCCACC TAATACTTTG 10920
 45 ATTGTAAGAA CTTACAGAGC ACCAGAGTTG TCTGCTACTT TCAAGCGTGT TTCTTGTTGG 10980
 ATCATTAGTT AAACCTCCCT TATCTCTAAA CTTGTATTAA ATAATTACTG ACTCTTCAAC 11040
 AATCTCTACT AAACGAAAAC GTTTTGTTGC TGATAAAGGA CGAGTTTCTT GAATTTTAAC 11100
 50 AATGTCTCCT AATTAGCTG AATTGTTTTT ATCATGAGTT TTGTATTTTT TAGAGTATTT 11160
 TACTCGTTTA CCGTATAATT TGTGTGTTTT GTAAGTTTCA ACAAGTACTG TAATAGTCTT 11220

55

TTTTGTAACC TCCTCTTACT TAATTATTGA TTAGCCTTAC TTTGTTCAAT TTCTCTTTCA 11340
 CGAGCAACAG TTTTGTAGACG TGCAATCGTT TTTCTTACTG TACGAATACG TGCAGTTTCT 11400
 5 TCTAATTGAC CTGTAGCTAA CTGAAAGCGT AGGTTAAAAA GCTCTTCTTT TGAAGATTTG 11460
 ATTTGTTCTT CGATTTCTGA AGTGGTTAAG TCTCTAATT CCTTAGCTTT CATTTGTTTC 11520
 ACCACCCAAT TCCTCACGTT TTACAACTT AGTTTTTACT GGAAGTTTGT GACTTGCTAA 11580
 10 ACGTAGTGCT TCACGCGCAA CTTCTTCAGA AACGCCAGCA ACTTCGAATA AAATTCTACC 11640
 TGGTTTAAACA ACTGCGATCC AGCCTTCAAC CGCACCTTTA CCAGCACCCA TACGTACTTC 11700
 TAAAGGTTTT TTAGTATATG GTGTATGTGG GAAGATTTTA ATCCAACTT TCCCGCCACG 11760
 15 TTTTCATGTA CGTGTCATTG CTATACGAGC AGATTGATT TGACGAGATG TGATCCAAGA 11820
 CGTTGTTGTA GCTTGTAAC CAACTCACC AAATGTTACG TATACCGCC TTTAGAACGA 11880
 20 CCAGTTGTTT TAGGACGATG TTGACGACGA TATTTTACAC GTTTTGGTAG TAACATTATT 11940
 ATTTTCTCC TCCACTAGTG TTCTTAGTAG GAAGAACTTC TCCACGATAA ATCCATACTT 12000
 TAACGCCTAA TTTACCGTAA GTAGTGCAG CTTACGCGTG TGCATAATCG ATGTCAGCAC 12060
 25 GTAACGTATG AAGTGAACA GTTCCTCTG AATATTGTT AGCACGAGCG ATGTCAGCTC 12120
 CGCCTAAACG ACCAGATACT TGAATTTTGA TACCTTTAGC ACCAAGTTTC ATAGCTCTAG 12180
 TGATTGCTTG TTTTGTACA CGACGGAATG AAGCACGGT TTCTAATTGA CGTGCGATGT 12240
 30 TTTTCAGTAC TAAACGAGCG TCAAGATCAA CTTTTTGTAT TTCAATTACG TTGATGTGTA 12300
 CTTTTTATC AGTTAACGCA TTTAATTGT TGCCTAATT TCGATTTCT GAACCGCCTT 12360
 TACCAATTAC CATACCAGGT TTACCAGTAT GAATTGCAAT GTTGATACGG TTTGCAGCAC 12420
 35 GTTCAATCTC TACGTGAGAA ACTGATGCTT CTTTAAATTC ATTATCAATA AATTTACGGA 12480
 TTTTAAATC TTCGTGTAAG AGTGAAGCGA AGTCTTTTC AGCATACCAT TTAGCTTCCC 12540
 AATCACGGAT AATACCAACA CGAAGTCCGA TTGGATTAAT TTTTGTACCC ACAGTATTCC 12600
 40 CTCCTTAAAA GTTAATTAAG CTTCTTTAGC TTCTTCTTTA CCGTCACTTA CGACGATTGT 12660
 AATGTGGCTT GTACGTTTGT TAATCGCACT TGCACGACCT TGCGCACGTG GACGGAAACG 12720
 45 TTTAATGTT GGTCTTCGT TAGCATATGC TTCTTAACT ACTAATTCAT CTGTGTTTAT 12780
 GTCATAGTTA TGTTGAGCAT TAGCTAAAGC GGACATTAAT ACTTTTCAA TTACTGGTGA 12840
 TGAAGCTTTG TTTGTTAATT TTAAAATTGC AATAGCTTCA GCAGCATTTT TACCTCTGAT 12900
 50 TAAGTCAAGA ACTAGTCTTA CTTTACGAGG TGCGATTCTT ATTGTTCTAG CAACCGCTTT 12960
 TGCTTCCATT AGGATGTCCT CCTCTACTTA ATAGATATTA TCTTCTGTT TTCTTGTCGT 13020

55

EP 0 786 519 A2

TATCTTCAGT TACATATACA GGTACGTGTT TACGTCCGTC GTATACTGCA AAAGTATGTC 13140
 CGATGAAATT AGGGAAAATT GTAGAACGAC GTGACCATGT TTTGATTACT TGTTCCTTTT 13200
 5 CGCTTCCTTC TTGAGCTTCA ACTTTTTTCA TTAAATGCTC ATCGACGAAA GGTCTCTTTT 13260
 TAATACTACG AGCCATTTGG GCGCCTCCCT TCTTATTATG TCGTGCAGC TTTAAGCCGC 13320
 ACACCCAAAT AAGTTGATTA TATTATTTTT TCTTACGTCC ACGAACGATA AGTTTGTCTG 13380
 10 ATGATTTTTT ACCACGACGA GTTTCTTAC CAAGCGTAGG TTTACCCCAT GGTGACATTG 13440
 GAGATGGTCT ACCGATAGGA GCACGACCTT CACCACCACC GTGTGGGTGA TCGTTAGGGT 13500
 TCATTACAGA ACCACGAACT GTTGACGGA TACCTTTCCA TCTTGAACGT CCGGCTTTAC 13560
 15 CAACGTTAAC TAATTCGTGT TGTAGGTTAC CAACTTGACC GATTGTAGCA CGGCAAGTAG 13620
 ATAAGATCAT ACGAACTTCA CCAGATCTTA ATCTGATTAA TACGTATTTA CCTTCTTTAC 13680
 20 CAAGTACTTG AGCACTTGCA CCAGCTGAAC GAGCGATTG TCCACCTTTA CCAGGTTTAA 13740
 GCTCGATGTT GTGTACTACT GTACCAACTG GAATGTTTTG TAATGGTAAT GCGTTACCAA 13800
 CTTTGATGTC AGCTTCAGCA CCACTTTCAA CGATTGACC TACTTCTAAT CCTTTAGGAG 13860
 25 CAATGATATA TCGTTTTTCA CCGTCTGCAT ATACAATAA AGCGATGTTT GCTGAGCGGT 13920
 TTGGATCATA TTGAATAGAA TCAACTTTTG CATTGATACC ATCTTGTTA CGTTTGAAAT 13980
 CGATAACACG GTATTGACGT TTGTGTCCAC CACCATGGTG TCTTACAGTC AATTACCTT 14040
 30 GGTGTTACG TCCCGCTTTT TTCGGTAGCG GTTTTAATAA TGACTTTTCA GGTGTAGTTT 14100
 TCGTGATTTC TCGGAAATCT AACGAAGTCA TATTACGACG ACCATTTGTT ATTGGCTTAT 14160
 ACTTTTTAAT AGCCATTGTC GCTTACCTCC TTAATGGTAA TTGTTTTATT AGTTAAATAA 14220
 35 GTCGATTGAT CCTTCTTTAA GAGTTACAAT CGCTTTTCTT CTTTGTTTG TATAGCCTTG 14280
 GTAACGGCCC ATACGTTTTT TCTTAGGTTT GTAATTCATG ATATTAACAC TTGCAACTTT 14340
 40 TACGTTGAAG ATTTCTTCAA CTGCCATTTT TACTTGTGTT TTGTTAACAC GAGTATCAAC 14400
 GTCGAAAGTG TATTTGTCTT CAGCCATTGC TTCAGAAGAT TTCTCAGTGA TTACGGGGCG 14460
 CTTAAGAATA TCTCTTGCTT CCATTATCCG AGCACCTCCT CAACTTTTTT AGCAGCAGCT 14520
 45 TCAGTAATTA CTAAGCTGTC AGCATTAGTG ATATCTAAAA CATTTAAACC TTGAGCAGTT 14580
 GTCACTTGAA CGCCAGGGAT GTTGGGTGCT GATAATTCAA CATTTACATC TTCGTTTTCA 14640
 GTAACACTA ATACTTTTTT AGGTGTTCT AATGTAGATA ATACATTTTT GAATCTTTA 14700
 50 GTTTTTGGAG CTTGGAAGTT GAATGCGTCA ACTACAGTTA AGCCATTCTC TTGAGCTTTG 14760
 AAAGATAATG CTGAGCGTAA AGCTAAACGA CGCATTTTCT TAGGCATTTT GTATGCATAA 14820

55

	CCTTGACGAG CACGACCTGT TCCTTTTTGC TTCCATGGTT TACGTCCGCC ACCGCTTACT	14940
	GCTGAACGAT TCTTAACAGC ATGCGTACCT TGACGTAATG AAGCACGTTG TAAATTAATA	15000
5	GCTTCGAATA AAACGCTATT ATTTGGCTCA ATACCGAATA CTGCATCGCT TAATTCGATT	15060
	GAACCTGATT TAGTTCGGTC TAATTTTAAA ACATCATAAT TAGCCATTAT GCATTCCTC	15120
10	CTTTCACCTC TTATTATTTA TTACCTTTTT TAATTGAAGT TCTGATTCTT ACTAAACCTT	15180
	TTTtaggtcc AGGTACGTTA CCTTTTACTA AGATAACTTT GTTTTCTGTG TCAACTTGAA	15240
	CTACTTCTAA GTTTTGAACA GTTACAGTGT TTCCACCCAT ACGTCCTGGC ATTTTTTGGC	15300
15	CTTTAAATAC TCTAGAAGCA TCTGAAGCCA TACCTACAGA ACCTGGTGCT CTGTGGAAAT	15360
	GAGAACCGTG TGACATAGGT CCACGAGATT GTCCGTGGCG TTTAATTGCA CCTTGGAAAC	15420
	CTTTACCTTT TGATACGCCT GTTACGTCAA TAACGTGCGC AGCTACAAA GTATCTACTG	15480
20	AGACTTCTTG AaCCTAcTcG TAAGCATCCA CGTCTACATT GCGGAATTCA CGAATGAAGC	15540
	GCTTAGGTGC TCGCTCAGCT TTTTtagcgt GACCTTCAGC TGGTTTATTA GCATATTTAT	15600
	TAGATTTTGC ATCTTTTTTG TATGCTTTTT TGTCTTCAA TCCAACTTGG ATTGCGTTGT	15660
25	ATCCATCAAC TTCTACAGTT TTCTTTTGTA ATACAACATT TTCTTTAGCT TCTACTACTG	15720
	TTACAGGGAT TAATTCACCG TTTTCTCCGA ATACTTGTGT CATCCCAATT TTTCTTCTA	15780
	AGATTCCTTT GGTcatcgaa AGTCCACCTC CTAAAATTGT CTATTATAAT TTGATTTCGA	15840
30	TGTCTACACC AGATGGTAAG TTTAAGCCCA TTAAAGCGTC AACTGTTTTT GGTGTGGGT	15900
	TTACAATATC GATTAAACGT TTGTGTGTAC GTTGTTCGAA TTGTTCACGT GAATCTTTAT	15960
35	ACTTATGCAC GGCACGGATG ATTGTGTAAA CTGATTTCTC AGTTGGTAAC GGAATTGGTC	16020
	CAGAAACATC TGCACCAGAA CGTTTCGCTG TTTCTACAAT CTTCTCTGCT GATTGATCAA	16080
	TTACGCGGTG ATCATAAGCT TTTAATCTGA TTCTGATTTT TTGTTTTGCC ATAATTTTCC	16140
40	CTCCTTATTC GTCTACATTT AGTGATAGAC TTCTCCACGA AACTATCTT ACACAGCGCC	16200
	ATGGCAAAGC GGCCGGGTGT GTCAGTAACC TTTGCTTCA TCGCTTTTCT TAAAGTCCAA	16260
	CGTTAGTTAT ATTACACGAA AAACATCGAT AAATCAAGGC TTTTCACATA ATTTTTCTAT	16320
45	CTGTCTAACA CATACTTTTA TATTTnACTT TATATACTTA GTCAGTTCAA CTATTTTCGA	16380
	GATATTTTnA ATTTCCn	16397

(2) INFORMATION FOR SEQ ID NO: 206:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29555 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

5	TnAGTTGTTT CTGCCACGAA AGATTCAATG GCTTTTCTTG CTTTACGCTT TTCTTTCAAT	60
	GGCAAATCAC CAATCATTTT TTTAAGGTGA TGTGGGTTTA CAACACCACT ATACTGGTAG	120
	TCATTTGAAn TTGTTTTTAG GGCTTGTTCA TCGATAGATC TCTCTCCAGC AAATCCTTTG	180
10	AACTCCGCTT CTTTTTTAAT ACTTTCGAAA TTAACATATT CTTGATCGAT ATCATCATCC	240
	TTATTTAAAG AAGGTACAAC ATTGTCGATG AATTCTCTAA TTAGATCTCG TTTTAACCTC	300
15	AATGtCGGAT CATCTGCATG ATCTAAAATG CGTCTAATTT GTTCTTGGTT ACGACGTTGT	360
	TCCGCTTTGT CTTCAAGATC AATTGTCTC AATATATTCA TAATATAATT CACATTAATC	420
	GTATCATTAC GCATCATTTT TATTTTCGAAA TCAATATCAT TTAAAATGGA TACTTTATTT	480
20	TTCTCAGCCG TCGCTCTTTT TACTTGATCG TACACAGCTA AATATTTACT TTTATAGTCT	540
	TCATTCTCTT GTTCATCCAT TCCAATTTCA TCAATTGTAA ACTCAAACCTC GTCAAATGCT	600
	TTTAAACGTA ATATTATTTT AGCTAATAAA CGATAAGCtT CAACAAAGCG CTTTAGCTCT	660
25	TCTTCATCtT GaATGtCATC AACCATGTGT GGTGTCGGCA CAATCATTTT AAGCTCACGA	720
	TAAGCGTCCA TAAATTCTTT TTTATACTCT TCATAACTGC GCATTAAAAT TGTATCCGTA	780
	TCATTTGTTT GTGAGAATAC TCTCAGTGCA TCGTCTGTCT CTTTTTTCAA GTCACGATAG	840
30	TTTACAATTT TACCAAATGG CTTTGATTCT TTTTCAACCC TATTGTACG TGAATACGCT	900
	TGAATTAAAT CATGATACAT TAAATTCTTA TCAACATATA AAGTGTTTCA TACTTTACTA	960
35	TCAAAACCAG TTAAGAACAT ATTAACAACG ATTAAGATAT CAATTTTACT ATCTTTAACG	1020
	CCCTTTTTTAA CGTTTTTtGA AATATGATTA AAATACTCAT TAGTTGtGGC TgnTgaAAAA	1080
	TTCGTCTCGA ACTTTTTTATT ATAATCACTA ATCATTATCT CTAATTTTTT CACGTGAATGA	1140
40	TATGGCACTT CACCATCAG ATCATCTTCA TTAGGTTTAA ACGTAAATAT ACCAGCTATC	1200
	GTTAACGGTT GTTCCAACCT TTTGTTAAGT CGCTTAAATG TCTCATAATA TTTAATAAGC	1260
	GCGTGAATAC TTTGGACTGT AAATATACTT GAATATTGAC GATTACGTGT ATATTTATCA	1320
45	TGATTATTGA TGATATGTCG TGTTACTAAT TCCACACGTT TATCCGCTAA CCATACTTCT	1380
	TCCGTATCAA TTGCTTCAAC CATGctGTTA TCTTCTGCTT TTAAAGCTTT ATTTTTAAAA	1440
	GTATTAATAT AGTCAACTGA GAAACCAAGT ACATTACCAT CATGAATGGC ATCTCTAATT	1500
50	AAATACGTAT GTAAGCATCT ACCGAAAATA TCTGCAGTTG TTCTACCATC TTGACTACTA	1560
	TTTTCTGGAA AACGTGGCGT ACCAGTGAAT CCAAAGTATT GGGCATTTTT GAAATGTTGT	1620

55

	ACTTTATTTCG TTTTATACTG TTCTAATAAAA GGGGCATTCC CTTGAATCGC TTTAGCCATT	1740
	TTTTGAATCG TCGTTACAAT AAGTGGCAAA CTTTTATCAT TTAGTTGGCG TACCAGTTGC	1800
5	GAGGTATTAA AAGTTTTGTC TACAGCACCC TTAGCAAATT TATTAAATTC CTCTTCTGTT	1860
	TGACTATCCA AGTCTTTACG GTCAACCAAA AAGATAACTT TCTTAATGTC ATCTTGCTGT	1920
	GATAAAATCT GACTCGCTTT AAAAGAAGTC AACGTCTTAC CACTTCCAGT TGTATGCCAT	1980
10	ACATATCCAT TATTCCCTGT CTCAGTCGCT TGTGAATAA GTGCTTCTAC CGCATACACT	2040
	TGATACGGAC GCATTGCCAT CAGTATTCTA TCTGTTTCAT TAATAATCAT ATAGCGCGAT	2100
	ATCATCTTAG CTAATTGACA AGGTCTCATA AATGACTCAG CAAACGATTG CAATGTATTG	2160
15	ATACGGTTAT TCTGTTTATC ACTCCAATAA AACATGTGAC TCTTCAATAG TTCGCTATCA	2220
	TTATTAGAAA AGTATCGCGT TTCAACACCA TTAATAATGA TAAACATTTG TATGTAGCGG	2280
20	AATAAGCCTG TGTAATTTTG TTTGCGGTAA CGTTTTACTT GGTAAACGC CTCATTAAATA	2340
	TCAATACCTC GACGTTTCAA TTCAACTTGG ACAAGGGGTA GTCCGTTGAT TAATATCGTT	2400
	ACATCATAAC GTGCTTTATA TGTATCCTCG ACAGATACTT GATTCGTCAC TTGAAACTTA	2460
25	TTTTTACACC AACTTTTCGT ATCTAAAAAC GACAAATAAA TCTCAGACTC ATCATCACGT	2520
	CTAAGTGGTA ATTTATCACG TAAAATACGG GCACTCTCGA AAATACTTTT TCCATCAATC	2580
	ATCGTTAACA GACGTTGaAA TTCTTTATCT GTTaaGGGAT TGCCTTCTAA TTTGTCCGCA	2640
30	TGACGCTCAT TTAAAATCGT TCTAAAATTA TCAAGCAATT GCTTATTATC ACGTATCGTT	2700
	ACTCTTTCGT AACCCAATTG TTCAAGTTGA TTCATCATTT CATTTTCTAA TGCCTATTCA	2760
	CTTTGGTATG CCATTCATAT CCCCTTCCAT ACACTTTCTA TTGCTCTAAA TATATCATAA	2820
35	ACTTTAATGA AAAATGTTTG TTTTTTATCT TCAAACGTAA ATTTATTCTA ATTTTATTGT	2880
	CTTATCTTTT AATATTTGTC TTTGAGGTAA GTCGTATACT AAAATTTGAA TACAAATAAT	2940
40	CAAATCATTG ATAAATTTTT TGTCTACGAT TAATGGAGGG ACTTGAATGG TGTTAATTAC	3000
	CTATCAAATC ATTTTATTTT TTATTATTAG TCTAAGTTAC TATTAACTT TAAATCATTA	3060
	CATGGCAGTC ACTGTAGGTA ACTTCACTTC AATATTCGGC ATGTTGCGAG CCATACTCTT	3120
45	TATGTACTAC TACCTACTCT ATAAAAGTCC CGAATACAAT CAACGCAAAC GATTTAAACA	3180
	TTTCATTCAT ATCACTAATT TGATAATAAT TGCTTTTAGC ACCTTCGTAT TAGTTCATTT	3240
	AGCATTAAAA TTATTCTTCA GCATTTAATT TCCATCTATG AAAAAAGCAA AGCTCAAATC	3300
50	TGAACTTTGC TTTAATTTGT CACGCCTTTA TCATTTTCAA AATAGCCTCT ATGCCAGTTT	3360
	TACAAACTTG TAGCAACAAT TTTTCATCAA GCAACTGAAT CACATCAAAA ACTTCAATTG	3420

55

	GTCGCAAGAT GCTTCCTGTA ATTATCAAGT GCCATTTTCG ATTGGGTTAT ACAATCTAGA	3540
	ATCGCATGAT AATTTAATGC TACAAATCGA TAGTACAATA TATCTACCGT GAATAACTGT	3600
5	GCAAATAGTG ACGTTGTAGC CGCCATACGC ATTTTCATTTT CATCAGTTCT GCCATAAATC	3660
	AATGCATAGT CTGCAATTTG AGCCACTGGA TTATTAGCTG TACTAGATAT AGTTATGATG	3720
10	GGAATACTGT AATGTGTGGC CACCTGTGCA ATTGACTGCA ATTCACTATG ACTACCTTGA	3780
	TTCGTCACAA AAATCATGCA ATCTCTATCA TCATGCGTCG CAAATGTTGA CACAAGTAA	3840
	TGCGTTTCAT GTAATAACCT GACATTTAAG CCAATACGAG ATAACTTTTG AAAAAGATCA	3900
15	CCAATAGTCA AACTCGATGC GCCAAATCCA AATAAAAATA TTGTCCTGGC ATTTTTC AAC	3960
	ACATCACAAA TTGCATCAAT TTGCGCATCC ATAATATTAG TAGCTACAAA TCGCATCGTA	4020
	TTCGTTGCTC TAGCAATCAT TTTATTTTTC AAAGTTTCTA CAGATTCATT TTCAATCAAT	4080
20	TCTAAATGTG GATTGGTTGC AATATCTTCG GGTAAGTATC GAGATATCGC AATCTTTAGC	4140
	TCTTGAAAAC CTTGATGTGT CATTTTCCGA CTAAATCTAA CAATTGATGC TGTACTAACA	4200
	TTCGTAACAT CTGCCAAATC ATTCACAGTC ATATCAATGA TTTTATGTGG ATTCTTTAAA	4260
25	ATGTAATCAG CGATTATCTT TTCTGTCTTC GTAAATCAC TCAACTGCTT ATCAATGCGA	4320
	TATAAAATAT TTGTCATCAT TAATCACCCA ACAAATCTGT CTGTCGCATC GCCTTTGTGCG	4380
	TTCCAAATAA ATATGTACAA ACGAATCCAC CAGCATACGC AGCAAGTAAT CCTGCAATAT	4440
30	AACCTAAATA CATATTATCT GAGATTAATG GTAATAGTGA CACACCACTT GGGCCTATTG	4500
	CTTTGGCACC AATATGTCCA ATTCCACCTA TTACAGCGCC ACCAATACCA CCACCAATAC	4560
35	AAGCAGTTAA GAAAGGTCGA CCTAATGGCA AAGTCACACC ATAGATTAAT GGTTCCTCGA	4620
	TACCTAGGAA ACCAACTGGC AATGCACCTT TTAAAGTATT ACGTAATGTT GTGTTGCGTT	4680
	TACATCTTAC CCAAAGTGCT AATGCGGCAC CTACTTGTC AGCACCAGCC ATCGCTGCAA	4740
40	TTGGCAATAA GTAAGTAGCA CCTGATTGGT TAATCATTTT TATATGAATT GCGGTAAAA	4800
	TATGATGAAG CCCTAACATA ACTAACGGTA GGAAGCTTGC ACCAATGATA AATCCACTAA	4860
	ATACGCCACC AATACTAATA ATTCCGTTAA CTACTGAAAC TAAACTGTCT GAAACAAAAC	4920
45	CTGCTAATGG CATAAAGATA AAGATAGTTA ATAGTCCTAC AATCAACAAT GCAATAGTCG	4980
	GCGTTACAAT AATATCAATC GCATTTGGCA CAATTTTATG TAATCTCTTT TCGACAATAC	5040
50	TTAAAAATCCA AACGGCAAAA ATAACGCCAA TAATCCCACC TTGTCCAGGT TGCAATGGTT	5100
	CTCCAGTGAA GACATTCATT AAAATATTTT TACCAGCAAT ACCCGTTAAT AACGTTGTAC	5160
	CACCAATCAC GCCACCAAGT CCTGGTGTG CACCAAATTC TTTAGCCGCA TTAATACCAG	5220

55

	GCGTAATCCA AGCACCTGAA ATATAGCCTG CCACCATTAA GTTACTCAGT ACTGCTGCAA	5340
	TACCACCAAT TAATCCAGCT CCAATAAATG CAGGAATCAA CCGTATAAAG ATATTGGCAA	5400
5	TTGATTTCAA TACTTTATTC AACTTACCAT TCTTTTGTTT TGCTTTATGC GCTTCCTTAT	5460
	TCGCCTTTGC TTTATCAGCT GCATATGATT TATAGTCCAT TTTTCACTA TCATTGTGAT	5520
	GGTGTGGTAT TGGGTCACCT AGTTTAACAC CACTTAATTC CGCCATATGA TTAGCCACTT	5580
10	TATTGatGTA CCAGGTCCAA CCACAACTTG AATGCGTTCA TCGTGTATAA CACCCATGAC	5640
	ACCATCAATA TGCCTTAGTT CTTGGTCAATC TACTTTATTC TCATCTAATA CTTTAATACG	5700
15	CACACGTGTC ATACAGTTCA TGACACTATC TATATTATCC ATACCACCTA CTGCAGCAAT	5760
	AATTCGTTCT GCAAGTTGTT GTTCTTTGGT CATTTAAATC CCTCCTAAGG TTGTCTATCT	5820
	CTGATTGCTC GTTTAAaATG TCACCATTGT TTAATAACCG TCTTGTGCT TCTTCCTTAG	5880
20	AAATGCCACA CATACCCATA ACTGTGCGAA CTTTCACATC ATGCTCAGAT ACCTGATATA	5940
	ACGCCATTGC TTCATCATAT GTGATAGCAC ATATTTCTTG AATAATACGC ACTGAACGGT	6000
	CGATCAGTTT TTGATTGGTT GCTTTAACAT CAATCATGAG GTTATCGTAA ACTTTTCCGA	6060
25	CACCAACCAT TGTGATGGTT GAAATCATAT TTAAATTAAT CTTTGTGCT GTACCAGACT	6120
	TTAAACGTGT TGAACCAGTT AATACTTCTG GACCAACTTT AACTTCTACT GGATACTGCG	6180
	CAATTTCACT TATAACTGCA TGTTCATTGC ATGAAATAGA TACTGTTGTA GCACCGATTG	6240
30	TGTTAGCAAA TGTTAAACCG CCTATAACAT ATGGCGTTTT GCCACTCGCG GCAATTCCTA	6300
	TAACGACATC TTTTGATGTT AAATCTATAT TTTTCAAATC TTCTTCCGCT AATTTTTTGT	6360
35	GATCTTCCGC ACCTTCTACA GCCATCGTCA TAGCATGTTG TCCACCAGCA ATAATACCTA	6420
	TAATTTCAATG AGGGTCAGTA TTGAATGTAG GTACACACTC CGCTGCATCT AAGACACCCA	6480
	ACCTTCCACT TGTACCTGCA CCGATATAAA TCAATCGTCC ACCCTTTTTTA TACTGTGCAA	6540
40	TTGTTTTTTTT AATTACTTTT GTCAATTGTG GTATTGCCTT TCGAACTGCT AACGGGACTT	6600
	GCTGATCTTC TTTATTCATC GTAATTAAAG CCTCTTCCAC AGTCATTTCA TCAAGATGCA	6660
	TCGTGCTTC ATTACGCGCT TCGGTCGTAC TATTTTCCAT CACTTCTTAC ACTCCCTAGT	6720
45	TTTTTGAAAA TCAAATGTAT CATTCGGCTC GATACAACTT AACAGTGGTA AGTCTTCTTT	6780
	AATAATTTGT GCAaCAACAT TCACATTGTC ATGTGCACTA AGCGTTTGTC TCACAATTTG	6840
	CATTTGCGCT TGATAACGTC CGTTATTCAA ATTATCAACG GTTACTGAAC CAATGCGTCG	6900
50	TTGCGTCGTA AACTGTGGTT GAATCGAATG TGGACATATT TGTCTTGACG TTTCCGAACG	6960
	AATGACATTT TCCGGATTAT CCGGGCGTAC TTTATGACAC ATATCGAAAA GGTAAGTCAC	7020

55

	AAGTTGTTTT GCCTGCCTCA TTTCAATCAA TGAGTCTCCA ACTAACACTT CAGATACACC	7140
	AGTTTCTTGT AATAATTTAG CTGCAACGAC AGGATGACTA TGTCTCGTTG CTTCAATTGT	7200
5	TGGCAAGCCT TTATGCAAAG GACCTCGCAA ATCACTCCCT ACAATAAAAC CATATATTTG	7260
	TGCCTTTGGA TTAAATTGAT AAATGAGTTC ATTTTCTTA TTGACCAAGT CAACAGATAA	7320
10	TCCCGTATCT GGTCTTGAT AATAGTTATG ACAAATGAA AGTAATGTAA AATCATTCAA	7380
	TTGTTGATGT AAGCTTGTTA ACAATTCCCG GGAAATAATA CTTGCATTCA AACAGCACTT	7440
	TAAACCCTGT GCCATTATCG CTTGATTGTC CTCAATTGAT GTACTATGAT CGATACGAAT	7500
15	CATAAATTGT GCATCATATT GTCGAAGATG GTCATAAAAA GATGGTGTTA AAATAGATGG	7560
	ATTAGCATCT ATGAGGTAAG TCACTTGTTT ATGTTTTAAT AAATTGAGTA GTTTTGTGAA	7620
	ATAATGATAT TTTGTCTCGT CATCTTCTTC TGGTATTTGT ACAGATGTAA AAATCATTTG	7680
20	GTAACCTTGT TTAATCATTG GCTTAATATA CGCTTCATCT AAAGGTTGTC CTAAATACAC	7740
	TGAAAAGCCT GTCAAAGTAG CCCTCCTTAA CAATATAATT ATTAGGAAAA TATAGTTGAT	7800
	TTGTGTAATC GCTTACATTT TACTATAAGA GAAAACACAT TACAATATTA ATCAGTTAAA	7860
25	GCCTGTTTCA TGTAATAATC TTACATATTT CTGTCACAAG TTAATTATTA CACCATCAAA	7920
	GATTATCCTT TCTTTTAAGT GCTGATAATA GCTGCTACTG CTGGATTATT ACAATAACTT	7980
30	TTATACATTT TATTCAGGAT TATCTTATAT TATGTTTTAA TAATAATCTG TGAACAATTA	8040
	AGAGATTTGA AATTGAATTT AATAATTGTA TTGAAAACGC ATACTTCACC ATGCTAAAAT	8100
	AGGAGTCGCA AACAAATAAG ATTCAATAAG ATGTGATGGT TACCAACACA GTCTATTTGC	8160
35	TCGTGCTTTT TTTTATTGAA TCTTAAATAA TAAATACAAC TTTGGAGGTT GGACAAGTGA	8220
	GGAAGAACT TTTTCGGTCAA TTGCAACGTA TTGGTAAAGC GCTAATGTTA CCTGTTGCGA	8280
	TTTTACCAGC AGCTGGTCTG TTATTAGCTA TCGGTACAGC TATGCAAGGT GAATCATTAC	8340
40	AACACTACTT GCCGTTTATA CAAAATGGTG GCGTACAAAC TGTGCTAAA TTAATGACAG	8400
	GTGCTGGTGG TATCATTTTT GATAACTTGC CTATGATTTT CGCATTAGGT GTCGCAATCG	8460
	GATTAGCTGG CGGTGATGGC GTAGcAGCTA TCGCAGCATT CGTCGGTTAC ATAATCATGA	8520
45	ACAAAACAAT GGGCGACTTT TTACAAGTTA CACCTAAGAA TATTGGTGAT CCAGCGAGTG	8580
	GTTACGCTAG CATTTTAGGT ATCCCAACAT TACAAACAGG TGTGTTCCGC GGTATTATAA	8640
50	TCGGGGCCCT GGCAGCTTGG TGTATAACA AGTTCTATAA CATTAACTTA CCATCTTATT	8700
	TAGGTTTCTT CGCTGGTAAG CGTTTCGTAC CTATTATGAT GGCTACAACA TCATTTATTT	8760
	TAGCATTCCC AATGGCATT AATTGGCCAA CGATTCAATC AGGATTAAAT GCATTGAGTA	8820

55

	TATTAATTCC ATTCCGGTCTA CATCACATTT TCCACGCACC GTTCTGGTTC GAGTTTGGTT	8940
	CATGGAAAAA TGCAGCTGGT GAAATTATTC ACGGTGACCA ACGTATCTTT ATCGAACAAA	9000
5	TTCGTGAAGG CGCACATTTG ACAGCTGGTA AATTCATGCA AGGTGAATTC CCTGTTATGA	9060
	TGTTCCGTTTT ACCTGCAGCA GCTTTAGCAA TTTATCACAC AGCTAAACCT GAAAATAAGA	9120
	AAGTAGTAGC AGGTTTAAATG GGTTCCTGCTG CTTTAACATC ATTCTTAACT GGTATTACAG	9180
10	AACCATTAGA ATTCTCATTG TTATTTGTAG CACCATTATT ATTCTTTATT CACGCaGTAC	9240
	TTGATGGTTT ATCATTCTTA ACATTGTAAT TATTAGATCT TCATCTAGGT TATACATTCT	9300
15	CAGGTGGTTT CATCGACTAC TTCTTACTCG GTATACTACC TAATAAGACA CAATGGTGGT	9360
	TAGTCATTCC TGTAGGTCTT GTATACGCAG TTATTTACTA CTTTCGTATTC CGATTCTTAA	9420
	TTGTAAAATT AAAATACAAA ACACCAGGTC GTGAAGATAA ACAATCACAA GCGGCTACTG	9480
20	CTTCAGCAAC TGAATTACCA TATGCAGTAT TAGAAGCTAT GGGTGGCAAA GCAAACATTA	9540
	AACATTTAGA CGCTTGTATC ACACGTCTAC GTGTTGAAGT TAACGACAAA TCTAAAGTTG	9600
	ATGTTCCCTGG TTTGAAAGAT TTAGGCGCAT CTGGTGTATT AGAAGTCGGC AATAATATGC	9660
25	AAGCAATTTT TGGTCCTAAA TCTGACCAAA TCAAACATGA AATGCAACAG ATTATGAATG	9720
	GTCAAGTAGT AGAAAATCCT ACTACTATGG AAGACGATAA AGACGAACT GTTGTGTTG	9780
	CAGAAGATAA ATCTGCAACA AGCGAATTGA GCCATATCGT GCATGCACCA TTAAGTGGTG	9840
30	AAGTAACACC ATTATCAGAA GTGCCTGATC AAGTGTTTCA CGAAAAATG ATGGGTGACG	9900
	GTATCGCTAT CAAACCTTCA CAAGGTGAAG TTCGTGCACC ATTCAACGGT AAAGTACAAA	9960
35	TGATTTTCCC AACAAAACAT GCAATTGGTC TTGTATCAGA TAGTGGTTTA GAACTATTAA	10020
	TCCACATCGG TTTAGACACT GTTAAATTAA ACGGAGAAGG CTTTACTTTA CATGTTGAGG	10080
	AAGGTCAAGA AGTTAAACAA GGTGATTTAT TAATCAACTT TGATTTAGAC TACATCCGCA	10140
40	ATCATGCAAA GAGTGATATT ACGCCTATTA TCGTGACACA AGGAAACATT ACAAACCTTG	10200
	ATTTTAAACA AGGTGAACAT GGCAACATTT CATTTGGCGA TCAATTATTT GAAGCTAAAT	10260
	AATGCTTACT ATAAACAGGT GCGTATACCT TCATAAGGTG ACGCGCCTGT TTTTCTTTG	10320
45	CTATTGTATT TTGCAGCATC ATTGATAGTT CGCTCTCCCC TTAAATTTTG AATTTTAAGA	10380
	TCATCAATTA AAGCCCCCTC TCATACTCAT TTCCTAAAAA ATATTAATTG TTCATTATG	10440
	TTAGCGTTTT CACAACAAAG TCAACTTCCT TGACCTTACA CTATATTCGA GGCTATCATT	10500
50	TTAAGTGTA ATATAGAGAA AAGGTGGCTT TTTTATGAA ACAACGCATT GGAGCTTACT	10560
	TAATTGACGC TATTCATCGA GCAGGCGTCG ATAAAATTTT TGGTGTTTCT GGTGATTTTA	10620

55

	ATGAATTAAA CGCAAGTTAC GCAGCGGACG GTTATGCCCG TCTTAATGGA CTCGCTGCAT	10740
	TAGTTACTAC ATTTGGTGTT GCGGAATTAA GTGCCGTCAA CGGTATCGCA GGTTCATATG	10800
5	CTGAACGCAT ACCTGTCATT GCGATTACAG GTGCGCCGAC ACGTGCTGTT GAACAAGGCG	10860
	GTAAATATGT ACATCACTCA CTTGGTGAAG GTACATTGA CCACTATCGA AAAATGTTTG	10920
	CACATATAAC CGTTGCACAA GGTATATCA CACCTGAAAA TGCAACAACC GAAATACCAC	10980
10	GTTTAATTAA TACAGCAATC GCCGAAAGAC GCCCAGTTCA TTTACATTTA CCAATCGATG	11040
	TCGCAATCTC TGAAATTGAG ATACCGACAC CATTTGAAGT GACGGCAACT AAATATACGG	11100
	ATGCATCAAC ATATATAGAG TTATTAGCAA CTAACTGCA TCAAGCGAAG CAGCCTATCA	11160
15	TCATTACTGG ACATGAAATT AACAGTTTTT ACCTCCATCA AGAATTAGAA GATTTTGTA	11220
	ATCAAACACA GATACCACTA GCACAACCTT CATTAGGAAA AGGTGCTTTT AATGAGGAAA	11280
20	ATCCATATTA TATGGGTATT TACGATGGGA AAATTGCCGA AGATAAAATA CGAGATTATG	11340
	TGGACAACAG CGATTTAATT TTAAATATTG GAGCCAAATT AACAGATTCA GCAACAGCAG	11400
	GTTTTTCATA CCAATTCAAT ATCGATGATG TCGTTATGTT AAATCATCAC AATATCAAAA	11460
25	TTGACGATGT TACAAATGAT GAAATATCTC TACCATCATT GTTAAACAG TTATCCAATA	11520
	TTTCATATAC GAATAACGCA ACGTCCCTG CGTATCATCG TCCAACATCA CCCGATTATA	11580
	CTGTTGGCAC AGAACCATTA ACACAACAAA CTTATTTTAA AATGATGCAA AATTTCTTAA	11640
30	AACCAAATGA TGTATCATT GCTGATCAAG GTACATCATT CTTTGGTGCT TATGATTTAG	11700
	CATTATACAA AAACAATACT TTTATAGGGC AACCGTTATG GGGTTCTATC GGCTATACAT	11760
	TACCTGCAAC ATTAGGTTCA CAATTAGCAG ACAAGATCG TCGTAACTTA TTATTAATTG	11820
35	GTGATGGCTC ATTGCAACTA ACTGTTCAAG CTATTTCAAC TATGATTAGA CAGCATATTA	11880
	AACCGGTATT ATTTGTGATT AATAATGACG GCTATACGGT AGAACGACTT ATTCACGGCA	11940
40	TGTATGAACC TTATAATGAA ATTCACATGT GGGATTATAA AGCTTTACCA GCTGTATTTG	12000
	GTGGTAAAAA TGTTGAAATT CATGACGTTG AATCATCAAA AGATTTACAA GACACGTTTA	12060
	ATGCAATTAA TGGTCATCCC GATGTGATGC ATTTGTGCGA AGTCAAAATG GCTGTGGAAG	12120
45	aCGCACCGAA GAAACTCATC GATAcCGCTA AAGCTTTTTT ACAACAAAAT AAATAATTTT	12180
	ATCGTATACA GGGTATAAGT TTAAGCGAAT ACTTTATTAA ACGAATAGGA CTCTGATATA	12240
	AGATGATTAA TTTTAATAAA ACCGCTTTAG TGTTAATCGA CCTGCAAGAA GGTATTCTTA	12300
50	AAATGGATTa TGCCCCATAT ACAGCTGAAA ATGTCGTTCA AAACGCTAAT AAATTAATAG	12360
	ATGTTTTTAG AAAAAACAAT GGCTTTATCG CTTTGTTCG CGTGAATTC TATGATGGTA	12420

55

	AGTCGTTTCC ATCATTATT AGACAAGAGA GATGACGATT TTGTCATAGA CAAACGACAT	12540
	TTTAGTGCAT TTGTAGGAAC AGATTGGGAC TTACAATTGC GACGTCGAGG AATTGATACG	12600
5	ATTGTTCTTG GTGGTGTGCG AACGCATATT GGCGTAGATA CGACAGCGCG AGATGCCTAT	12660
	CAATTAACT ACAATCAGTT TTTTGTACATA GATATGATGA GTGCACAAA CGAAACGCTA	12720
	CATCAATTC CAATAGATAA TGTATTCCCA TTGATGGGAC AAACAATAAC TACAAACGAC	12780
10	TTTCTAAATA TATTGAACTA AACATATACT TCCCCCTTC GATCATGTTG AGGGGGATCT	12840
	TTATTTTACA AAGTATTAAT ACGTCGGGT GTCTAACCTT CTATATTAA CATATTCTAT	12900
	ATCTGTTAAA TCGTTCTTAA CTTACGCCCC TACTACATAA AAAACAGTAT TTATTCCGGA	12960
15	ATTTTCAAAA AATTTAGTAT TTATTGCAAA ATTATGTATC ACTTTATGTT TAATTTTGA	13020
	TATTATCTTA ATTAAGTAGA TTTTATAAG TTCTAAAAAG GAGAACAAAT ACATATATGA	13080
20	AGAAGAACT AACATTTAAA GAAAACATGT TTATAGGTTT TATGTTATTT GGTTTATCT	13140
	TTGGTGCCCG CAATCTTATC TTCCAATAC ACTTGGGTCA AGCTGCTGGT TCTAACGTTT	13200
	TTATCGCTAA CTTAGGATTT TTAATTACAG CAATTGGCTT ACCATTTCTA GGTATCATTG	13260
25	CTATTGGCAT TTCAAAGACA TCTGGTTTAT TTGAAATTGC ATCGCGTGT AATAAAACAT	13320
	ATGCTTACAT TTTCACGATT GCCTTATATC TAGTTATCGG ACCATTTTTC GCCTTACCTA	13380
	GACTGGCAAC GACATCATTT GAAATTGCAT TTTCGCCATT TTTATCACC AAGCAAATCA	13440
30	CTTTATATTT ATTTATTTT AGCTTCGTCT TCTTTGTGAT TGCATGGTTT TTTGCGAGAA	13500
	AGCCATCAAG AATTTTAGAA TATATCGGTA AATTTTAAA TCCGGTATTC TTAGTATTAT	13560
	TAGCAATTAT TTTATTATTT GCTTTTATCC ATCCATTAGG TGGCATATCT GATGCACCTA	13620
35	TTAGTAAACA ATATCAATCA CATGCCTTAT TTAACGGCTT TTTAGATGGA TACAATACCT	13680
	TAGATGCGCT AGCGTCATTG GCATTTGGTA TTATCATGTG TGCAACGATT AAAAAGTTAG	13740
40	GTATCGAAAA TCCAACGTAT ATCGCTAAAG AAACAATTAA GTCTGGTACT ATCAGTATCA	13800
	TTATGATGGG GATCATTTAT ACCCTACTAG CAATCATGGG TACATTAAGT ATTGGTCATT	13860
	TCAAACCTAG TGAAATGGT GGTATTGCCT TAGCGCAAAT TACTCAATAC TACTTAGGTA	13920
45	ACTACGGTAT CGTCCTGTTG TCACTTATCG TTATGGTTGC TTGTTTAAAA ACAGCCATCG	13980
	GTTTGATTAC GGCATTTTCA GAAACATTG AACACCTTT CCCTAAAATG AATTACCTAG	14040
	CGATTGCAAC AGTTGTAAGC TTTATTTGCT TCTTATTCGC GAATGTTGGT TTAAC TAAGA	14100
50	TTATTATGTA CTCAGTCCCA GTGTTAATGT TCTTATATCC ATTAGCAATT GCCTTGATTG	14160
	TACTAACATT ATTTAGTAGC AAATTCCATC ATTCAAACT TATTTATCAA TGTACCATT	14220

55

GCACATCATT CTCACAAACT TTGATTAATT TCAGCCAAAA ATATTTACCA TTATCAGACA 14340
 TTGGTATGGG CTGGGTTGTT CTCAGTTTGA TTGGTTTCAT TATCGGCTTC ATTATTTATA 14400
 5 AAATTAAGCA TCGTAAAATT CCACAAGCAT AATACTATGC CACAGTCATA TGT'TAAACAT 14460
 ATGCTTGTGG CATTTTTTAT TCATACTACA TTAAACTGCA ATCGTATACA TACATATCAA 14520
 TGATTATCCA CAAAAAATAT TAGTACTTTC ATTTTACAAA TCACATTAAT ACAAACACAA 14580
 10 CCTTATCTTT ATATTATTAA ATTTATATTT GACACTTATA TTGAACAACT GTAATATATT 14640
 AATATTAATT CTTTAAAATG TATAAATATA AAGGAGGGAG ACCGATGaAT TCAATCATTG 14700
 15 aATTAAGTGA TTATTATAGC TCTAATAATT ATGCACCACT TAAGCTTGTC ATTTCTAAAG 14760
 GTAAAGGTGT CAAAGTTTGG GATACTGATG GCAAACAATA TATAGATTGC ATTTCGGGTT 14820
 TTTCAAGTGC AAACCAAGGC CATTGTCATC CAACAATTGT TAAAGCGATG ACAGAACAAG 14880
 20 CTTCAAAGTT GTCTATCATT TCACGTGTCC TTTATAGTGA CAATCTCGGG AAATGGGAAG 14940
 AAAAAATTTG TCATCTTGCT AAGAAAGACA AAGTACTCCC CCTTAACTCT GGTACTGAAG 15000
 CTGTTGAAGC AGCCATTAAA ATTGCTAGAA AATGGGGCTC TGAAGTTAAA GGCATTACTG 15060
 25 ACGGACAAGT TGAAATCATC GCTATGAATA ACAATTTTCA CGGTCGTACA CTTGGCTCAT 15120
 TATCACTATC TAACCACGAC GCATATAAAG CAGGATTTCa CCCCCTACTT CAAGGCACTA 15180
 CAACAGTAGA TTTTGGAGAC ATTGAACAAT TAACACAAGC TATTTCAACG AATACAGCAG 15240
 30 CAATTATTTT GGAACCAATT CAAGGTGAAG GTGGCGTTAA TATACCACCG AAAGGATATA 15300
 TTCAAGCTGT GCGTCAACTA TGTGATAAAC ATCAAATATT ATTGATTGCA GATGAAATTC 15360
 AAGTTGGTCT TGGTAGAACT GGGAAATGGT TTGCTATGGA ATGGGAGCAA GTCGTTCAG 15420
 35 ACATTTATAT TTTAGGTAAG GCATTGGGTG GCGGCTTATA CCCTGTATCT GCTGTACTTG 15480
 CAAATAATGA TGTATGCGT GTTCTAACAC CAGGTACACA TGGTTCAACA TTTGGTGGA 15540
 ACCCTTTAGC CATTGCAATA TCGACGGCAG CGCTTGATGT ACTTAAAGAT GAACAACTGG 15600
 40 TTGAACGATC AGAACGCTTA GGTTCATTTT TATTAAAAGC GTTGCTACAA CTTAAACATC 15660
 CTAGTATTAA AGAAATTAGA GGTGCTGGTT TATTTATAGG CATAGAGCTT AACACAGATG 15720
 45 CTGCACCTTT TGTGGATCAA CTGATTCAAC GTGGAATCTT ATGCAAAGAC ACGCATCGTA 15780
 CTATCATTCG ATTGTCTCCA CCTCTAGTCA TTGATAAAGA GGAAATCCAT CAAATTGTTG 15840
 CAGCTTTTCA AGACGTTTTT AAAAATTAAC AATTAATCAT TTATATATGA CATAGGAGGG 15900
 50 ATTCATGATG ATTAAAGTAG GTATCGTTGG CGGTAgcGGT TATGGCGCAA TTGAATTAAT 15960
 TCGATTGTTA CAAACACATC CTCATGTAAC GATTGCACAC ATCTACTCAC ATTCAAAAGT 16020

55

ACTTACAGTG GaTAATAATG ACTGTGATGT AATTTTCTTT GCGACACCAG CACCCGTAAG 16140
 TAAACATGT ATCCCTCCCT TAGTAGAAAA AGGTATTCAT GTTATCGATT TATCTGGCGC 16200
 5 ATTTAGAATT AAGAATCGTG AAATATATGA AGCATATTAC AAAGAACTG CTGCAGCACA 16260
 AGATGATTTG AATCATGCTA TTTACAGCAT TTCAGAATGG CAATCGTTTG ATAACAATGG 16320
 AACGAAGCTC ATTTCTAATC CTGGCTGTTT CCCTACAGCA ACATTATTAG CATTACATCC 16380
 10 ACTTATTAGC GAAAAAATAG TAGATTTGTC ATCTATTATT ATTGATGCTA AGACCGGCGT 16440
 GTCAGGTGCT GGTCTTCAT TATCACAACG gTTcATTtTT CAGAAATGAA TGAAAATCTA 16500
 15 AGCGCTTATG CAATCGGAAA CCATAAACAC AAACCGGAAA TCGAGCAATA TTTATCTATC 16560
 ATTGCGGGTC AAGATGTATC AGTCATATTT ACACCACATC TCGTACCAAT GACACGAGGT 16620
 ATTTTATCAA CAATATATGT CAAATTATCA TCTGAATATA CGACTGAATC ATTACATAAA 16680
 20 TTAATGACCT CTTATTATGC TAATCAGCCA TTTGTGAGAA TTAGAGATAT TGGGACTTTT 16740
 CCAACCACAA AAGAAGTACT CGGTAGTAAC TACTGCGATA TCGGCATCTA TGTAGATGAA 16800
 ACAACGCAA CAGCAATTTT AGTATCAGTG ATTGATAACC TTGTCAAAGG CGCAAGTGGG 16860
 25 CAAGCCATTC AAAATTTAAA TATATTATAT GATTTTGAAG TGACGACTGG CCTAAATCAA 16920
 TCACCAGTTT ATCCATAAGG GGTGTTAGAA TGAAACATCA AGAAACGACA TCACAACAAT 16980
 ATAACTTTTC AATTATTAAA CATGGCGATA TCAGTACACC TCAAGGCTTC ACGGCTGGTG 17040
 30 GTATGCACAT CGGTTTACGC GCTAACAAAA AAGACTTTGG GTGGATTTAC TCATCGTCTT 17100
 TGGCAAGTGC AGCTGCCGTA TATACTTTAA ATCAGTTTAA AGCTGCACCA CTTATTGTCA 17160
 CTGAAGACAC TTTACAAAAG TCTAAAGGAA AATTACAAGC ACTTGTTGTT AATTCAGCTA 17220
 35 ATGCAAATTC TTGTACCGGT CAACAAGGCA TAGATGATGC ACGACAAACA CAAACATGGG 17280
 TTGCTCAACA ACTTCAAATA CCATCTGAGC ATGTTGCTGT TGCTTCAACT GGGGTCATTG 17340
 40 GTGAATATTT GCCTATGGAT AAAATTAAGA CTGGGACCGA ACATATTAAG GATGCTAATT 17400
 TTGCAACGCC AGGTGCGTTT AACGAGGCAA TTTTAACAAC TGATACCTGT ACAAACATA 17460
 TCGCTGTATC ACTAAAAATC GATGGTAAAA CCGTTACAAT TGGTGGTAGC ACCAAAGGTT 17520
 45 CAGGTATGAT TCACCCAAAT ATGGCTACCA TGCTTGCTTT TATAACAACC GATGCATCGA 17580
 TTGAATCGAA TACACTTCAT CAATTATTAA AATCTTCGAC TGACCATACA TTTAATATGA 17640
 TTAAGTTTGA TGGCGATACA AGTACAAATG ACATGGTATT AGTCATGGCA AATCACCAG 17700
 50 TTGAACACCA AATACTTAGT CAAGACCATC CACAATGGGA AACATTGTGT GATGCATTCA 17760
 ATTTTGTCTG TACATTTTAA GCTAAAGCTA TAGCCAGAGA TGGCGAAGGC GCAACAAAGT 17820

55

	CTATCGTAAG TTCAAATCTA GTAAAATCAG CTATTTTTGG CGAAGATGCC AATTTTGGTC	17940
	GAATCATTAC AGCTATTGGC TACAGCGGAT GTGAAATTGA TCCTAACTGC ACATATGTTT	18000
5	AACTGAACCA AATACCTGTC GTTGATAAAG GTATGGCTGT ACTATTTGAT GAGCAAGCTA	18060
	TGTCGAATAC ATTAACATCAT GAAAATGTCA CAATTGACGT TCAGCTTGGT TTAGGTAACG	18120
10	CTGCAGCGAC TGCATACGGT TGTGATTTAT CCTATGATTA TGTGCGTATC AACGCATCAT	18180
	ATCGAACATA AGGTGGTGTT GGTTAGATGA AATTTATTGT CATTAAAATT GGTGGCAGTA	18240
	CACTTAGTGA CATGCATCCA TCAATTATTA ACAACATTAA GCATTTACGA TCAAACAACA	18300
15	TCTACCCCAT TATCGTTCAT GCGGGTGGCC CATTATTATA TGAAGCATTG TCAAAACCAGC	18360
	AAATCGAGCC ACACTTTGTT AATGGCCTAA GAGTGACTGA TAAAGCAACC ATGACCATTA	18420
	CTAAACACAC GCTCATTGCA GACGTTAACA CTGCATTAGT AGCTCAATTT AACCAGCACC	18480
20	AATGTTCTGC AATAGGCTTA TGTGGTTTGG ATGCACAGCT GTTTGAAATT ACATCTTTTG	18540
	ATCAACAATA TGGATATGTC GGTGTTCCGA CCGCTTTAAA TAAGGATGCT TTACAGTATT	18600
	TATGTACTAA ATTTGTACCT ATCATCAATT CGATTGGTTT CAATAACCAT GATGGAGAAT	18660
25	TTTACAATAT TAATGCTGAC ACGCTTGCCT ATTTTATTGC ATCATCATTG AAAGCGCCTA	18720
	TTTATGTATT AAGTAATATT GCAGGTGTAC TCATCAATGA TGTGTTTATA CCTCAATTGC	18780
	CATTAGTCGA TATTCATCAA TATATTGAAC ATGGTGATAT TTATGGAGGT ATGATCCCA	18840
30	AAGTGCTAGA TGCCAAAAAT GCGATTGAAA ATGGCTGTCC TAAAGTTATC ATTGCATCAG	18900
	GAAACAAGCC AAATATCATT GAATCTATTT ACAATAATGA TTTTGTTGGC ACAACAATCC	18960
	TTAATTCATA ACTATGAAAT TAAGGCCTAA CAAGTTTTGA CACGCGAGAT GATTCCAGTT	19020
35	CGATTATCCA TTGCGCTAAA ACATTTATTT ACCGTTTCATC TCGTTAACAA TTTTGAATAC	19080
	AGTACGATAC AATATGAGAT GTAAAAAACT AATAACCTTT TACAAATTTG TTTATCAAAA	19140
40	TATTTTAAGT TTGCAAAGC TTTTATTGT GATTATTTTC ACAAATACT ATAATGAGGA	19200
	TAGTAAATAG AGAGGAGTCC TTAAGTTGAC GAAACGACAA ATGGGTATAT TCATTTATGC	19260
	TGGAATTATC GGTGGCTTGT TATCTGGAAT TGTAATAATTA GGTGGGAGG TCATGTTTCC	19320
45	ACCTCGCACA CCAGAACGTA ATGCAACGAA CCCACCTCAA GAGTTATTGC AACAATTAGG	19380
	ATTTAGTAGT GAGTTTACGC ATCAAACATA TACATTTTCA AATATGGAAT TGCCTTGGGT	19440
	AAGCTTTATT GTCCACTTTA GTTTTTCTAT CGTCATTGCA ATTATTTACT GCATATTAGT	19500
50	TAAAAAATAC GCTTACTTAG CAATGGGACA AGGTGCTGTT TTTGGTATTG CTATTTGGGT	19560
	ATTATTCCAC CTTATCATTG TGCCAATCAT GCATACTGTA CCTGCTGTGT GGGATCAACC	19620

55

	AGTGCGACAA	CATTTTGTCT	ATCGCTATAA	ATTAAATTAA	TACACTGACT	AACATTAACG	19740
	TGAGTTTTAA	ATCATCGTTT	GAGTATGATG	ATTGATGCTC	ACGTTATTTT	ATTAAGTGAC	19800
5	ATGATATGAT	TCCAGCCAAC	TTACGTGAGC	ATTAAAGTCT	CAAATGCGTC	GTAACAAACT	19860
	ATTATTTTCG	GTAATTTCAA	TATTGCTCAG	TATATTTTAA	CCTTATCACT	TACTTTAATC	19920
	TCGTCATGAT	TTGAATGAT	GCCATCGTGT	ATTCACCTTT	CATTTTCCA	ATAAAAAAAC	19980
10	ATCTAACAGT	AAACATTTAG	GCAGTATAGT	TTAAAATCAC	TGCGCAATGA	TACTGTCAGA	20040
	CGTCATATTA	ACTACTCAAT	AACTGAAATA	CAGACACTTT	TTTATAACCC	CAGGGTGCCT	20100
	GTCCTAAGAA	ACATACCTGT	ACCATAAACT	GATCAAAAAT	AAATTGTTTG	AACTTCACTT	20160
15	CACGTGATTG	ATAAAAGTGT	GATTGTGTCA	TATCATAAAT	GTCCAATCCT	TTGATTAAAC	20220
	CTTCACCAAT	CAATTTTGTA	AAACTTTCTT	TTTGTGTCCA	TATTTGATAA	AAATCATTTA	20280
20	AACTACATAT	TTGATGTGCT	TCGTTTGTAG	AGAAACACGT	CACTAACGTA	CGCCAGTCTA	20340
	AACGTTGTGA	TATCTTTTCG	ATATCAATAC	CAACTGGTTC	TTTATCGACA	ACACACACGA	20400
	TATAAGGATA	ACTATATGAT	AAGCTCACAT	AGATGGGCTG	TCCATCACGA	TTGTGTTGAA	20460
25	CAATATCTGC	CTTACCTCGT	GGCGAAATGT	GATAATGCCA	TTGATGTGGT	AATAAACCTG	20520
	TGTCATGTTG	AATTCCATAT	TGCACTAAAA	TATCTCCCAA	TCTGTGCATG	AGTTTATCTT	20580
	GATTGTATCT	ATAGTTGACT	GTACGCGGTT	TTTTATATGA	CCAACGACTT	TGTGATATTA	20640
30	ATTCTTCAAT	ACTTTTCAAG	TACTCTGTGA	ATTGCATTAC	AAATACTGTC	ATAACTTTCC	20700
	CTACTTACTT	ATTGAATATT	GTTTTGATAT	ATTGTGCCCA	ATGATACAGC	CAATTGTTAG	20760
	TTATCGTTGG	CCATTTTTC	CTGATGTGAT	TCATTATTTT	TAATGTTAAT	GTTGTATCTA	20820
35	TCATTGCTAG	TTGTTGTTC	CGGTCAACAC	TAGTTAATCC	AATCGTTTCG	TACATGTCTT	20880
	GTTTCTGTAA	AATTTCAAT	AATGATTCAT	CGCTGACGAG	TTCAATTTCT	TTGCGCTTAA	20940
40	CGCATTCTAA	CAAAGATTT	ACCGGCATT	TATTAGGTGA	TAGCACATGG	TAAATGATTT	21000
	GTGGTGTGTT	GACCTGTGCT	AATGCGACAA	TTTGTCTTGC	AGTCGTATCC	ACAAAAGAAA	21060
	AATCTACAGG	CATTTAGCC	ATGCTAACCC	CGATACAATC	CAGTTGTAAC	AAATCATTCA	21120
45	TTACCATTGA	AAAACGGTTA	GTCTTTATAT	TTCTCATATG	CCATCTTCCA	TTGTAAGGAT	21180
	TCGTCAAATT	ACCAACACGT	ACAATCCGAC	CATCTAAGCC	ATTATTTACA	GCTTCTAATA	21240
	CTTTTAATTC	ACTATAAAAT	TTGCTCCGTG	TATATGGTGA	TGTTAGTAGT	TGCCCTTTAT	21300
50	AGACATCCGC	TTsTGAAAT	GTCACATCTT	CTGTGTCTAT	ATCAAAATAA	GTTCCACAC	21360
	TTATCGTAGA	CACATATATT	AACCTTGCAAT	GATGTTGTTG	TGCCAAACGT	ATGACATCAA	21420

55

CACCTGCATG AATAATCGTA TCCATGTTTT CTGGTAAAAC AACATCATCC ATACACTCGA 21540
AATCACCAAC AATGACTTCA ATGTTTGATA ACATTATTTT AACCGTCTCT TCTGAAAAAT 21600
5 AATCATTTAA ATTCGTCATC AACTTATACC ATGCTATTTT CTCATTATCA GCACGTATGA 21660
AACAATAAAT GCGATGACTG TATCCTTGTA GTACTTCAAT CAGATAAGCA CCTAAAAAAC 21720
CTGTCCGCC AGTCAATAGT GTATTTCTTA GAGGTGATG ACTTAGACTA TCCTCTAAAA 21780
10 TACCCAAGTT ATAACGAGAC ATAACAATCT TTTGTAATTC CGAAAGATTA TCCGGTAATG 21840
CAACTAATGA TTGTTGATTT TGGTACATAT AATTAACAAT CTGTCCGACG GTTTTATATT 21900
GGTATAATGT CTGCATTGAA ATATGATGGC CAAATCGTTT TAAATGCGAG ACAACTAACA 21960
15 TCGCCTCTAA TGAGTTACCA CCAAGTTCAA AGAAATCATC GTCAACACCG ACATCATTTT 22020
GTTTCAATAC CTCTCCAAAT ACATCAACAA ATGTCTGCTC AATTTTATTA GAGGGTTCGC 22080
20 TATACACTTT ATTAGACTGT TGTATAGGTG ATGGATTGG CAAACGCGTA GTATCCACCT 22140
TGTCATTGCT GGTAAATGGC ATACAATCGA TATGCGTTAT AGTCTTAGGA ATCATATACT 22200
TAGGCAGCTG ATCATTTAAA TATTGCTTCA AATCCTGTTT CACTTGTTGC TCTCCGACAT 22260
25 AATAAGCATT CAATATATCA TGCCTATCAA AGTGACTTAC TGTTACAACA CAATCAGATA 22320
TACCACGAAT AGCTAATATT GCATTTTCAA TTTCATCAAG TTCAATACGG TACCCGTTAA 22380
CTTTCACTTG TTTATCTATT CTTCTAAAA ATTCAATTTG ACCATCAGAT GTATAACGTG 22440
30 CTAAATCACC ACTATGATAC AACTTTCTTT TACCAAATGG ATTATTTTGC CATTTATCAG 22500
CCATTAATTC TGGACGATTA ATATATCCTA TCGCTAAACT ATCACCTGCA ATACACAAC 22560
CGCCTGGCAT ACCAATACCG CATAACAAAC CATCTGACAT AATATACACT TGGATGTTAG 22620
35 ATAAGGGTTT GCCAATTGGA ATCGTCTCAG GTATCAAATC ACCACAATGA TGTGACCAAT 22680
ACGATGTGAT GACTGTTGAC TCAGATGGTC CATAGGCATT GAAATACGTG CCACAATGCT 22740
TCTCAATATA TTTAACAAAG GATGCCGTAC TAGTTGCCCC GCCTGTAATC AACTTTTCAA 22800
40 TATAAAAGTC TTCCATAACA CTACACATCT GTAACGGAAT CGACGCAACC GTCACACGAT 22860
GCTTATTAAT GAGTTGTTGT AACTGTTCTG GATTAACACG TTCCTCTCTA TCTGGAATCA 22920
CAAGCGTATG ACCATTTAAC AAACAACAAT AAATCTCCAT AACTGATGCA TCAAAAACAA 22980
TATTTGCATG TTGCAAAAAT ACTTCATTGT CGCCTAATTG CAATTCAGTT GACCATGCAT 23040
GCACTAAATT CAACAAATTT CGTTGTCGTA TGGCAACCCC TTTAGGCATC CCGGTCGTAC 23100
50 CAGATGTGTA AATAGCATAC ATCTCATTAT CTAACATCGC TGTGTTTTCA AGTTGATTGC 23160
CATGTAAATC ATCATATTGT TCATTTTCCT TTGATTCAAC AAAGCCTTTA GCATTTTCCA 23220

55

TAGCATCCTC CAAAATTGCA CCTTGTCGTT TATTCGGAAA ATCAATATCG ATAGGTATAT 23340
 AAGATGCACC TACTTTAACT GTCGCCAACA TCGCCGCAAT CATTTCAAAA CTACGTTCTG 23400
 5 TAAACAAGGC AACCCGTTGA CCATTGCCCA CACCATTGTA TAGGAGCATG TCGCCAATGG 23460
 CATCCACATA GTTGCCTAAT GTTTCATACG TCATTGTCAA ATCATTTCATG ACTAGCGCAA 23520
 CATGATTACC TTGTCGTGAG ACAACTTCAT TAAAGTAACT TATGATAGAT TTATTTCCCG 23580
 10 GGACATTAAG CATTGATCG TTAACATGCG TATTGACCCA ATTTAGAAGT TCCTCCGTGC 23640
 CGTTTGGTAT ATCACAAAT TGTAGTGTAT CTTGATGCTT CAAAATATAA TCAATCATAA 23700
 TCATACATTG ATTACCCATG TGACGAACTG TTTCTGAGTG ATATAAATCG GTATTATACT 23760
 15 CGATATTGAT TGTATAGTCA TCGCGATCTT CTTCAATGAT GAAAGATAAA TCAAATTTTCG 23820
 CCGTCACTGA TTGGGTTGA ATGTGTGTTA ATTTACTATG CCCAAAATGA GCATGATTTCG 23880
 TTTCAATTGTT TTGTAGTACT AACATGACAT CAAATAATGG ATTCCGTGAG GCATCATGTG 23940
 20 ATTGATCTAA GTCATTTACT AAACATTCGA ATGGGTATTC TTGATGCTCG TATGCCTCCA 24000
 AACTCATTTT CTTAACCTCT TGTA AAACT GTGTCCACAT TTTATCAGGT GACGGTTGCC 24060
 25 CTCTATATAC CAACGTATTA GCAAACATGC CTAGCATTTG CTCCGTGCCT TTATGCATAC 24120
 GCGCACTCAT CACACTACCG ACAACAACAT CATCTTTTCG AGCATATCTA CTTAACAACG 24180
 TCATGACCAC ACTCATAAAG AACATAAAAT CAGTAATTTG ATGCTTTTCT ACATACTTTT 24240
 30 GAAGTAGCTG TCTCATTTGT TGATTCAATG TAAATGACAT CATTGCTCCA TTTGTCGTTT 24300
 TAATATTTGG TCTAACATAG TCTGTCGGTA AGCTTAAAT AGGTACTTCA TCTTTGAATT 24360
 GAGATAACCA ATATTGTCTA TGTTTCGTCA TATCAGGATG CGACATCCAC TCACTATAGT 24420
 35 CTTTATATTG CAATTTAAGT GGTAACAATA ATTTATGTTG ATAAAGTGCG TTAAGATCAT 24480
 TCAATTAAATG TATATTACTC ATACCGTCAT TAATGATATG ATGCGTATCT ATAAAGAGGT 24540
 ATGCATGTAA GGGACTTCTA ATGTATCTCA CTCTAATTTG ACTTGGCTTT TCCAAATTAA 24600
 40 AAGGTGCTAC AAATTGGCGC ATGATTTCTT GTTCATCCGT AAAATGCGTG TTAACCTCTT 24660
 CAAAGTCAAC TGCAACATCT GCCACAATAC GTTGTGGAAC CTCATCATCT ACAACAATAT 24720
 ATTGTGTTTCG TAAAATCTCA TGTCGCGCTA TCAAACGCTG CACTGCTTGT CGCAATTGAG 24780
 45 CTACATTAAG TTCTGATGAT AACCGCCATA AAAAAGGTAC GTTATACACC GTATCTTTAT 24840
 GGTTTGATTT CCATAATAAA TACATACGCT TTTGTGCAGA GCTCAGCACA TAATCATCTT 24900
 50 TAACTATAGT TTCTGGAATC ACTTCATAGT TTTGTTCTTG AACCTTAGCA ATCGCTTGTG 24960
 CTAGTTCAAA TACAGTTGGC TTTTGTAATA AATCACCAAT TTGTAATCGT TTCCAGTAG 25020

55

	AATTATCATG AATACCTACT TGATTCACAT GTAAAATATC TGCAAAAATT TGGCATAGCA	25140
	AGTGTTCCGGT ATCTGTACTC GGTGCTACAT AGGCATCCGT ATCGACATAG TCCATGATAG	25200
5	GCAATGCCTT CTTATCTAAT TTCCCATTA TAGTAATAGG AATTGCTCA ATATGCATGA	25260
	AATTAACCTGG TATCATGTAC TCCGGTAAGG TCATACGTAA TTGTGATTTA ATCTTATTAT	25320
	GTGATAATGT ATGCATCGCT TCATAATAAG CAACGATATA CTGATCTTGA TCATGATTTT	25380
10	GAACAATAAC AACTGCTTTA TTAATACCTT GTATACGCTC GAGCGCATGC TCAACCTCTG	25440
	ACAACTCAAT CCTAAACCCT CGAATCTTAA CTTGTTTGTC CTTTCGATAT AAATAATCTA	25500
	TGTTGCCATC GGGTAACAAA CGAACGATAT CACCATTCT ATACATCAGC TGaTTTATAT	25560
15	TTGAATCTTT GATAAATTTA TCTGCTGTCA ATTCTGGCTG ATTTAAATAA CCTGCAGCTA	25620
	ACCCAAAGCC ACTTGTTACAT AATTCTCCAG GAATACCAAC GCCACACCGA CGCTCGCCTT	25680
20	GCATGATATA AACATGAGTA CCCAGAATCG GTTTACCAAT AGGAATACGA TTTGGAACCTT	25740
	TGTTAGGTAT ATTATACGTC GTTGTAATG TTGTATTTTC AGTTGGTCCA TAACCATTAA	25800
	TAATTTGAGG ATGCTTCGGT TTTTGATTAA GCAAAATCCAC CCACTTAGCA TTCAATACCTT	25860
25	CTCCACCAAT TAATAAATAC TTTAACGGTA CCAATACTTC TATTCGTTCA CTAGCAATCT	25920
	GATTAAATAA TGAGGAGGTT AACCACATAG TATTAACGTC ATTTTCATTG ATTAATTGTT	25980
	CTACCGCTAT TGGATTTAAT AATTGTTCTT TTTTAGCAAC AATCAGCTTT CCACCATTGA	26040
30	GCAATGCACC ATATATTTCA AATGTTGCAG CATCAAAGGC TATAGTTCCT GATAACAAAA	26100
	TCGTCGTCTC TTCATTTAAT GGTACATAAT GATTTTGATG GACCAAGCGA ACAATACCTC	26160
	GGTGCGGAAT TAGTGTCCTT TTAGGGTTAC CAGTTGTCCC CGACGTGTAA ATAACATAAG	26220
35	CATGATCTTC TAACGTGTTA CATTAGAAA GATTATCAAT ATTTTCCAC GCTATCTTAT	26280
	TCAAATCAAT GTGATTAATA TTTTGTTTAC CATTTTCATA TAAAGCTTGG TACGTTATTA	26340
40	CAACTTTAGG CGTTACATCT TTTAAAATGT ACTCCTGACG ATCACTTGGA TAGTTCGGAT	26400
	CAATTGGCAC GTAAGCCCCA CCAGCTTTCA ACACACCTAT CATCGCTATT ATCATCTCAA	26460
	TACTTTTTTC AGCTATGACA GCGACACGAT CATTAGGTTT AACACCATAC TGGTTTCTCA	26520
45	AACGGTGTGC TAAATCATTC GCGCGTGCAT TCAATGTTTG ATATGTTATA AACaCTCCGT	26580
	CAAAATTGCAC AGCGACATGA TTCGGCGTTG CTTCAACTTG TTGCTCAAAT AAGGTAACAA	26640
	CTGTTTGCGC ATCATCTATC TCAGGCAAAC TTAAATTGAT ATCGTCATAT AATTGAATAT	26700
50	CACGTTCTGT CATCAAATTA AGTTCATCTA CAGTTGTTTG TTTATTTCCA TTTTCTTCAG	26760
	TAATTTGCAA ATAAATATTT CGAACTAAGT CACTCAGCGT CTCGATTGAG AGCAAATCAT	26820

55

	CTAAAGATGT ATGTGCATCA TGTATTTGAT GTACATCCTC AATAACATCA TTACAACAAG	26940
5	ACATCATATG ATGATAACAA TGAAAAATAG TCTCTAGTGA AAGCGAAGAC TTCGCGCACT	27000
	GTAATTGCGA CATATTTTGC AACACACATT TATTAAAATC TGTTGTAAAA CGTTGACATA	27060
	CATCTTTTGC ATCGATTGTT AACGTTAACG GCACAATATT TCCGTGTAAA TCATTTGGTA	27120
10	AATGTGATGG TACATGTATA CCTAATGTGA CATCATGTTG TTGACTCATT ATATGATTAG	27180
	CTAAATACAC ACTAACAGCC AACGATGCCA TATCTATTGA TGTCATATCA TCAATCAAAT	27240
	ACGTTTGATA TAAAGCTTGT TCAAATGGAT GCTTAATTGG AAAATAACTA TCAATATGGA	27300
15	TGTCAGAGTT ATTCTCTAAC CGAAAAATAGT TTGAGTCTAA TGCTATATGC GATGCATCTT	27360
	GATTGTCTTT ATCATCATTT CTATTTATAT GTGCATGCTG TCGAGTATTG TTAATAACAG	27420
	TATTGCCACG ATATGCATTG CATAAATCAT CAAGAAaAAT ATCAATTTGA CTATCATCga	27480
20	AAATGGaCAC ATGAAATCT aTAGTATAT ATGcAGCATC AGCGAACTGm AACAAITTA	27540
	CTTTGAATAA AGGTGAATCA TTAAATGGT AAGTACTTAA TTCTTGCTTA AAAAAAGCTT	27600
	CTAAATCATA GTTTGCGGAA GAAGATGGAA CTTGTTTTAT CTCAATAAAA GGCAGAAATT	27660
25	CATGAAGTAT CATTGTtAAA TTGTCATCGG TAGTAACATC AAAAAATGT CTTATAGATG	27720
	CATGTTGTgC ACAATTGTCG ATAATGCATA CATCATTTTA GTAGCTTCAA CATTTTTAGC	27780
	GAGTTTAAAC CAATACGCAT TACGGTGTGT CGTTGATTCT GTATTATTTT TGTATATACG	27840
30	AAAATATTCC TGTTGAAATC TCAAATTACC CATAATCATA AAAAGTCCTT CTTTCATATC	27900
	ATAATACTCA TTACTTACTG AAATTGCATG ATGATATGAT AACCAGACGAA ATGTTAATTA	27960
35	ACTCGTTATG TAATGaTTAA TATaAAACAC CATTGCGAAC ATATGAGCGA TATATTCTAC	28020
	CCTAAAATAC ATCTTGATC ATCGTTACAA TTGGTATATT TTTCAATGTA AATTACATAC	28080
	ATCTTCGATA AATAGCACAC TACAAATCGT TAATCACTTT CTGTTGTTCA CATCTCATTG	28140
40	CAAACTCAAT ATTGTTGTTA CAAAATATCC ATGAAGCAAG TTTATATTAA ACAAACAACT	28200
	CGCATAAAAC AATTGTTATC CTTAAATTTT AACAAATTCT TAATAAATTT ATCTCTATTT	28260
	TAATTACGAC CAAATTAATA GGTMTTCCAT ATAAAAAGAT GCATAAAATA AATATTTAAA	28320
45	TAAATTCAAT TTGTATTCAC TTGTTTTTGT CCCCCAAATA CACCAGCAAC AAGCATGCTA	28380
	GCACCAATTG TTAAACGAT AACATATAC AGTCCCATTT GTAATGACGT TAAGAAAACA	28440
	CCCAACACAA TCCCTAACCT AGCTAGTGTT TCTGAAAAAT GAATACCTAA TGCATTAAC	28500
50	GCACTATATG TTCCTCTTTT AGCTTTAGGA ATAATTTTAA AGCGTTGTTT TGAAACTATA	28560
	GGCGAATAAA TAATTTCAAC TACAGTCGCA ATTATCATAA AAACAACATA TAAGCCAAAC	28620
55		

GCTTTTTTAA AATCTATTTT CAATACAACT TTCGAGATTG AATACGTGAG TAAATGACG 28740
 ACGACCGTAT TAATCATTAG CAAGATTGCT AACATCTTAG CACCTGTAAT ATCATATGAA 28800
 5 CCTATACTTA TTGTTTCAAA CTGATCCTTT AGTCTAATAG CAATATATGA GGAGATTGAA 28860
 AATTCACCCA TCATGATGAT ACTGAACCCC GAAATCAATA ACATATAATT ACGGTCTTTC 28920
 AAAACTAATT TATAACTGCG AAATATATTC ATTATTTGTA ATTTTGTATA ACGACTTGCA 28980
 10 TGCCTCTTGT CATCACTTTG CTTTACTTGA TTTTCGGTCTT GAGGTAACCA AATATATAAA 29040
 ATAAAGAGTA CAATTAAAAA TATACAAGCT GCTATTAAGA AAAGTAGTAA CATACTGTAG 29100
 CCATACATCA AGCCACCTAA CAATGCCCCA ATAGCTACCG ATAAGTTTGT CATCCAATAG 29160
 15 CTAATCTTGT AAATATAATG TTCCACGTCT TCGGTAATTG CATCCATAAT TAATGTGTCC 29220
 ATAACTGGAA ATTGTAATCC CCAAACGATT GTAAATATGG CATATGCAAC ACAAAAACCA 29280
 20 ATAATTTGCC ACAATTGATG TGACCCAAAT ACGCCCATGA ACACAAGCAT TATCACCATC 29340
 GTCGCTTGAT AAATAAGTAC TAGCAACTTT TCGGAAATA TCTCAATAAG GTAACCAGAT 29400
 ATAATGGACA ATGGAAATTT NAGAACCACT AAACCAACAA GATATATACC GACAATTGAT 29460
 25 TGACTTAACA TATCTGTAA ATATAGTGCT ATAAACGGTA TAAATGCTGT CGTAATAATT 29520
 AGCTGTAAAA NATTGCTAAT CAATCGTACT TTCAA 29555

(2) INFORMATION FOR SEQ ID NO: 207:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

40 AAAAAAAAAA AAAAAAGGTG AATCTTTAAT TAAACACTAA TATTGTAAAA GATGTTAAGT 60
 AAACGCTTAA TGACACTTAT TTTTGAAAA TAATAGTAAT ATCATTTTGT TAAATGAAAG 120
 AATAAGCTA TAATMATTAT AGAATAACTA TTAAAGGAG ATTATAACA TGCCAATTAT 180
 45 TACAGATGTT TACGCTCGCG AAGTCTTAGA CTCTCGTGGT AACCCAACTG TTGAAGTAGA 240
 AGTATTAACT GAAAGTGGCG CATTTGGTCG TGCATTAGTA CCATCAGGTG CTTCAACTGG 300
 TGAACACGAA GCTGTTGAAT TACGTGATGG AGACAAATCA CGTTATTTAG GTAAAGGTGT 360
 50 TACTAAAGCA GTTGAAAACG TTAATGAAAT CATCGCACCA GAAATTATTG AAGGTGAATT 420
 TTCAGTATTA GATCAAGTAT CTATTGATAA AATGATGATC GCATTAGACG GTA CTCCAAA 480

55

AGCTGACTTA TTAGGTCAAC CACTTTACAA ATATTTAGGT GGATTTAATG GTAAGCAGTT 600
 ACCAGTACCA ATGATGAACA TCGTTAATGG TGGTTCTCAC TCAGATGCTC CAATTGCATT 660
 5 CCAAGAATTC ATGATTTTAC CTGTAGGTGC TACAACGTTT AAAGAATCAT TACGTTGGGG 720
 TACTGAAATT TTCCACAAC TAAAATCAAT TTTAAGCAAA CGTGGTTTAG AACTGCAGT 780
 AGGTGACGAA GGTGGTTTCG CTCCTAAATT TGAAGGTACT GAAGATGCTG TTGAAACAAT 840
 10 TATCCAAGCA ATCGAAGCAG CTGGTTACAA ACCAGGTGAA GAAGTATTCT TAGGATTGTA 900
 CTGTGCATCA TCAGAATTCT ATGAAAATGG TGTATATGAC TACAGTAAGT TCGAAGGCGA 960
 ACACGGTGCA AAACGTACAG CTGCAGAACA AGTTGACTAC TTAGAACAAT TAGTAGACAA 1020
 15 ATATCCTATC ATTACAATTG AAGACGGTAT GGACGAAAAC GACTGGGATG GTTGAAACA 1080
 ACTTACAGAA CGTATCGGTG ACCGTGTACA ATTAGTAGGT GACGATTTAT TCGTAACAAA 1140
 CACTGAAATT TTAGCAAAAG GTATTGAAA CGGAATTGGT AACTCAATCT TAATTAAAGT 1200
 20 TAACCAAATC GGTACATTAA CTGAAACATT TGATGCAATC GAAATGGCTC AAAAAGCTGG 1260
 TTACACAGCA GTAGTTTCTC ACCGTTGAGg aACAGAAGA TACAACAATT GCTGATATTG 1320
 25 CTGTTGCTAC AAACGCTGGT CAAATTAAAA CTGGTTCAAT ATCACGTACT GACCgTATTG 1380
 CTAAATACAA TCAATTATTA CGTATCGAg TGAATTATTT GAAACTGCTA AATATGACGG 1440
 TATCAAATCA TTCTATAACT TAGATAAATA ATTTCTnTA TAATCAAATG CTGACATAAT 1500
 30 TTTAGTTGAG GATTATTATG ACGGTATAAA TAAATAAAG 1539

(2) INFORMATION FOR SEQ ID NO: 208:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 846 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

40 CAATTTCTAT CTATCAATGA TGTGCATACT TCCANTTAAA TTAAtCGAAA TGaATCAAGG 60
 TATATCATTC CTGCCTCTTT ATATAACaAC AAATAGTGAT TACAATATTT CGGTTATTAA 120
 45 CACGAAAATT TTACAAGCAC CTATTTCAAT TACATATATA TACAGCAAAA AAGAAAGCCC 180
 AGAAATATTG GTGTTTATTA AATCATTTAA AAAGTATATT GCCAATGAAC AATTATAATA 240
 50 AATTTCAAAT CTAAAAAACC AAGAATGCGA TTAATCATCA CATTCTTGGT TCAATTTTAT 300
 TCATGAATTT TTTCAACATT AAACGTTAAG TTATTGTCTG AATTTAAATT AACTTTAATC 360

55

CGTTGTACAA AACGTTTTAA TGGTCTTGCA CCGTATTGAG GTTCATAAGC TTCTTGACCT 480
 AGCCAAGCTT TAGCATCATC AGAAACTTCA ATTGAGATTC GTTGTTCTAA TAATCTTATA 540
 5 TTTAATTGCG TTAAGATTTT ATCTACAATC ATACTCATGT CATCAATAGA TAATGGTTTA 600
 AATAATACGA TATCATCCAT ACGATTCAAA ATTTCTGGTT TGAAATATGC ATTTAAACTT 660
 GTCATAACAG CTTTTTCTGT TGATTCTGTA ATTTACCAG TCTCTTTTAC GTTTTCTAAT 720
 10 AAAACTTGAG ATCCAATATT ACTTGTCATA ATAATAATAG TATTTTTTAA ATCAACGCTA 780
 CGTCCTTTAG AATCAGTTAA ACGGCCTCAT CTAAAATTTG CAATAATACA TTAAAGACGT 840
 CAGTAT 846

(2) INFORMATION FOR SEQ ID NO: 209:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1674 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

NTGGGAACAG TAAGCCAGTA TTTTGTAGAA GTTGCCATAC ATGAGCGTCG ATTTTTCCAA 60
 TATGGCTATG AACTAGAAC AATGGGAATT TGGAGGAAAA GTAAATGATT AAACCTAAAA 120
 30 TAGCATTAAAC CATTGCAGGT ACTGATcCaA CAGGTGGTGC CGGCGTAATG GCTGATTTAA 180
 AATCATTTCa TTCATGTGGT GTATATGGTA TGGGCGTCGT TACAAGTmTT GTTGCTCAAA 240
 ATACATTGGG CGTACAACAT ATTCATAATT TAAATCATCA ATGGGTAGAT GAACAACTTG 300
 35 ATAGTGTCTT CAATGATACC TTACCTCATG CTATTAAAAAC GGGGATGATT GCTACAGCAG 360
 ATACTATGGA AACGATTcGT CATTATTtAA TGCAACATGA ATCTATTCCA TATGTAATtG 420
 ATCCTGTTAT GTTGCGGAAA rCggtGATTc ATaATGGwTA ATGACaCAAg CaAAACTTGC 480
 40 AGCATAcGTT ATTGCCATTA GCTGACGTAG TAACACCGAA TTTACCAGAA GCTGAAGAAA 540
 TAACGGGACT AACCATTGAT AGTGAAGAAA AAATTATGCA GGCTGGCCGC ATCTTTATTA 600
 ATGAGATTGG TAGTAAAGGT GTCATCATTa AAGGCGGTCA TTCAAATGAT ACTGATATAG 660
 45 CAAAAGATTA TTTATTtACT AACGAAGGTG TTCAAACATT TGAAAATGAA CGATTtAAAA 720
 CAAnACATAC GCATGGAACA GGGTGTACAT TTTcAGCAGT TATAACGGCA GAACCTGCAA 780
 AAGGTAGACC ATTATTTGAG GCTGTACACA AGGCTAAAAA GTTTATTtCA ATGAGTATAC 840
 50 AATATACGCC TGAAATCGGC CGTGGTAGAG GTCCAGTGAA TCATTTTGCA TATTTAAAGA 900

TGTATACAA ACCATGTAGT TAAAAATTTT ACAGCGAATG GTTTATTAAG TATTGGTGCT 1020
 AGCCCTGCAA TGAGTGAAGC TCCCGAAGAA GCTGAAGAAT TTTACAAAGT TGCACAAGCG 1080
 5 CTATTAATCA ATATCGGTAC TTTAACAGCA GAAATGAAC AAGATATTAT TGCGATTGCT 1140
 CAAACGGCAA ATGAGGCAGG CTTACCTATT GTATTTGACC CTGTAGCTGT TGGTGCTTCT 1200
 ACATATCGAA AGCAATTTTG TAAATTATTA TTGAAATCAG CGAAAGTATC AGTAATTAAA 1260
 10 GGCAATGCAT CTGAAATATT AGCGTTGATT GATGATACAG CAACTATGAA AGGTACAGAT 1320
 AGTGATGCTA ATCTTGATGC GGTGCAATA GCGAAAAAGG tTACGCAACA TATAAACTG 1380
 15 CAATAGTAAT CACAGGTAAA GAGGACGTTA TTGtTCmAGA TAATAAGCC TTCGTATTAG 1440
 CTAATGGATC TCCATTATTA GCACGAGTAA CTGGAGCTGG TTGTTTATTA GGAGGCGTTA 1500
 TTGCTGGATT TTTATTTAGA GAAACAGAAC CAGACATAGA AGCGTTAATT GAAGCGGTAA 1560
 20 GCgkATTTAA TATTGCTGCT GAGGTAGCTG CTGAAATGA AAATTGTGGT GGTCTGGTA 1620
 CGTTTTCAAC ATTGTTGCTT GATACGTTAT ATCATTAAAA TGAAACAACC TATC 1674

(2) INFORMATION FOR SEQ ID NO: 210:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2232 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

35 ATGAGTTGCC GATGAATTTA GCACCACCAA CGATTGCnTT TGATACTGTG TCCCAACCAG 60
 CTTGTTTAGC ATATTTAATA CCTTCACGTA AAGGATCGTT ATCATATGCA GCAATACCAA 120
 ATACGTTATG GTATTTCTGTG TTTGAGTTAG TTACAACTTT GTnTTGCACT ACATCTGCAC 180
 40 CTTTCGCTAA TTGAGAAGTA CCGTTACCTG TTTCTAATAG GGCATGTGAG ATAAGATAAA 240
 CTTCAATTAAT GCCATACATT TGAGCAGCTT TGTAAATGC AGCACCTTGG TTTTCTAATA 300
 45 CACCTTTACC TTTTAAGAAT TGATTAATTT TATCAATAGA AATATTTTGT GGTGTTGCTA 360
 AGCGTAAGAA TTGATATTTT AATGCTGGAT CTTGAGCTAA ACGCTTCGTA TCCATTGCAT 420
 GCTTAACATC ATTAAATTTA GCATCTGTCC ACTTACCTGG TACACGTTGT ACTTGTGGTT 480
 50 TATATTGTAA ACCAGCTTGT ATTTGAGCAA CTGTTTTAA TGTCATACCT GTTTGATTAT 540
 ACTTAATTAA TTCTTTAGCT AAATCAGTTG ATTTAATCCA TGCLAAATTA CCGTTAGATA 600
 ATTTACCATA GTACCAAGTT TGTCCATTAA TGACTTGTTC TTTAACAACT GCGAATGGTT 660
 55

AACCATTACC ATTTTAAATT ACATAAGTGT AGTTATAATC TTTGGCAGCT GATGTAGTTG 780
 GTTTCACAGC AGTTGGTGCA GTTAAATCTT TTGCATTTAC CCAACCAGTG CGGTTATTAA 840
 5 TAGTACCGTA TAAATAAACA TCTTTGCCTA CAGATACTTG TTTCGTTGCA TTAAATGTAC 900
 CTTGAGCAAT GTTATTGCCT GTTAAATGA CTTGGTTTTT AGTACCCCAA GGAACCATTG 960
 10 ATAAGCCGTT ATTTGATTTA TTAACAGTAT ATTTTGTAGT CGTTTTAACT TCTTTGCCTA 1020
 AGTTTTGAAC ATTTAAGTCT TTTACATTGA ACCAACCTAA TGGGATGTTA TGGCTTGTAT 1080
 TGTTTAATAA TACATACGTT TCATTACCAT GAGCACGCTC TTTTGTTACA TAGAACGTAC 1140
 15 GGTCTGCATA TTTCGCACCG TTTTTCGCTG TTTTTCATA AACAGAAGCA CGAATACCAG 1200
 TGTGTTTTGG TTTAACTTGA GCAATCTTGC TAACTGTTTG AGTCGTTTGT GGTTTAGTAA 1260
 CAGTATAAGC TTTTACAGCT GTTTTTGGTT GTGCTACTGC TTTTGTAGGT GCAGCAGGTA 1320
 20 CAGCTAAATA TGCTTTACTT ACCCAACCAG ATTTACCATT TACAGTTCCA AATAAATAGA 1380
 TAGATTTATC AATTTGTTGT TGCTTAGTCG CTTTAAAAGT TTGGTTACCT GTACCAGAAA 1440
 25 CTGCACCAGC TTCTTGTTTA TAAGTGCCCC AAGGTACTGA ATATAATTTA GTGCCTGGgT 1500
 TTAAGTGTATA TGTTTGCATT ACATTTACAG GTGATTTTGC ATtGtTATAA ATACGTCACC 1560
 TTGTTTAACC CAACCAATTA AAGTTGGACT ATTGTAATCT TTAACCTAAGT AGAATTTGTT 1620
 30 TCCACCTAAA CTGCTTCTT TTGTTACAGC AAATGTTTTT TGAACCTCTT TCGTTGGCTT 1680
 ACCAGTTTTG TCATAAACTG TAGTGAATAA GCCATTGTTT TTAGCATTAA TTTGAGCAAC 1740
 ACCGTTTAAT GATGAAACTG TTAATTTATT ATTTGTTGTA GGTGTTGATG GCTTAGGTGT 1800
 35 TGGTGTAGGC GTAGGTTTAG CAGTATCAAC TAAATATGCT TTAATTACCC AACCAGATTT 1860
 ACCATTCACA GAGCCATATA AATAAATTGA TTTATCAATT TGTTGTTGCT TTGAAGCCTT 1920
 40 AAATGTTTGG TTTCAGAGC CAGACACACT ACCAGCAACT TGTTTAGATG TACCCCAAGG 1980
 TACTGTATAA AGTTTCGTAC CAGGTTTGAT TGAATATGAT TGATTTACAT TTACAGGTGA 2040
 TTTAGCTGTG TTGTAAACCA CATCGCCTTC TTTAACCCAA CCAAATTTAT TACCAGAATT 2100
 45 GTAATCTTGA ACAAGATAGA ATTTTGTGAT ACCTAATGTA GCTGTTTTAG ATACAGCAAA 2160
 TGTTTTTTGA ACTTCATTAG TTGCTTTACC AGTTTTGTGC TATACAGTAG TATATAAACC 2220
 ACTATTTGTT GG 2232

50

(2) INFORMATION FOR SEQ ID NO: 211:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2082 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

5	GATTTAAATA AAATTAATGG ATATCGTGAT CGTACGATGT TAGAACTTCT GTACGCAACG	60
	GGAATGCGTG TATCTGAATT GATACATTTA GAGTTAGAAA ACGTGAACCTT AATAATGGGA	120
	TTTGTACGCG TATTTGGTAA AGGCGATAAA GAAAGAATTG TACCATTAGG CGACGCAGTC	180
10	ATTGAGTACT TAACTACTTA TATTGAAACG ATTAGACCGC AACTTTTAAA AAAGACTGTT	240
	ACTGAAGTCT TATTTTAAA TATGCATGGT AAACCTTTAT CACGACAAGC AATATGGAAA	300
	ATGATTAAAC AAAATGGTGT AAAGGCAAAC ATTAAAAAGA CGTTAACGCC ACATACGTTA	360
15	CGCCACTCTT TTGCGACACA TTTATTGGAA AATGGCGCAG ATTTAAGAGC AGTGCAAGAG	420
	ATGTtAGGtC ACTCTGaCmT ATCTACTACC CmaCTCTATA CmCATGTTTC GrAATCTCAA	480
	ATTAGAAAAA TGTATAACCA ATTTTCATCCT AGAGCATAAA GTGAACAATA ACTCAAAAGT	540
20	CACAATACAC ATGACTAAAA ATGTCTGTGC TATTGTGGCT TTTTAAATT GGTGATTAA	600
	TTACGTCTAT GTTTTCTTAA TTGAATCGCT TCTTCTTTTG CTGCAATCAC TTCTGAACGA	660
25	TCACGGCGCA TGTGATGGTC TACAATAAAA GGATCTGTTG CTGTTTCCTG ATTATAATCA	720
	TAGTCTGGAT AGTTGGCCTT GATGATGCGT TCAAAGACTG GAGTTATTGG TAATATAACA	780
	GATGAAAAAG GCTTTGCTGC ATTCAATTTT GCAATCTGTT GCTCAATTAA CAACTGATAA	840
30	TCATTTAAAT TAAGGTATAA CGCATCTCTA TCTTTAGCAT TTTGTATTAT TTCTTTAGAT	900
	TTATTAAAAG ACTTATAGGC GCCTTTTAAA TTATTGCGGC GATAATGGTA ACAAGCAGTT	960
	GCAAACAAGA TTAAACTAAC AACTGCATCT TGCTTACTGT AGTTATTTTC AGCTTTCCAT	1020
35	GCATCTTCTA AAATGTCATG ACATAGGAAA TAATGTTGCT TAGTATGAAA TTGATAATAG	1080
	AAATTTATCA GTGCCTGTTG CATTTTGTGA TCACCCCAAT TTAAAAGTAA GTTATTTTCA	1140
40	TGCTATAATA TTTTAGAGAA TTATGCACAT ATGACGCAAT ACGAGGTAGA TATTATGTAT	1200
	GAAGTTAAAT TAGATGCTTT CAATGGACCA TTAGATTAT TGCTGCATCT TATCCAAAAA	1260
	TTTGAAATAG ATATTTATGA TATTCCTATG CAAGCATTAA CAGAGCAGTA TATGCAGTAC	1320
45	GTTTCATGCAA TGAAACAGCT TGAAATTAAT ATTGCAAGTG AATACCTAGT ATTAGCGTCA	1380
	GAACTCTTAA TGATTAAAAG TAAGATGCTA TTACCACAAT CAACATCAGA TATGGATGTT	1440
	GATGATGACC CACGGGAAGA TTTAGTtGGG CGTTTAATAG tATATCaAAA TTATArAGAA	1500
50	TATACTGCTa TTTTAAATGA CATGAAAGAA GAAAGAGATT TTTATTTTAC CAAAAAGACC	1560
	GACAGATTTA TctCATTtGG AAacAGATGA ATCyTGGGAT CCaATCATA CGATTGATTT	1620

55

ATCTGTTGAA ATCCGAAAAG AGACATTTAC CATTCAACAA GCTACAGAAC AAGTGACATC 1740
 GAGATTGAAA GATAAAGATC ATTTTAACTT CTTTAGTCTG TTTACGTTTT CTGAGCCAAT 1800
 5 TGAACAAGTA GTCACCTCACT TTTTAGCTAT TTTAGAGATG TCAAAAGCAG GAATAATTAA 1860
 TATTGAGCAA CAACGTAATT TTGAAGATAT TAACATTATT AGAGGAGTGA ACTACCATTT 1920
 10 TGGATAATCA TGGTATATTA GAGTCGCTTT TATTTACAGC TGGCGATGAA GGTTTAGATG 1980
 AAAACAACCT ATTAGAAATA TTAGATATGT CGAAAGACCA ACTCGTTGAA TTAATTGAAA 2040
 ATTATTCATC ACATGGATTA ATGATACAAC GATTTGGAAT GA 2082

15 (2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4219 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

TCTATTCTCG TTCTTCCAAG ACCCTGaATT AGAAGTTAAG AAAATCGAAG AAGATGAGAA 60
 AGAATCTATT AAAAAAGCTC AAAAAGGTAT TTATAAAGAC CCTAGAGACA TCAATGATGA 120
 30 CGAACAAGAT GATGATACAA AAGATACTGT TGATAAAAAG GAATGATTGT AATTGCCTAA 180
 CAAAAACACT CAAGAATATT GGGAAGAACG CGGACGCAAA GCAATCGAGA ATGAGTTGAA 240
 GCGTGATAAA ACTAAAGCTG AAGAAATAGA ACGTATATTG AATATGATGA TTAAGCGCAT 300
 35 TGAAAAAGAG ATCaATGCGT TTATTGTCAA GTACGGAGAT TTTGCAGGCG TTACATTACA 360
 AGAAGCACAA AAGATTATTG ATGAGTTCTGA TGTAAGAAGCG TTTCAAGAAG AAGCAAAAAG 420
 ATTGGTTCGAA AACAAGGAGT TTAGCGATAG AGCAAATGAA GAATTAAAGA AGTATAACAC 480
 40 GAAAATGTAT GTATCTAGAG AACAGATGTT AAAGATTCAA ATAGAATTCT TAATTGCTTA 540
 TGCAACAGCT CAAACAGAAT TATCGATGAG GGAATATTTT GAATCAACAG CTTATCGTGT 600
 GTTCAGTGAT CAAGCGGGTA TTTTAGGTGA AGGTGTACAA GTAGCTAAAG AAGTTATAGA 660
 45 TACAATCGTT GATACACAAT TTCATGGTGT CGTTTGGTCA GAGCGATTAT GGACTAATAC 720
 CGAAGCAATG AAACAAGAAG TAGAAGAAAT AATTGCTAAT GTAGTTATTA GAGGTCGACA 780
 50 TCCTAATGAA TATGTTAAAG ATATGCGCAA CACTTAAATA AATTCGAAGG CACAGCACGA 840
 CAAAAGACCG CAGCAATTAA ATCATTGCTT TATACGGAAT CGGCACGTGT TCACGCACAA 900
 TCAAGCATTG ACAGCATGAA AGAAATTTCA CCGGAAGgAT ATTATATGTA TATTGCAAAA 960

55

	GACGCTAAAA TTGGTGTTAA TTTCTATCCT ATGCAATATCA ATTGTCGTTT AGATTGCGCT	1080
	TTACTACCTA AATCTATGTG GCCGAAAAAA CCAAGCAAGA AACGAAAAAC AAAATACTTC	1140
5	GGAGGGAAAG TGAAAAGCGG TGATTGATTT AAAAGTGAAG TTTTTTAAAG GCAAGTTAGT	1200
	TTTGTATGAC AGTAAATTAA ATGTTTGGAG GATACTAATA TGAGTAATAC TGACAAATAC	1260
	CTTAGAGACA TAGCAAGAGA ATTAAAAGGT ATACGTAAAG AGTTACAAA GCGAAACGAA	1320
10	ACAGTTATTA TTGATGCAAA CTTAGACAGT TTAAGGTCGG CAGTATTAGC CGATAAAGAA	1380
	AAATCGAAAT ATAATGAACC TCTCTTTTAA TAGCTAGCAC TTAATTGTGT TGGCTATTTT	1440
15	TTATGTCCAA AACGTGCTGA TGACATAAAA AGCACGCATG GAAAAACAGT CGACAGACTA	1500
	TAAATGGAGG TATATCTCAT GGAAGAAAAT AAACCTAAGT TTAATTTGCA aTTTTTTGCA	1560
	GACCAATCAG ATGATCCGGA CGAACCAGGC GGAGATGGTA AAAAAGGAAA TCCTGATAAG	1620
20	AAAGAAAATG ACGAAGGTAC TGAAATAACT TTCACGCCAG AGCAACAAAA GAAAGTTGAT	1680
	GAAATACTTG AACGTCGTGT AGCCACGAA AAGAAAAAAG CTGATGAGTA TGCAAAAGAA	1740
	AAAGCAGCAG AAGCTGCTAA AGAAGCTGCT AAATTAGCGA AAATGAACAA GGATCAAAAA	1800
25	GATGAATATG AACGCGAACA AATGGA AAAA GAACTGGAAC AATTACGTTT AGAAAAACAA	1860
	TTAAACGAAA TCGGTTTCAGA AGCACGAAAA ATGTTGAGTG AAGCGGaAGT TGATTCATCA	1920
30	GATGgGGTTG TCAATTTAGT TGTAACAGAT ACTGCTGAAC AAATAAATT GAATGTTGAA	1980
	GCTTTTCTA ATGCAGTAAA AAAAGCGGTT AATGAAGCGG TTAAGGTTAA CGCTAGACAA	2040
	TCGCCATTGA CTGGTGGAGA TTCATTTAAT CACTCGACTA AAAATAAACC GCAAAACTTA	2100
35	GCTGAAATAG CTAGACAAAA AaGAATTATT AAAAATTAAC GGAGGCATTT AAATGGAACA	2160
	AACACAAAAA TTAAATTTAA ATTTGCAACA TTTTGCAAGT AACAATGTTA AACCACAAGT	2220
	ATTTAACCTT GACAATGTAA TGATGCATGA AAAGAAAGAT GGCACGTTGT TAAACGACTT	2280
40	TACAACACCT ATCTTACAAG AGGTTATGGA AAACCTCTAA ATCATGCAAT TAGGTAAGTA	2340
	CGAACCAATG GAAGGTACTG AGAAGAAGTT TACTTTTTGG GCTGATAAAC CAGGTGCTTA	2400
45	CTGGGTAGGT GAAGGTCAAA AAATCGAAAC GTCTAAGGCT ACTTGGGTTA ATGCTACAAT	2460
	GAGAGCGTTT AAATTAGGGG TTATCTTACC AGTAACAAAA GAATTCTTGA ATTACACTTA	2520
	TTCACAATTC TTTGAAGAAA TGAAACCTAT GATTGCTGAA GCTTTCTATA AAAAGTTTGA	2580
50	CGAGGCAGGT ATTTTGAATC AAGGTAACAA TCCGTTCCGT AAATCAATTG CACAATCAAT	2640
	TGAAAAAACT AATAAGGTTA TTAAAGGTGA CTTACACAA GATAACATTA TTGATTTAGA	2700
55	GGCATTGCTT GAAGATGACG AATTAGAAGC AAATGCATTT ATCTCAAAAA CACAAAACAG	2760

	TGATTTCGTTA	GACGGTCTAC	CTGTGGTTAA	CCTTAAATCA	AGCAACTTAA	AACGTGGTGA	2880
	ATTAATCACT	GGTGACTTCG	ACAAATTGAT	TTATGGTATC	CCTCAATTAA	TCGAATACAA	2940
5	AATCGATGAA	ACTGCACAAT	TATCTACAGT	TAAAAACGAA	GATGGCACAC	CTGTAAACTT	3000
	GTTTGAACAA	GACATGGTGG	CATTACGTGC	AACTATGCAT	GTAGCATTGC	ATATTGCTGA	3060
	TGATAAAGCG	TTTGCTAAGT	TAGTTCCTGC	TGACAAAAGA	ACAGATTCAG	TTCCAGGAGA	3120
10	AGTTTAATAA	ATAATTAGGA	GTGGTAACAT	GCCCGAAATC	ATTGGAATTG	TTAAAGTAGA	3180
	TTTTACAGAT	TTAGAAGATA	ACAGACATGT	CTATATGAAA	GGGCATGTCT	ACCCTCGTAA	3240
15	AGGTTATAAT	CCTACAGATG	AACGTATCAA	AGCTTTAGCT	AGTGTTGAAA	ATAAACGCAA	3300
	CAAACAAATG	ATTTACATTG	TAAATGACAA	ATTAACCAA	AAAGAACTTG	TCGAAATAGC	3360
	AAGTGTTGCT	GGCTTACAAG	TTGATGAAAA	ACAAACAAAA	GCTGAAATTA	TCAATGCTTT	3420
20	TGAGTCACTA	GAGTAGGTGG	TTATATGACT	ACGCTAGCTG	ATGTAAAAAA	ACGTATTGGT	3480
	CTTAAAGATG	AAAAGCAAGA	TGAACAATTA	GAAGAAATCA	TAAAAAGTTG	TGAAAGCCAG	3540
	TTGTTATCAA	TGTTACCTAT	TGAAGTTGAA	CAAATACCGG	AAAGgTTTAG	TTACATGATT	3600
25	AAAGAAGTTG	CAGTTAAACG	CTACAACAGG	ATTGGTGCTG	AAGtATGACA	TCAGAAGCGG	3660
	TTGACGGACG	TAGCAATGCG	TATGAATTGA	ACGATTtCAA	GGAGTATGAA	GCTATTATTG	3720
30	ATAATTACTT	TAATGCTAGA	ACGAGAACTA	AAAAAGGAAG	GGCTGTGTTC	TTTTGAGATA	3780
	TGAAGATAGA	GTTATTTTTT	AATTAGAACA	AGTAGCAACT	TACAATCCTA	AAACTAGCAA	3840
	AAAAGAAAAC	ACACTAATCA	CTTATGATGC	GATACCATGC	AATATTAACC	CCATTTCCTAG	3900
35	AGCAAGAAAG	CAACTTGAAT	TTGGTGATGT	AAAAAACGAT	GTAAGTGTTT	TGAGGATAAA	3960
	AGAATCAATA	TCTTACCCTG	TTAGCCACGT	GTTGGTTAAT	GGCATTTCGCT	ACAAGATAGT	4020
	TGATACAAGG	ATATACAGAC	ACGAAAACGTC	ATATTATATC	GAAGAGGTCA	ATTGATGAAT	4080
40	ATAGATGGAT	TAGACGCACT	GTTAAACCAA	TTTCACGATA	TGAAAACCAA	CATTGATGAT	4140
	GATGTAGATG	ATATTTTACA	GGAAAACGCC	AAAGAATATG	TAGTACGAGC	TAAATTGAAA	4200
	GCTAGAGAAG	TAATGAATA					4219

(2) INFORMATION FOR SEO ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

GCTTACAAGT ATATTCATAA TTACATATTC AAGGTCCTTG CATGTGGTAT TTTGCTATGG 60
 yCtTTaACTA CAACGGGGTC TAAGACTGCG TTTATCATAT TAATCGTCTT AGCCATTtAT 120
 5 TyCTTTATKa AAAAGTTATT TAGTAGAAAT GCGGTAAGTG TTGTGAGTAT GTCAGTGATT 180
 ATGCTGATAT TACTTTGTTT TACCTTTTAT AATATCAACT ACTATTTATT CCAATTAAGC 240
 GACCTTGATG CCTTACCGTC ATTAGATCGA ATGGCGTCTA TTTTGAAGA GGGCTTTGCA 300
 10 TCATTAAATG ATAGTGGGTC TGAGCGAAGT GTTGATGGA TAAATGCCAT TTCAGTAATT 360
 AAATATACAC TAGGTTTTGG TGTCGGATTA GTGGATTATG TACATATTGG CTCGCAAATT 420
 15 AATGGTATTT TACTTGTTGC CCATAATACA TATTGCGAGA TCTTTGCGGA ATGGGGCATT 480
 TTATTCGGTG CATTATTTAT CATATTTATG CTTTATTTAC TGTTTGAATT ATTTAGATTT 540
 AACATTTCTG GGAAAAATGT AACAGCAATT GTTGTAAATGT TGACGATGCT GATTTACTTT 600
 20 TTAACAGTAT CATTTAATAA CTCAAGATAT GTCGCTTTTA TTTTAGGAAT TATCGTCTTT 660
 ATTGTTCAAT ATGAAAAGAT GGAAAGGGAT CGTAATGAAG AGTGATTCAC TAAAAGAAAA 720
 TATTATTTAT CAAGGGCTAT ACCAATTGAT TAGAACGATG ACACCACTGA TTACAATACC 780
 25 CATTATTTCA CGTGCAATTTG GTCCAGTGG TGTGGGTATT GTTTCATTTT CTTTCAATAT 840
 CGTGCAATAC TTTTGTATGA TTGCAAGTGT TGGCGTTCAG TTATATTTTA ATAGAGTTAT 900
 CGCGAAGTCC GTTAACGACA AACGGCAATT GTCACAGCAG TTTTGGGATA TCTTTGTCAG 960
 30 TAAATTATTT TTAGCGTTAA CAGTTTTTGC GATGTATATG GTCGTAATTA CTATATTTAT 1020
 TGATGATTAC TATCTTATTT TCCTACTACA AGGAATCTAT ATTATAGGTG CAGCACTCGA 1080
 35 TATTTTCATGG TTTTATGCTG GAACTGAAAA GTTTAAAAAT CCTAGCCTCA GTAATATTGT 1140
 TGGCTCTGGT ATTGTATTAA GTGTAGTTGT TATTTTTGTC AAAGATCAAT CAGATTTATC 1200
 ATTGTATGTA TTTACTATTG CTATTGTGAC GGTATTAAAC CAATTACCTT TGTTTATCTA 1260
 40 TTTAAAACGA TACATTAGCT TTGTTTCGGT TAATTGGATA CACGTCTGGC AATTGTTTTG 1320
 TTCGTCAATTt AGCATACTTA TTACCAAATG GACAGCTCAA CTTATATACT AGTATTTCTT 1380
 GCGTTGTTCT TGTTTTAGTA GGTACATACC AACAAGTTGG TATCTTTTCT AACGCATTTA 1440
 45 ATATTTTAAC GGTCGCAATC ATAATGATTA ATACATTTGA TCTTGTAATG ATTCCGCGTA 1500
 TTACCAAAT GTCTATCCAG CAATCACATA GTTTAACTAA AACGTTAGCT AATAATATGA 1560
 50 ATATTCAATT GATATTaCA ATACCTATGG TCTTTgGTTT AATTGCaATT ATGCCATCAT 1620
 TTTATTTATG GTTcTTGGT GAGGAATTCG CATCAACTGT CCCATTGATG ACCATTTTAG 1680
 CGATACTTGT ATTAATCATT CCTTTAAATA tGTTGaTAag CaGGCAATAT TTAtTAAtAG 1740

55

TATGTAyTAT TTTGATATAT TTTTATGGAA TTTACGGTGC TGCTATTGCG CGTTTAAATTA 1860
 CAGAGTTTTT CTTGCTCATT TGGCGATTTA TTGATATTAC TAAAATCAAT GTGAAGTTGA 1920
 5 ATATTGTAAG TACGATTCAA TGTGTCATTG CTGCTGTTAT GATGTTTATT GTGCTTGGTG 1980
 TGGTCAATCA TTATTTGCC 1999

(2) INFORMATION FOR SEQ ID NO: 214:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7769 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

20 TCATTATTAA GACTATTATA TATAATGAAT TTAACTGGT TTATTAAACG AGAACGTCGG 60
 GAATTAAGTA ACTACAATAA AAATAAGATA TGACAATAAG GAGACTACAC GCGTGATCAT 120
 TGCCATAATT ATATTGATAT TTATTTTCGTT TTTCTTTTCA GGAAGCGAGA CGGCATTAAAC 180
 25 GGCTGCCAAT AAAACAAAAT TTAAACTGA AGCTGACAAA GGTGATAAAA AAGCAAAAGG 240
 CATTGTAAAG TTACTTGAAA AACCAAGTGA GTTTATTACA ACGATTCTAA TTGGGAATAA 300
 TGTCGCGAAT ATTTTATTAC CAACACTTGT TACAATTATG GCTTTACGTT GGGGGATTAG 360
 30 CGTTGGTATT GCATCAGCTG TTTTAACAGT TGTTATCATT TTGATCTCCG AAGTGATTCC 420
 CAAGTCTGTC GCTGCAACAT TTCCAGATAA AATAACAAGG CTTGTATATC CAATTATTAA 480
 35 TATTTGTGTC ATTGTGTTCC GTCCTATCAC ATTACTTTTA AATAAGTTGA CGGACAGTAT 540
 TAATCGAAGT TTATCTAAGG GCCAACCTCA AGAACATCAA TTTTCAAAAG AAGAATTTAA 600
 AACAAATGTTA GCAATTGCTG GACATGAAGG TGCTTTAAAT GAAATTGAGA CGAGTAGGTT 660
 40 GGAAGGTGTC ATTAATTTTG AAAATTTTAA AGTAAAAGAT GTAGATACAA CACCTAGAAT 720
 TAATGTGACG GCATTGCTT CAAATGCGaC ATACGAAGAA GTTTATGAAA CGGTTATGAA 780
 TAAGCCATAC ACTAGATATC CAGTGACGA GGGAGATATT GATAACATTA TTGGGGTGTT 840
 45 TCATTCTAAA TATCTGTTGG CTTGGAGTAA TAAAAAGAA AATCAAATTA CAACTATTC 900
 AGCTAAGCCA TTATTTGTGA ATGAACACAA TAAAGCTGAA TGGGTATTAC GTAAGATGAC 960
 50 TATTTCTAGA AAACATTTAG CAATTGTGTT GGACGAATTT GGTGGTACTG AAGCGATAGT 1020
 GTCACATGAA GACTTAATTG AAGAATTATT AGGTATGGAA ATTGAAGATG AGATGGATAA 1080
 AAAGGAAAAA GAAAACTTT CTCAACAGCA AATTCAATT CAACAACGGA AAAATCGCAA 1140

	GTATTGAATA TCCAATTATA CAAGCAGGTA TGGCAGGAAG TACGACACCG AAATTAGTTG	1260
	CATCAGTAAG TAACAGTGGT GGGTTAGGCA CAATAGGCGC AGGTTACTTT AATACGCAGC	1320
5	AATTGGAAGA TGAAATAGAT TATGTACGCC AATTAACGTC AAATTCTTTT GGC GTAAATG	1380
	TCTTTGTACC AAGTCAACAA TCATATACCA GTAGTCAAAT TGAAAATATG AATGCATGGT	1440
	TAAAACCTTA TCGACGCGCA TTACATTTAG AAGAGCCGGT TGTA AAAAATT ACCGAAGAAC	1500
10	AACAATTTAA GTGTCATATT GATACGATAA TTA AAAAGCA AGTGCCTGTA TGTGTTTTA	1560
	CTTTTGAAT TCCAAGCGAA CAGATTATAA GCAGGTTGAA AGCAGCGAAT GTCAA ACTTA	1620
15	TAGGTACAGC AACAAGTGTT GATGAAGCTA TTGCGAATGA AAAAGCGGGT ATGGATGCTA	1680
	TCGTTGCTCA AGGTAGTGAA GCAGGTGGAC ATCGTGGTTC ATTTT TAAAA CCTAAAAATC	1740
	AATTACCTAT GGTGGAACA ATATCTTTAG TGCCACAAAT TG TAGATGTC GTTCAATTC	1800
20	CGGTCAATGC CGCTGGTGGA ATTATGGATG GTAGAGGAGT TTTGGCAAGT ATTGTCTTAG	1860
	GTGCAGAAGG GGTACAAATG GGCACCGCAT TTTTAACATC ACAAGACAGT AATGCATCAG	1920
	AACTACTGCG AGATGCAATT ATAAATAGTA AAGAAACAGA TACAGTCATT ACAA AAGCGT	1980
25	TTAGTGAAA GCTTGACGC GGTATCAACA ATAGGTTTAT CGAAGAAATG TCCCAATACG	2040
	AAGGCGATAT CCCAGATTAT CCAATACAAA ATGAGCTAAC AAGTAGCATA AGAAAAGCCG	2100
	CAGCAAACAT CGGCGACAAA GAGTTAATAC ATATGTGGAG TGGACAAAGC CCGCGACTAG	2160
30	CAACAACGCA TCCCGCCAAC ACCATCATGT CCAATATAAT CAATCAAATT AATCAAATCA	2220
	TGCAATATAA ATAATCGACC GCAATCCACA AAAGCACAAG CACCCCCAAA CATTATTTTA	2280
35	GTGCTTGCCA TTTTGTGGA TTGCGTTTCT ATTTTACCAA TTTAATCAA CGAAAACATC	2340
	AAGCTGAAGA TCGCCGAAAG ATTTTAATCA AGCAAAAACA TCAA ACTAAA GTTCGCTGAA	2400
	ATGATTATGA TAAAAGTTAT ATGGTATGAT GACATTGGTG ATATATATGA TAAACATCGG	2460
40	ATTAACAGGT TGGGGTGATC ACTATTCATT ATATGAAGAT TTAGAACGCC AAACCGATAA	2520
	ACTTAAAACA TATGCTGGAC ATTTTCCGGT TGTCGAATTA GATGCGACAT ACTATGCGAT	2580
	ACAACCGGAA AGAAATATAT TGAAATGGAT AAAAGAAACG CCTGATACAT TTGAATTTGT	2640
45	GGTCAAAATT CATCAAGCAC TcACATTGCA TGCAGACTAC AAAACATTTG CAGATACAAG	2700
	GCAAGAACTA TTTGATCAAT TTAAGAATAT GTTAGAGCCC TTACATACAC AGAAAAAATT	2760
50	AGCAATGGTA TTGGTTCAAT TTCCGCCATG GTTTGACTGC AATGCACAAA ATATCAAATA	2820
	TATTTTGTAT GTAAGACAGC AATTACAAGC ATTTCCAATG TGTGTAGAAT TTAGGCATCA	2880
	ATCATGGTTT AGTGATGCAT TTAAAGAACA AACATTGGCA TTTTAAACAG AACATCAAAT	2940
55		

	AATCACAAAT GAAATTGCGT TTGTACGTTA TCATGGACGT AATCATTACG GTTGGACTAA	3060
	GAAAGATATG TCAGATCAAG AATGGCGCGA TGTACGCTAT TTATATGATT ATAATGAGCA	3120
5	AGAATTAATA GACTTGGCAC AAAAGGCACA AATATTAGCA CAAAAAGCTA AGAAAGTTTA	3180
	CGTCATATTT AACATAATT CTGGTGGTCA TGCAGCAAAT AATGCCAAAA CATATCAGCG	3240
	ATTATTGAAT ATAGAATATG AAGGGTTAGC ACCACAACAA TTAAAATTAT TTAAAGAGGC	3300
10	GACGACTATG TTATTAACAA TTACATTATT AGTTTTAATC GGAGGTTTGT CAGCGATTAT	3360
	AGGGTCTATC GTAGGCATTG GAGGCGGTAT TATTATCGTT CCAACAATGG TTTACCTCGG	3420
	TGTTGAACAT GGATTACTAC ATAATATTAC AACACAAGTA GCGATAGGGA CGTCTTCAGT	3480
15	CATTCTAATT GTGACAGGAC TTTCTTCATC ACTTGGATAT TTAAAAACAA AACAAAGTTGA	3540
	TATTA AAAAT GGTTCATCT TTTTATTTGG ACTATTACCA GGTTCAATTGC TTGGGTCCTT	3600
20	CATTAGTAGA TATTTAACAT TTGAGTCATT TAATTTATAT TTTGGTATCT TTTTAATTTT	3660
	CGTAGCCATT TTATTAATGG TAAGAAATAA GATTAAACCG TTTAAAATTT TCGATAAACC	3720
	CAAGTATGAA AAGACTTATG TAGACGCTAA AGGTAAAACA TATCATTATA GTGTTCCACC	3780
25	ATTGTTTGCT TTTATTACAA CGTTTTAAT TGGTATATTG ACAGGTTTAT TTGGTATTGG	3840
	AGGTGGCGCA CTAATGACGC CACTAATGCT TATTGTATTT AGATTTCCAC CTCATGTAGC	3900
30	TGTTGGAACA AGTATGATGA TGATTTTCTT TTCAAGTGTC ATGAGTTCTA TAGGGCACAT	3960
	TGCTCAAGGT CACGTAGCTT GGGGTTATGC AATCATTTTA ATTATTTCTA GTTATTTTGG	4020
	TGCGAAAATC GGTGTCAAAG TGAATCAATC AATTAAGTCA GATACGGTAG TAACATTATT	4080
35	GAGAACAGTA ATGTTGTTAA TGGGTATATA TTTAATTATT CGTGCGTTGA TTTAATACAA	4140
	CTTTAAAAGG AGGACGTCAA TTTGAGGCTT ACAATTTATC ATACGAACGA TATTCATAGT	4200
	CATTACATG AATACGAACG CATTAAAGCA TATATGGCAG AACATCGGCC ACGACTTAAT	4260
40	CATCCTTCTT TATATGTTGA TCTAGGTGAT CATGTAGATT TATCCGCACC TATAACTGAA	4320
	GCAACTTTAG GTAAAAAGAA TGTGGCATTG CTAAATGAAG CAAAATGTGA TGTTGCAACA	4380
	ATCGTAATA ATGAAGGGAT GACCATTTC A TACGAAGCTT TAAATCACCT TTACGACGAA	4440
45	GCAAAATTTA TAGTGACATG TAGCAATGTT ATAGATGAAT CAGGTCATTT ACCAAATAAT	4500
	ATCGTTTCTT CTTATATTAA GGACATAGAC GGTGTGAAAA TACTATTCGT TGCAGCGACA	4560
50	GCACCTTTTA CCCCATTTTA TCGTGCACTA AATTGGATTG TTACCGATCC ACTTGAATCT	4620
	ATAAAGAAG AAATTGAACT TCAACGAGGT AAATTTGATG TATTAATCGT GCTAAGTCAT	4680
55	TGTGGCATT TCTTCGATGA AACATTATGC CAAGAATTGC CTGAAATTGA TGTCAATTTT	4740

	GCAGCTGGAA AGTATGGTAA TTATCTTGGG GAGGTTAATT TAACTTTTGA GGCACATAAA	4860
	GTAGTACATA AAAC TGCAAA GATTATTCCT TTAGAAACAT TACCTGAAGT TGAAACTTCA	4920
5	TTTGAAGAAG AAGGAAAAAC GTTAATGTCC AATTCAGTAA TTCAACATCC AGTAGTGCTT	4980
	AAGCGTAGTA TGAATCACAT AACTGAAGCT GCATACTTAT TAGCTCAAAG TGTTTGTGAG	5040
	TATACACATG CACAATGTGC CATCATCAAT GCTGGCTTAC TCGTTAAAGA TATTGTAAAA	5100
10	GATGAAGTGA CAGAATATGA CATTCAATCA ATGTTACCGC ATCCGATTAA TATGGTAAGG	5160
	GTTAGACTTT TTGGTGTGAA ATTAAAAGAG ATTATAGCTA AAAGTAATAA ACAAGAATAT	5220
15	ATGTATGAAC ATGCACAAGG TTTGGGTTTC AGAGGGAATA TATTGGAGG ATATATTCTT	5280
	TATAATTTAG GGTACATTCA TTCTACAGGG CGTTACTATC TGAATGGAGA AGAAATCGAA	5340
	GACGACAAAG AATATGTACT AGGTACGATA GATATGTATA CGTTCGGTCG TTATTTCCCA	5400
20	ACATTGAAAG AATTACCAA AGAGTATTTA ATGCCAGAGT TTTTAAGAGA TATATTTAAA	5460
	GAAAAATTAT TGGAATATTA AAAAGTAAGA TTATTGGATT TTCATTTGTC ATGAATTTTCG	5520
	ATATAATGTT TAAAGATACA CTTAACAGGA GGGTATGTGT TGTTATGGCG ACAAAAAACG	5580
25	AGGAAATATT ACGTAAACCG GATTGGTTGA AAATAAAATT AAATACCAAC GAAACTATA	5640
	CAGGACTTAA GAAGATGATG AGGGAAAAAA ATCTTAATAC TGTATGTGAA GAAGCTAAAT	5700
	GTCCTAATAT ACATGAATGT TGGGGTGCAC GTCGTACAGC GACATTTATG ATTTTAGGTG	5760
30	CCGTATGTAC AAGAGCTTGT CGTTTTTGTG CGGTTAAGAC AGGTTTACCT AATGAACTTG	5820
	ATTTAAATGA GCCTGAACGT GTAGCTGAAT CAGTTGAATT AATGAATTTG AAACACGTTG	5880
35	TTATCACTGC TGTTGCGCGT GATGATTTAA GAGATGCTGG TTCAAATGTT TATGCTGAGA	5940
	CAGTACGTAA AGTTAGAGAA AGAAATCCAT TTACAACGAT TGAAATTTTA CCATCAGATA	6000
	TGGGCGGGGA CTATGATGCG TTAGAAACAT TAATGGCGTC AAGACCTGAC ATTTTAAACC	6060
40	ATAATATTGA AACTGTTTCG CGCTTAACAC CGAGAGTTCG TGCGCGTGCG ACTTACGACA	6120
	GAACATTAGA GTTTTTACGT CGTTCAAAAG AATTACAACC GGATATCCCA ACTAAATCAA	6180
	GTATTATGGT TGGATTAGGT GAAACTATAG AAGAAATTTA TGAAACGATG GATGATTTAC	6240
45	GTGCGAATGA TGTAGATATT TTAACGATTG GTCAATATTT ACAACCTTCA CGTAAACATT	6300
	TAAAGGTTCA AAAATATTAC ACGCCTTTAG AGTTTGGTAA ATTAAGAAAA GTGGCAATGG	6360
50	ATAAAGGGTT TAAACATTGC CAAGCTGGAC CTTTAGTACG TAGTTCTTAT CATGCGGATG	6420
	AGCAAGTAAA TGAAGCTGCT AAAGAAAAGC AACGCCAAGG TGAGGCACAG TTAAATAGTT	6480
	AATATTTAAC CATTAAATAAG GCATAAAGGC TTAGTTTGTA CAAAACGAAC GTGTCATAGA	6540
55		

AGGTGAAGAA TTTGATAAAA GTAGATCAAC ATTACTTTGA ATTAATAGAA AATTATCGCG 6660
 AATGTTTTAA TGAAGAACAA TTTATTGCTA GGTATTCAGA TATTTTAGAT AAATATGATT 6720
 5 ACATAGTTGG TGAATATGGT TACGATCAAT TACGATTAAA AGGTTTTTAC AAAGATTCTA 6780
 ATAAAAAGC AGAGATGAGT AAACGTTTTT CAAATATTCA AGATTACATA TTTGAATATT 6840
 10 GTAACTTTGG TTGTCCTTAC TTTGTATTAA GACATTTGTC TAAACAAGAG GTTAAAAAGT 6900
 TAATCGAAGA AGTTCATCCG TCTGATGTGA TAGATGACGA CAATAAACTT CAAGATGTGA 6960
 AGATTAAAGCC AACCATTCAA GATACTGAAC ATTAATAAAA CCCTTAGCTA GATTGAAAAT 7020
 15 GGAATCATG CAATTCAAGC ATGGACCTGT AATCTAGTTA GGGGTTTTTA TCTTTAATGA 7080
 ATGACTTCAT TTAAATACTC AGTAATTTCA TCGCCTTCTT CAGCATTTAC ACCTAAAATA 7140
 TGAGCGATAT AGCCTTCTTC TTTTAAATCA TCAGTACCGA TAATACCGAA TTTATTTGTT 7200
 20 TGCATATTAA GTACGAGTGT CTTACCATAA TGTCTATTTG TATGGACTAA CATCAAATCA 7260
 TATCGACTAT GCTCGCCAAC AAAACCAACA AACTGAACTT GACTCTCTTC GTTGTCTATCA 7320
 TATAAATACA TATCAATCAT TTTGTAGCGA CTCCTTTTAA AAGTAGTAAA GTTAGTATAA 7380
 25 CGACAAATGA AGTATACTGC AAAATTATGA TAATATATAA GTGAGAGGTG ACAAGGAATG 7440
 TATTTTGTAG ACAAAGATAA ACTAACTCAG AAATTAGCCT ATTTACAAGC ATTAAGTATG 7500
 30 GATTATCATG AGAGCAAGCA CAATCATTAT GCATTTGAAC GCATTGCTCA AATGTTGATA 7560
 GAATCATCGG TAGATATAGG GAATATGATT ATCGATGCAT TTATTTTAAG GGATCCTGGT 7620
 AATTATAAAG ATGTGATTGA TATATTAGAA CTAGAAAATG TTATTACTAA AGAAACACAG 7680
 35 CAGGCGATTA ATAAAACTGT CGGTATTCGT AAACAATTTA CATATGATTA CACAGCCTTA 7740
 GATGTTGAGA TTATCATGCC AATGTTTGA 7769

(2) INFORMATION FOR SEQ ID NO: 215:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 644 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

50 ACCGCCACCC ATTAATGATT GCTTAAAATC AATAGTCGTA CCATTTAATA CGGGTGCATC 60
 TTTTTGTCT ACTAATACTT TTAATCCAAA GTATTCTAAG ACTTCATCAT TTTCACCAGG 120
 CGCTTCTTCT GCACCCATAC CGTATGTTAA ACCAGTGCAC CCGCCACCAT TCACTTTAAT 180

TGCTTCTGTT AATATAACTG TTGGCATGAT AACTCCTCCT TAAAAAATCC AAGTTTCTTT 300
 TATATGTGCA TATATATTTT GTAATAATTC TTCCGGCGAA TCACCTTCAA CAATATCACC 360
 5 ATTTACTAAA GCATACAACC CGGCTGAACA TATACCACAA TGTGTCAGGC AACCATACTC 420
 TAACACATCG ACATCTGGGT CATTTTCCAG TTGATTAAAA ACATAATCTC CACCTTTTGC 480
 10 CATGTTAGAG AGACAAAATT CTACGATCGG ATTCATACTT CACCTTCTTA TTTCAATTTGT 540
 TACAATATTA TAGCATTTTA AACTGGTAT TTTAACATGA TGTGCTCAAT TAGCAACAAC 600
 TGATGTTTCT TATCCCAGTT ATGTAATAGT GCCTTAGTTA GTAC 644

15 (2) INFORMATION FOR SEQ ID NO: 216:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1578 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:

GAATGATGAA AGGAATAGAA AAGAAAAGAT AAATAATGTA ATAGATTTAT CCGAGAAAAT 60
 TGAAAGAACA AAAGATATGC CAATCAAGAA TACTATAACT ACTCAATTAG GAAATAAACT 120
 30 TATTGGCACA AAAAAAGCTC GTTTTGATGA TAAGAAAGTA GTGTCGTTTG GAGCATTTGA 180
 AGATGAATAA AATAAATGAT AGAGATTTAA CAGAATTGAG TAGCTATAGG GTTTATCAAG 240
 ACATCAATAA AGATAATGAC TTTACAGTTA ACGAAAAACG ATTTAAGCAG GCAGATGTAT 300
 35 TTGAAGATTT ATATAGAGAG AACTAAAAG ACACAAATAA ATTAAGAGAG TATAATTATT 360
 TACAAAATGA AACTTTTAAA AGCGCATAAA TAGGTGATGA GATATGCTTA AAAAAGCAAA 420
 ATTTATCTTA ATGGCAACGA TACTACTATC AGGATGTTCA ACTACCAATA ACGAATCCAA 480
 40 CAAAGAACA AAATCTGTAC CAGAAGAAAT GGATGCTTCA AAATATGTAG GACAAGGATT 540
 CCAACCACT GCAGAAAAAG ATGCGATTGA ATTTGCAAAG AAGCATAAAG ATAAAATTGC 600
 45 TAAGCGAGGC GAACAATTTT TTATGGATAA CTTCGGTCTA AAAGTTAAAG CTACAAATGT 660
 TATAGGTAGT GGCGATGGTG TAGAAGTATT CGTGCAATTGT GATGACCACG AyATCGTATT 720
 TAATGCGAGT ATTCCATTTG ATAAATCAAT wATTGAsAGT GATAGCTCAT TAAGAAGTrA 780
 50 GGAYAAAGGY GATGATATGA GTACTTTAGT TGGTGCAGTA CTCAGTGGGT TTGAATATCG 840
 AGCACAAAAA GAAAAATATG ATAAATTATA TAAATTTTTC AAAGATAATG AAGAGAAATA 900
 TCAATATACA GGATTTACAA AAGAAGCAAT TAATAAGACG CAAAATAGTG GTTATGAAAA 960

55

5 ACCATTGTTA AACAAAAGTG ACAGTGAATT TTCAAAAGAA TTGTCAAATG TTAAGAAGCA 1080
 ATTAAGAAGAT AAGTCTAAAG TTTCGGTAAC TACTACTCTA TTTAGTAAAA AAAAGAAGTA 1140
 TACTAAAAAA AGTAACAGTG AAAATGTAAT AAAAATGGCA GAAGAAATAA AAAAAGATAA 1200
 AGAGATACCA AACGGTATAG AGCTTAGTAT AAAATTTTCG GACAATAAAA TAAATACGGT 1260
 10 TAAACCAAAT TTAAACGGTG aAGCACTTC AGAATATGGT GTGTTTGATC AAGAATAAAA 1320
 TTAATGATGa AAATTTAACG GAGAATAGTG TATATTGAGT AGATCmAGAA TAAAAAGATA 1380
 ATTCTACTAT TGTTGTGAAG GCAAATAAGT AGAAGATTTT AAGTGTAATT TCTGGTGATT 1440
 15 TAAATAATAA TATAnATGGn AGTACTGATA TAAnACTTTT TAACCTACTA GATTCTTATA 1500
 ATTTGCTTTC CATTTTATGA CGATTTTAC TCCAATTGAG TGATAGAATC CAAAAAGCC 1560
 ATCTCCAAAA ATTAATCC 1578

(2) INFORMATION FOR SEQ ID NO: 217:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5137 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

30 TGTTTTCTTT GGGTTAAAC ATGCTTGCTA TCGTTTTGTA AATATGACTT GCTGTTTTnA 60
 CCTGnATACC CGTCACACCA TGGAAGTAAA AATGTTTCTT GCTCTTGGCT TACAATTTTA 120
 35 GCTTTAATCG CTTCATATGC TTTATATTGG TCTTCTGTTA ATTGCTGTTT TGATTCTTGT 180
 TCGAAACAC GATCTTTAAA TGGGTCTCTT TCAACAACCG CGTCATATT TTCAACATAA 240
 40 CCTTTTTTGA TAAGTCCATC TAACTGGAT TTTGAAAAGC CCATATCCTC AATATCAGTT 300
 AAAAATATTG TTTTATGTTG TTCTTCAGAC AAGTAAGCAT ACAAATCGTA TTGTTTAATA 360
 ACTTTCTCCA ACTTAGCTAA TACTTCATCA GGATGATACC CTTCAATGAC ACGAACAGCA 420
 45 CGCTTGGTTT TTTTAGTTAT ATTTTGTGTG AGAATCGTTT TTTCTTCAAC GATATCATCT 480
 TTTAACAAC TCAATAGCAA TTGAATATCA TTATTTTTTT GCGCATCTTT ATAATAATAG 540
 TAACCATGCT TATCAAATTT TTGTAATAAA GCTGAAGGTA GCTCTATGTC ATCTTTCATC 600
 50 TTAAATGCTT TTTTATACTT CGCTTTAATA GCACTCGGAA GCATCACTTC TAGCATAGAA 660
 ATACGTTTAA TGACATGAGT TGAACCCATC CACTCACTTA AAGCTATTAA TTCTGATGTT 720
 55 AATTCTGGTT GTATATCTTT CACTTCTATG ATTTTTTTTA ACTTCGAAAC GTCAAGTTGT 780

	ACAATTACAC GCACACCAGG TTGGATGACA GATTCGAGTT GTTCGGGAAT TATATAATCA	900
	AATTTATAGT CAACGCTCTT CGACGCGACA TCGACTATGA CTTTCGCTAT CATTATTGCC	960
5	ACCTAGTTTC TAGTTCATCT AAAATTTGTG CAGCTAATAC TACTTTTTTT CCTTCTTGA	1020
	TATTTACTTT TTCATTATTT TTAAAATGCA TTGTCAATTC ATTATCATCA GAACTAAATC	1080
10	CGATAGACAT ATCCCAACA TTATTTGAAA TAATCACATC TGCATTTTTC TTGCGTAATT	1140
	TTTGTGTGTC ATAATTTTCA ATATCTTCAG TCTCTGCTGC AAAGCCTATT AAATACTGTG	1200
	ATGTTTATG TTCACCTAAA TATTTAAGAA TGTCTTTAGT ACGTTTAAAA GATACTGACA	1260
15	AATCACCATC CTGCTTTTTC ATCTTATGTT CTAATACATC AACCGGTGTA TAGTCAGATA	1320
	CGGCTGCTGC TTTTACAACA ATATCTTGTT CGTCAAATCG GCTTGTCACT TGTTCAAACA	1380
	TTTCTTCAGC ACTTTGAACA TGAATAACTT CAATATCTTT TGGATCCTCT AGTGTGTAG	1440
20	GACCAGCAAC TAACGTCACG ATAGCTCCTC GATTTTCGAA TGCTTCAGCT ATTGCATAGC	1500
	CCATTTTTCC AGAAGAACGA TTGGATACAA ATCTGACTGG ATCGATAACT TCAATAGTTG	1560
25	GTCCTGCTGT AACCAATGCG CGTTTATCTT GAAATGAACT ATTAGCTAAA CGATTACTAT	1620
	TTTGAAAATG AGCATCAATT ACAGAAACGA TTTGAAGCGG TTCTTCATA CGTCCTTTAG	1680
	CAACATAACC ACATGCTAGA AATCCGCTTC CTGGTTCGAT AAAATGATAC CCATCTTCTT	1740
30	TTAAAATATT AATATTTTGC TGCGTACGTT TATTTTCATA CATATGCACA TTCATAGCAG	1800
	GCGCAATAAA TTTCCGTGTC TCTGTGCTA GCAACGTTGA TGTCACCAA TCATCAGCAA	1860
	TACCTACACT CAATTTTGCA ATTGTATTTG CCGTTGCAGG TGCAACAATG ATTGCATCTG	1920
35	CCCAATCACC TAATGCAATA TGCTGTATTT CTGAAGGATT TTCTTCTATA AAAGTATCTG	1980
	TATAAACAGC ATTTGACTT ATTGCTTGAA ATGCTAATGG TGTCACAAAT TTTTGTGCGT	2040
	GATTGGTTAA CATAACGCGA ACTTCATACC CAGATTGTGT TAACTTACTT GTCAAATCAA	2100
40	TTGCTTTATA TGCCGCAATG CCACCTGTAA CGGCTAATAA TATTTTCTTC ATATTCAATC	2160
	TCCCTTAAAT ATCACTATGA CATTTACGCT TTACATCATC ATATGCGCAC AAATGCTCAT	2220
45	TACTTTTTTA TAGATACAAA TTTAGTATTA TTATAACATC AATCATTGGA TAACTAAAA	2280
	AAACACACCT ACATAGGTGC GTTTGATTTG GATATGCCTT GACGTATTTG ATGTACGTCT	2340
	AGCTTCACAT ATTTTAAATG GTCGAAACTA TTCTTTACCA TAATAATCAC TTGAAATAAC	2400
50	AGGGCGAATT TTACCGTCAG CAATTTCTTC TAACGCTCTA CCAACTGGTT TAAATGAATG	2460
	ATATTCACCT AATAATTCAG TTTCAGGTTG TTCATCAATT TCACGCGCTC TTTTCGCTGC	2520
55	AGTTGTTGCA ATTAAATACT TTGATTTAAT TTGTGACGTT aATTGGTTtA AAgGTGGATT	2580

	TTTA ^r GTG ^c T CAGCTTCTAC AATACATTGA ATT ^c cATT ^c y TCG ^c AAG ^t TC TACTTCAT ^c CA	2700
	TTAACTACAA cGTAAYCGTA TAAATTCATC ATT ^c TcACTT CTkTACGCGC yTCGTTAATA	2760
5	CGACTTTGTA TTTTCTCATC AGATTCTGTT CCTCTACCTA CTAATCGCTC TCTCAAGTGT	2820
	TCTAAACTTG GAGGTGCTAA GAAAATAAAT AGCGCATCTG GAAATTTCTT TCTAACTTGC	2880
10	TTTGCACCTT CTACTTCAAT TTCTAAAAAT ACATCATGAC CT ^c cGTCCAT TGTATCTTTA	2940
	ACATATTGAA CTGGTGTACC ATAATAGTTG CCTACATATT CAGCATATTC TATAAATTGG	3000
	TCATCTTTGA TTAAAGCTTC AAACGCATCC CTAGTTTTAA AAAAGTAATC TACGCCATCA	3060
15	ACTTCACCTT CACGCATTTG ACGTGTGTC ATTGAAATAG AATACTTATA TGATGTACTT	3120
	GGATCTTCAA ATATnCGTnT TCTAACAGTA CCTTTACCTA CTCCAGATGG TCCTGATAAA	3180
	ACGATTAACA ATCCTTTTTT ATTATCCATG CCTTACGACC TCTCTAAGCT AATCTTCTAT	3240
20	TATTTAAATA TGATATCACA TTGTTCTTTA TATTGTATAG CATATTTGAA ATTGCATGCC	3300
	ATAATTTCTA TTAAGTCTAA CAATATCGTT ATATTGCACG ATTAATTTTA ATTAAATAAA	3360
	TTGAATTGCA AACTTTTAGA TAATGTAAAA TGTATGGCAT AATGTATGGT TCAATAACTA	3420
25	TACTGAAAAG TTACAATCAT GTTAAATGA AACGAATGAT ATGAAGAAGG TGGAAGATAA	3480
	ATTATGGCTT ATGATGGCTT ATTTACAAAG AAAATGGTTG AGTCTCTACA ATTTTAAACA	3540
30	ACAGGACGTG TTCACAAAAT CAATCAACCT GATAATGACA CGATACTAAT GGTGTACGT	3600
	CAAAATAGAC AAAACCATCA ATTGTTATTG TCAATCCATC CAAACTTTTC AAGATTACAA	3660
	TTGACTACTA AAAAATATGA TAATCCATTT AATCCACCCA TGTTTGCGCG TGTTTTTAGA	3720
35	AAACACTTAG AAGGTGGTAT TATCGAATCG ATTAAGCAAA TTGGTAATGA TCGTCGCATT	3780
	GAAATCGATA TAAAGAGTAA AGATGAAATT GGCGATACTA TTTACCGCAC TGTCACTCCTT	3840
	GAGATTATGG GTAAACATAG TAACTTAATT TTAGTAGATG AAAATCGCAA AATAATTGAA	3900
40	GGATTTAAAC ACTTAACACC AAATACGAAT CACTATCGTA CAGTAATGCC AGGATTTAAT	3960
	TATGAAGCAC CACCTACTCA GCACAAAATA AATCCGTATG ATATTACAGG TGCAGAGGTG	4020
45	TTGAAATATA TCGATTTTAA CGCAGGTAAT ATTGCTAAAC AATTATTGAA TCAGTTTGAA	4080
	GGATTTAGCC CTTTAATTAC GAATGAAATC GTTAGTCGTC GTCAATTTAT GACTTCATCA	4140
	ACATTACCAG AAGCATTTGA CGAAGTAATG GCAGAAACCA AGTTACCACC TACTCCTATT	4200
50	TTTCATAAAA ATCATGAAAC AGGTAAAGAG GATTTCTATT TTATAAAGTT AAATCAATTT	4260
	AATGATGATA CAGTTACATA CGATTCAATTA AATGATTTGC TTGATCGTTT TTATGATGCG	4320
55	CGTGGCGAAC GTGAACGCGT TAAACAACGT GCGAATGATT TAGTTCGATT TGTTCAACAG	4380

ATAAAGATAC TGAACAGTTA TATGGTGAAT TGATCACTGC TAATATATAT CGAATTAAGC 4500
 AAGGCGATAA AGAAGTGACG GCATTGAATT ATTATACGAA TGAAGAAGTT GTCATTCCTT 4560
 5 TAAATCCTAC AAAATCCCCA TCAGCAAATG CTCAATATTA TTATAAACAA TATAAycGTA 4620
 TGAAAACGAG AGAmCGTGAA TTACAACATC AAATTCAATT GACGAAAGAC AATATAGATT 4680
 10 ATTTTTC AAC AATCGAACAA CAATTACATC ATATTTCTGT CCATGACATT GATGAAATTA 4740
 GAGATGAATT AGCAGAACAA GGCTTTATGA AACAGCGTAA AAATCAAAC T AAGAAAAAGA 4800
 AAGCGCAGAT TCAATTACAA CATTATGTAT CAACTGATGG CGACGATATA TATGTTGGTA 4860
 15 AGAATAACAA GCAAAATGAT TATTTAACAA ATAAAAAGC TAAAAAACT CACACATGGT 4920
 tACACACAAA AGATATTCCT GGTTCCATG TCGTTATATT TAATGATGCA CCAAGTGATA 4980
 CGACAATCAA GGAAGCGGCT ATGTTAGCAG GATACTTTTC AAAAGCTGGT AATTCTGGAC 5040
 20 AAATACCTGT TGATTATACA TTAATTAAAA ATGTGCATAA ACCATCaGGT GCAAAGCCTG 5100
 GGTGTGTAAC ATATGACAAT CAAAAAATT TGTATGC 5137

(2) INFORMATION FOR SEQ ID NO: 218:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2267 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:

35 GTTTTATCGC AGCAGTAAAG CTATCAATCG GCGGTTCAAT TGATGATGCA TTAGCAGAAA 60
 TCAnACAATC ATTTTAGTTA AAATTTACTA ATAATGAaAA ATGTAAACCT TTTTCAAATG 120
 AAAC⁻TTTATa AaAAATATGA TAGTATATAT GTAAATGTTT AATAAAATCT GGAGAAATAG 180
 40 GAGGACATTG CCATGCAACA CCTTATAAAA AAACATGTAT TGAATGGCGA GTTTGATTTA 240
 GTACGACAAT TGATGTCCGA AACAGATTTT ATGGAATTTG AAGAAGCATA TATTTCAAGT 300
 45 GCGCATGAAG TAGAAAGTAT GATGTTTTAT ACATGTATTT TAGATATGAT TAAGTACGAA 360
 GAATCATCTG AAATGCATGA CTTAGCATT TATTGCTTG TGTATCCACT AAGTGAATAT 420
 GAAGGTGCTT TGGATTCTGC TTATTATCAT GCAGACGCTT CCATAAACT TACTGACGGC 480
 50 AAAGAAGTTA AAAGTTTGTT ACAAATGTTA TTATTGCATG CGATACCAAC ACCTGTTATT 540
 TCAGATAAGA AGGCTTTTGA TATCGCCAAG CAAATTTTAA AATTAGATCC TAATAATAAT 600
 GTTGCTCGTA ACGTCTTAAA AGACACTGCC AAACGTATGc gACAaCGTTG TTGTTGATAT 660

AGTTTTAACA TTTGGTTGGG TTGGGCATAT GTTCCAGCCT TTTTAAATAC TTAAAACTA 780
 ACGAAgTATA CTTGTGTGCA CAAATGGTTT TTATACAACA TTTTATAAAT TTATACATT 840
 5 TAATAAAGAA CATACGATAG ATGGTTTAAA CCTTGTTAAC TGAGAAATTT TGATATGTAT 900
 TCTTCGAAAT TTAATAAAT ATACGAAATT CAAGAAGCAC AATAATTAAT CATTTTTCCT 960
 10 ATACAAAAGT TCGTATGACT GCATTATAAA AGCATAAATT TATAATTTTT TTAAATGTCA 1020
 TTGAACGTGA TAATGTGAAT GGATTGAGCA ATTTTGAAAA AGTGAAAAAT AACCTATGCG 1080
 ACTTGCAATT AATTTTCAGT ACGTTATAAT GCACACTGTG CAAAATTAAG GAGGTCTATT 1140
 15 ATTCACATGA TGATGAaTAA AGAAGCAACA AAAATTGGAT TGCCTACGT CGGCATTGTA 1200
 GTGGGCGCAG gATTTTCAAC TGGACAAGAA GTTATGCAAT TTTTCACTAA ATATGGCTTG 1260
 TGGGCTTATT TAGGTGTTAT TATATCTGGT TTTATTTTAG CTTTATTGG GCGCCAAGTA 1320
 20 GCAAAAATTG GTACTGCCTT TGAAGCGACA AATCATGAAT CAACATTACA ATACGTATTC 1380
 GGTGAAAAGT TTAGTAAAGT CTTTGaTTAT ATTTTAATCT TCTTCTTATT TGGTATAGCT 1440
 GTAACCATGC tAGCTGGTGC AGGCGCAACA TTTGAAGAAA GTTATAACAT ACCTACATGG 1500
 25 CTAGGTGCTT TaATTATGaC ATTAGCGATT TATATTACGT TGCKATTAGA CTTTAATAAA 1560
 ATAGTACGTG CACTAGGTAT CGTTACACCA TTTTAAATTG TTTTAGTTGT ATTAATCGCT 1620
 30 GGC GTTTATT tATTTAAGG TCATGtTTCA TTAGCAGAAG TTAACCAAGT AGTGCCtGAA 1680
 GCAAGTATTT GGAAGGGAAT CTGGTTTGGT ACAATATATG GTGGATTAGC TTTTCTGTGA 1740
 GGTTTTAGTA CCATCGTAGC AATCnGTGGG GATACTGAAA AGCGTACAGT GTCAGGTGCA 1800
 35 GGC GCGATGT ATGGTGGTAT TATCTATACT GTATTACTAG CATTGATCAA CTTTGcATTG 1860
 CAAGTGaATA TCCAAC TATT AAAAATGCCT CAATTCCTAC ATTGACGTTA GCAAATAATA 1920
 TCCATCCTTT AATAGCAACA GTGkTATCTG TTATTATGCT GCGGkTATG TATAATACTA 1980
 40 TTCTAGGACT AATGTATTCA TTTGCAGCAC GTTTTACAGA ACCATACAGT AAAAATTATC 2040
 ATATCTTTAT TATTATAATG ATGGTAGCAG GTTATTTATT AAGTTnCGTA GGATTnGCTG 2100
 AATTAATTAA TAAGTTATAT ACnATTTATG GGATATGTAG GCTTATTnTA TTGTAGTAGC 2160
 45 TGTAATTATn AAATATTTCC AAACGTAAAA ATGGCGGATA AAAAACATAT TGCTTTAATA 2220
 TCATATGGAG GGGATATCCG AAAC TTTTACA ATTTGAATCA CTTTGGT 2267

50 (2) INFORMATION FOR SEQ ID NO: 219:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6336 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

5	GTCAATGTAA CCTAATAGTT TATGTCTATC TTGTGTACCA ACTACTACAT CGACACCAGG	60
	AATTTCCATA ATTTTCAGCTG ATGAAGTTTG CGCATAACAA CCTGTTACAC AGATTACAGC	120
10	ATCAGGATTT TGTCTTATTG CACGTCTAAT TATTTGACGA CTTTTTTTAT CACCCGTATT	180
	CGTTACTGTA CAAGTATTAA TAACAAATAC ATCAGCATTG GCTTCAAAGT CAACGCGCTC	240
	ATAGTTTGCT TCTTTAAATA ATTGCCAGAT TGCTTCAGTT TCATAATGGT TTACTTTTACA	300
15	ACCTAATGTG TGaACGCAAC TGTTGACATA AATATTCACC CCATTAATTC TTTTTCATAA	360
	CTTATTGCAC TTAACGCATA CAATGGCGCA GTTCTGCCC GTAAAATTCT CGGCCCAAGA	420
	CCAACAAC TG TACTAGTATT ACTAAATAAT GAAATTTTCT TTTCTGACAA ACCACCCTCA	480
20	GGaCCAAAA TCATCAACAC TTTATCCTGA GCATTGAATT GTTGTAAGT TTGCTTGAAA	540
	TTGCTTAACT CACCATCTTT TGCTTCCTCT TCATATGCAA TAAGAATATA GTCATAATTA	600
	TCAATAGTAT CACAAATTAA TTTTAAATTC GACTCGAATT GAATAGATGG AATCACTAAA	660
25	CGATAGCTTT GTTCAGCAGC TTCTTTAATT ATTTTTTGCC AACGCTCTAT CTTTTTGCCA	720
	ACTTTTGCCT CGTTTAAATT AACAATTGAA CGTTCCATGC TCACAGCTAT AAATGATGAA	780
30	GCACCCAATT CAGTAGCTTT TTGTAGCAAC CACTCATATT TGTCAGCTTT GATTAGTCCA	840
	CTGCAAATCG TAACATCAAC TGGCAATTCT GTATTAATAT TTTGTTTTTC TTTTAAATCA	900
	ACTTCAATTT TATCACTTGT TATGTCAGCA ATTTACATA AATAAACTGT TTGATCATT	960
35	AAAGTTAAAA TAATTTTACT ACCAACATCA TATCTCATT CATTGTTAT ATGATGAATA	1020
	TCTTCTTTTT TTGTAATAAA AAAACGCTGA CTTACATCAG CGTTTTGGnT CTATGAAATA	1080
	ACGTTGCACA TTATTCCTC ACTTTCTGGC CAACAAGACA AACCCAACCG TTGTCATGTT	1140
40	GTTCTGAAAT AATTTTAAAA CCTACACGCT CCATATGTGA CTGTATACCT TCATACTTCT	1200
	CTTTTATAAT ACCAGAAGTA ATAAAATAAC CGCCTTCATT TAGAGTATTA TAAGCATCTT	1260
45	CAATCATTTT ATCAATAATA TGCCTGCTAAA TATTTGCTAT TACAATATCA AATTTTTCTG	1320
	TTTCGTCTTT CAATAAGTTA CCTGGAACAG CTTCAATTAA CGTTTCACAA TGATTTCTTC	1380
	TGAAGTTTTT TTTAGCTACA CTCCTGCTCA TTTTCATCAAT ATCCAACGCT TTAATACGTT	1440
50	TTACACCGAT TAGATGACTT GCAATACTTA ATATACCTGA GCCAGTACCA ACATCAATTA	1500
	CTGAATGCTG TGGCAATACA TATGTTTCTA TTGCCTTCAA ACACATACTT GTAGTCGGAT	1560
55	GATCACCTGT TCCAAAAGCC ATACCTGGGT CGAGCTCAAT GCAAAGCTCT TCATCCGCTT	1620

	GGAAATAGTT TTTCCATTCA TTTTCCCAAT CCGTCTCTGC AATAATTGTC TCACTGAATT	1740
	GAACGTTATG TTGATCAAGT TCATCTAAAT TTAATAACTC ATCTTTAATT TGCTGTCGCA	1800
5	ACTTATCATC ATAAGTCATT TCATTAAAAT AGGCTTTCAA TCTTACTCCC TTATCTGGAT	1860
	AATCCTCTTT TTTCAAAGCG TAAATTTAC CGTATTTATC TTCTGGTTGG TTAATTAAAT	1920
10	CATCTGAATC TTCTATCAG ACACCATTG ATCCATGATT TTCAAGTATA TTGGTAGCCA	1980
	ATTCTACTGC TTCATGATTA ATAATAATTG AAAGCTCTGT CCAGTTCATA CTTTATTCTC	2040
	CCTTAAAGAA TCTTTTTGCT CTATCTTTAA AATTCGAAGG TTGTTCAATTA ATTTCTTCAC	2100
15	CATTTAATTG GGCAAATTCT TTCATTAGTT CTTTTGTCT ATCTGTTAAT TTAGTAGGCG	2160
	TTACTACTTT AATATCAACA TATAAATCTC CGTATCCATA GCCATGAACA TTTTTTATAC	2220
	CCTTTTCTTT TAAGCGGAAT TGCTTACCTG TTTGTGTACC AGCAGGGATT GTTAACATAA	2280
20	CTTCATTATT TAATGTTGGT ATTTTTATTT CATCGCCTAA AGCTGCTTGT GGAAGCTAA	2340
	CATTTAATTT GTAATAAATA TCATCACCAT CACGTTTAAA TGTTTCAGAT GGTTTAACTC	2400
	TAAATACTAC GTATAAATCA CCAGCAGGTC CTCCATTAC GCCTGGAGAG CCTTCACCAG	2460
25	CTAATCTAAT TTGTTGTTCA TTGTCGACAC CTTCAGGTAC TTTCACTTCT AATTTAACTG	2520
	TTTTATTTTC AGTACCTTTT CCGTGACATG TTGGACAAGC TTCTTCAAAT TCTTGACCAC	2580
30	TTCCATTACA TTTAGGACAA ACTTGTTTCAG TACGAACTCT ACCTAAAATT GTGTTTTGTT	2640
	CTACAGCTAC ATGACCAGCG CCATTACAGT AACTACAAGT CTTTTTACTT GTTCCAGGCT	2700
	TTGCACCATC ACCATGACAT GTTTCGCATG TTACATCTTT ACGGATTGAA ATTTCTTTTG	2760
35	TTGTACCAA TACCGCTTCT TCAAATGTTA ATGTCATTGT ATACTGAAGA TCATCACCTT	2820
	TTTGCGGTGC ATTTGGATCT CTTTGTCTGC CGCCACCGAA GAAAGAGCTA AAGATATCTT	2880
	CAAAAGCGCC GCCACCGAAG CCACTAAAAC CGCCAAAGTC AGAGCCATTG AATCCTTGTC	2940
40	CACCAAACC TTGTGGACCA TCATGTCCAA ATTGATCATA GCTTGCGCGT TTATTATCAT	3000
	CACTTAAAAC TTCATAGGCT TCAGAAATTT CTTTAACTT TTCATCTGCA CCTTCTTCTT	3060
45	TGTTAATATC TGGATGATAT TTTTTCGAAA GCTTTCGATA CGTTTTTTTG ATTTCATCTT	3120
	TTGAAGCATC CTTACTAATG CCTAAAACCT CATAATAATC TCTTTTGGCC ACAGCTATCT	3180
	CTCCTTTTCT TAATTAATC ATATAGTTTA ACGTAATATG TCATACTATC CAAATAAAAA	3240
50	GCCAAAGCCA ATGTTCTATT GACTTTGACT TTTCAGATCA TGACAACATT CTAATTGTAT	3300
	TGTTTAAATTA TTTTTTGTCG TCGTCTTTTA CTTCTTTAAA TTCAGCATCT TCTACAGTAC	3360
55	TATCATTGTT TTGACCAGCA TTAGCACCTT GTGCTTGTG TTGCTGTTGA GCCGCTTGCT	3420

	TATCTTCTAT ATCTTGACCT TCTAAAGCAG TTTTAAGAGC GTCTTTTTTC TCTTCAGCAG	3540
5	ATTTTTTATC TTCTTCACCG ATATTTTCGC CTAAATCAGT TAAAGTTTTT TCAACTTGGA	3600
	ATACTAGACT GTCAGCTTCG TTTCTTAAGT CTACTTCTTC ACGACGTTTT TTATCTGCTT	3660
	CAGCGTTAAC TTCAGCATCT TTTACCATAC GGTCGATTTC TTCGTCTGAT AATGAAGAAC	3720
10	TTGATTGAAT TGTAATTCTT TGTCTTTTAT TTGTACCTAA GTCTTTTGCA GTTACATTTA	3780
	CAATACCGTT TTTATCGATA TCAAACGTTA CTTCAATTGG AGGTTTACCA CGTTCAGCTG	3840
	GTGGAATATC AGTCAATTGG AATCTACCAA GTGTTTTATT ATCCGCAGCC ATTGGACGTT	3900
15	CACCTTGTA TACGTGTACA TCTACTGATG GTTGATTATC TACTGCTGTT GAATAGATTT	3960
	GAGATTTAGA TGTAGGAATC GTAGTGTTAC GTTCAATTAA CGTATTCATA CGTCCACCTA	4020
	AAATTTCAAT ACCTAAAGAT AGTGGTGTTA CGTCTAATAA TACTACGTCT TTAACGTCAC	4080
20	CTGTGATAAC GCCACCTTGG ATTGCAGCTC CCATTGCCAC TACTTCGTCC GGGTTTACTC	4140
	CTTTGTTAGG CTCTTTACCG ATTTCTTTTT TGACAGCTTC TTGTACTGCT GGAATACGAG	4200
25	TTGATCCACC AACTAAGATA ACTTCATCGA TATCTGAGTT TGTTAAGCCA GCGTCTTTCA	4260
	TTGCTTGGCG TGTAGGTTCC ATTGTTCTTC TAATTAATGA ATCTGATAAT TCTTCAAATT	4320
	TAGAACGAGT TAAGTTTACT TCTAAGTGTA ATGGACCGTT TTCACCAGCT GAGATAAATG	4380
30	GTAATGAGAT TTGAGTTTGT GATACACCTG ATAAGTCTTT TTTAGCTTTT TCAGCAGCAT	4440
	CTTTCAAACG TTGTAATGCC ATTTTATCTT GAGATAAGTC TACGCCATTT TCTTTTTTGA	4500
	ATTCTGCAAC TAGGTAGTCA ATAATTACTT GGTCAAAATC ATCACCGCCA AGTTTGTGTG	4560
35	CACCGGCTGT TGATAGTACT TCGAATACAC CGTCACCTAA TTCTAGGATA GATACGTCAA	4620
	ATGTACCGCC ACCTAAGTCA AAAACAAGAA CTTTTTCATC TTTATCAGTT TTGTCTAAAC	4680
	CATAFGCTAA TGCTGCAGCT GTTGGTTCAT TAATGATACG CTCAACTTCT AAACCAGCAA	4740
40	TTTTACCAGC ATCTTTAGTT GCTTGACGTT CAGCATCGTT AAAGTATGCA GGTACTGTAA	4800
	TTACAGCTTT GTCAACTTTC TCACCTAAaA TAgtTTTCAGC TGTATTTTTT AAGTTTTGTA	4860
45	AAATCATAGC TGAGATTTCT TGTGGTGTGT ATGATTTACC TTCAATATCT ACTTTATAAT	4920
	CAGTACCCAT ATGACGTTTA ATAGATTGAA CAGTGTTTGG GTTTGTAATA GCTTGACGTT	4980
	TTGCTACTTC aCCAACTTGA GTTTCTCCAT TTTTGAAAGC TACAACAGAT GGTGTTGTAC	5040
50	GTGAACcTTC AGGGTTTTGA ATTACTTTTG GCTCATCGCC TTCTAATAcT GTnACACATG	5100
	AATTTGTTGT ACCTAAGTCT ATACCAATAA TTTTACTCAT AATAAAATTC CTCCATTTAA	5160
55	TCATTAAATT AATTTAATTT TAAACAATGT CTTTTCGCCA AATTTAAGTT ATTGGTTTAC	5220

AGTGATTTCG CCAGATTCAA AATCAGGGTT ATCATCTTGA ACTACAGCTT GGTGAATATT 5340
 TGGATCAAAT GCTTCACCTT CAGTTTTAAT AACTTCAAGA CCATTATCTT TTAGTGCCTT 5400
 5 AATCAAACCTT TCATGCACCA TTTGTACACC TTTTGAAGA GATTTAAAAG TCTCATCATC 5460
 ACCTTCAATT TGAAGTGCAC GTTCTATATT GTCTATTGCT GGTAATAATAT CTGTTAACAC 5520
 10 ACGTTGTGCT TGATATGTTT TGTTTATTTT ATTTTCTTTT TGAATTCTAC GCTTATAATT 5580
 TTCAAACTCA GCGTAGAGCC TTAAATATTT CTCTTCGTTT TCATCTGCTA ATTGTTGAAG 5640
 TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC 5700
 15 TTCTATTGCT TCATCTTGTA AATGACCTTT ACTTTCTTCA GCTTGTTCAA CTGAATCATC 5760
 AATATTTTGT TTGACGTTTG TTTCTTCAAC TGTTGATTCA GTGTTTTTTT CAACTGATTC 5820
 GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT 5880
 20 TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC 5940
 CTTTAAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT 6000
 CATCAATTTT ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTATAATT 6060
 25 CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT 6120
 CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG 6180
 30 AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT 6240
 GAACAAAAGA GACAATATCA TCTTGTAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTA 6300
 CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 6336

35 (2) INFORMATION FOR SEQ ID NO: 220:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13059 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

TTCATGATTA TTATCTGTTG TAGACACTGC TGGATCTTCC GATGTATCTT TCGATGCATC 60
 TTTCGATTTG TGTATTTGCT GATTCAAATG GTCTAGGTCT TCTAACGCCT TATTTACCAT 120
 50 TGCTTCATCA TTTTATCAT CTTTTCTCC ATGTTTTGTT GTAGCCGTTT GTGACATATC 180
 ATTTTTCATT GCATTAAGAT CGTCCTCGCC ACTTTGTTGA CCCCTATCAA CATTGAAGA 240
 AACCTCATTT AAATCTTTAA GCAATTGATC TAATTTACTG TCTATATCAC TTTGACCGTT 300

55

	TTCATCTATT	TGCGATGCTG	TTTTCGCTTC	ATTTAGTTGT	GCTTTATAAT	GTTCTTTAGA	420
5	TGAAGCCGAT	AACTGTTTTA	ATTGCTCAAT	TTGACGAATT	GCCTTGTCAA	CTTTGTCTAA	480
	TAAATCTTGC	TTAGATAATA	TCTCTTTTGT	AATTTTCAGTA	TCCTTTTCAG	ATGCAGCTTG	540
	GGCATCGTAC	GGCAAGATAT	TCGTTAAAAAT	GATACTTGTC	GCCATCATTG	TCGAACACGA	600
10	TAACTTTACA	TATAATTGAA	ACGGTTTCCC	TCGATATTTA	GCCATCAACA	TACTCCTTCC	660
	TCACTTACTT	CCTTCAAAGA	ATTACATACT	ATTATATAACC	TGTTTACAAG	AAATTTACAC	720
	TTATCTATCT	AGTTATTGTT	GTTAGTAATT	ATCTACTTAT	TACTTAGCTT	ATATTTAAGT	780
15	AAACAAAACA	AGCATGACGT	AATATCATAT	TGTCCATGTC	GCTAACATCA	TATTACGTCA	840
	AATCTTTTAT	ATTAAATGAT	GTTTTATTTT	AGACTGCTTT	TTCTTTTAG	CTTTGAGCG	900
	CCTGTTTAAA	AACTTGCTCG	AATTGTTTAC	GCGAGATTTT	GTGTGCATGT	GCTTTTTGTG	960
20	CTAATAAAGC	ATCTCGAAAC	TGTTGTTGAT	CTTTCAAAC	TTCTAACATT	TGTATTAATT	1020
	GGTCTTTACT	TTCCATTGTT	ATCTCATCAT	TATGCTCAA	TAAGTGCTCT	GATAATGTTA	1080
25	CTTTAGCATG	GTGTGCGGTT	TGACGATAAC	CTAAATCAA	CAACTCATAG	TCAAACGCTT	1140
	GTTCCACCGC	ATTTAAAATT	TCATTACCCT	CATTGATATC	AAGATAAATA	TCACATAACT	1200
	GGTATAGTTC	ATTTACCCTG	TCAATATTAA	TAGATGGGTA	TAAATGCACA	TTAGCATATT	1260
30	GATCAAGTTG	CATTAGCTTA	TCAGACATCT	CTGTAATAGC	AGCGATGTGn	AACTTAAAAT	1320
	CTGGTAAAGt	TyCAACCAAT	ACCTTGATGT	TACGAatTGa	TCCgAGTTAG	TTAATATTAC	1380
	AATTTCTTTA	GTATATCTAT	TACGACTACG	ATAGTTATAT	AGATATCCGC	CTTGTAATAAT	1440
35	ACGAGATTGA	ACCTTTGCGT	CTGCTATATT	GAGCATCGTT	TCATATTTCG	TTTATCTGG	1500
	AATAATAATA	TTACAATGTC	GTTTCATATC	ACCTTTACAC	ATCAATTGCA	TATTTCCCGG	1560
	GACATTACCA	TTACAGTGTT	CTTGCCATAC	CAAAACATCA	CTACCTTTTG	ATGGCAAATT	1620
40	ATATAACACT	GAAAATGGTA	GGGCTAGTGA	GTTAATAACG	AAATGATGTT	CCGTAATTTT	1680
	AAGTTGCTTG	ATAAAAAATA	ATGCGAATGC	GAGCTTTGAA	GGGAAAAAGT	AAGACTTCCC	1740
45	TTGCCAATCC	AATATGACAT	CAGATGTTAC	AAAATTTTCA	TAAATCACTT	CTTTACCTTC	1800
	TGCTGTCATA	TATTTCTTCA	AGATCGCTTT	ACGATTTAAA	TCGTAAACAG	TTGTGCAAA	1860
	TTTAATACCA	TTCTTAGAAT	AATAATCGAC	AAATCGGACA	CGTTGTTGGT	CATCAAACCA	1920
50	TTGACACGA	CTAACAATT	TAGGGCGCTC	TCCACTTTGA	yAAAAATATT	TACCTCGyAG	1980
	ACGTCCCAT	TCATTaATTG	TAGCCGAATT	GTTGTTACCT	TTAATTTCCC	AAAAAGCTGG	2040
55	TACAGTAACC	TGATTAAAAA	ATCGTGGTTT	CATATTTTCT	GTATTATGAT	TATCTGCAAA	2100

	TAAATCTTCT	TCCAACCTTAC	TGGCTTTTAAA	AGACTCATAT	AACCTTCGTG	AATGATCGTT	2220
5	AAAGTAATCA	AATAATTTAA	TCATGTAGCA	CCTCTTGaAC	TAATGTTTCC	CATTTTAAAA	2280
	TAATATCTTG	AGTCATAAAT	TGCTGTGCCA	CTTCATAAGA	GATGTCATGT	GGTGTCTGGG	2340
	GACCATTGTT	AAAATACATT	ACAATGGcAT	GAGCTAGTTT	TGCGATAACA	TCATCCACAC	2400
10	TATCTTCGTC	GGTATCAAAA	GGTACCAAGT	AGCCATTTTC	CCCATCTCGA	ATAAAGGTTG	2460
	GGTTACCATA	ATTCACATTT	AATCCAATCA	TACCTAGTCC	TGAGCCTACC	GCTTCCATTA	2520
	GTGTTAACCC	AAAACCTTCG	CTAGTTGATG	CAGAAAGAAA	TAACTCATAA	TCATTATAAA	2580
15	TTTCATCAAG	TTTAACATGC	CCTAGTAAAC	GAATATAATC	TTGTGCGCGG	TGTGTATCAA	2640
	TAATTTTACG	CAGTCGCGTC	TTCTCGCTAC	CTTCTCCATA	AATATCAAAT	GTTAATTCTG	2700
	GCACTTGTCT	TTTAGCCACG	ATAACCGCTT	TGACAAGCCA	ATCAATATGT	TTCTCATTCG	2760
20	CTAAACGAGA	TGCACTAATC	ATCGCATATG	GCTTTCTTGA	TAATGTTGGA	TATGATAATG	2820
	CATCAATGCT	TCCCACAGGA	ATAGTATAAA	CACGTGGGCG	ATAACCTTGA	TATTGCTCAA	2880
	ATTGTCGACA	AACCATATGA	TTTTGAATAT	CTGTTGCTGT	AATAAAGAAA	TCAATGTATT	2940
25	TAGCTTTTGA	AAATTGATAT	TCATAATAAT	TGTTCCATAG	TATATGCTGC	TCACTCATCA	3000
	TATTATTACT	ATAATGATCA	GCATGAATCA	CAACACCAAC	TTTACTATCA	CCTTTATGCT	3060
30	GCAAAACAGC	CTGACCAATA	TCAGAAGCGC	GGTCTAATAT	GACAATATCG	TCTCGGGTTA	3120
	AATTCAATCG	TTGTAAAAAG	TATGCAATAA	ATTCCGTTTT	GTTATACAAC	ACCGCATCTT	3180
	CAACACATA	TATAGAGCTG	TCTCCATCAA	TATATTCGTT	ATAAGCGATG	GAACCATCTT	3240
35	GATTATAAAA	TTGTCGCATA	TATAATTTTC	CTTTATTATC	AGCTGGTGCA	TAATACTCAG	3300
	AAAATATGCG	CGTATAACTA	TAAAAATCTT	TACGTACTAA	CATACTATTA	ATTACAAATT	3360
	CTGCACGATC	CACAATATCT	TTTTGTTTAT	TTTGAGATA	ACATGTTACA	AATGATGATT	3420
40	TCCCATTAATA	ATATAGGCGG	ACTATCTTAC	CATTTCTTTC	TCTAAACTA	ATGTCATGAC	3480
	CAAGCTCACG	TTCAATGTCA	TCTAACGTGT	ACGTTGTTGG	TGCTAAAGAA	ATATCACTAA	3540
45	AATACTGATA	CAACCAAATA	ACTTCTTGAT	CTTTAAACCC	AATGTTTTGC	GTTAATGTCT	3600
	GTATGTTCTC	TGACTGTATA	AAATCTAAAA	ACACAAATTT	AGTGTCTTGA	TTGTACGTC	3660
	TCAATAATTT	AGCACGGTAA	GCTTGTGCAT	ATTCAACACC	GCTACTCGCC	CAGCCTATAC	3720
50	CAAAGTTTAT	ATTATATATT	GTGATGCGCT	ACCCCTTTTC	ATTTATGGAA	AATGTATAAC	3780
	TGGCATACCC	TCTTTATCAA	ATGTAATCAT	GCTTTGACAA	ATATTTTTCA	CCATTCTTTT	3840
55	TTTGATATTT	CGTGTCAATA	CTTCAAATGA	ATCTAAGGCA	ACTCTATGGT	ATTCAAAAAT	3900

	GACTTGTTCT AACCAACATG AATCAATTGC TTTCAAAAAG ACTTTTTGAA CGAAAATATT	4020
5	ATAATAATAT GCACTTTGCA TGTTTTTACG ATTCAAAGCT AATTGCTTTT CAAATTGCTC	4080
	TAATAAAAAT GTCACACTG CTTGCTTATC TTTAAAATTA ACACAAGCCA CATCTTTATT	4140
	AAATTGGAAA CTTAAATTTT GATAAATATA CTCGACAACA CGCGATTTTG TTAGCACCTT	4200
10	TTCTCATTT ACAACATTT CAAATACATC TTTAGCTAAC GCTTTAAAAT CTTGATTCTC	4260
	AGCATCATCT ATTTCTAAAA CTCGATTGCG TTCCTCGTAT ACAAGATCTC GCTGTATACT	4320
	AATGCTTTTT TCAAATTCAT TAGCCATTTT ACGAGCTTTA ACCCCTTGTT CTTCCGAGAT	4380
15	aCGcTGCGCT TTAAC TACAA TTTGCTTAAC TTTGCGATTA AACAAATTAC TTTGCGATAA	4440
	TCGTTGTGCA TCTAATGAAT ATAATTGATT ATTTTCCGCT AAATTACTAT CGCTCCATCG	4500
	CTTAACTAAA TAATCATCTA GTGAAATATA TATACAAGAT GATCCCGGAT CCCCTTGCTT	4560
20	ACCAGAACGA CCACGTAATT GCCTGTCTAC ACGGCTATTT TCCATATGTT CATGAATAAT	4620
	AACAGCTAAT CCACCTAATG CTTGACACC TTCACCAAGT TTAATATCTG TGCCTCGACC	4680
25	TGCCATACTA GTCGCAACAG TCATGGAACC AATTGCCCCT GCTTCAGCTA TCATCTGCGC	4740
	TTCTTTTGCA ACATTTTGCG CAATGAGTAA ATTATTAGGA ATATCCATTT GGAATAATAC	4800
	TTTCGAAAAG TATTCAGCCG CTTGAGCAGT TCTCGTTATG AGTAAAACCG GTCGCCCCGT	4860
30	TTTCATGAAGT TCAACTATAT CATGAATCAT CGCGATGTTT TTCTCATCAA CTGAACGAAA	4920
	CACTTTATCT GGTTCATCGA TACGTTGAAT CGCTTTATCA GTTGCTACTT GTACGACTAT	4980
	TTTTGAATAC AAATCAAAGA ACTCTGATTC GCCTAATTTT CCTGTAGCTG TCATACCTGA	5040
35	AAATGATTCA AAAAGTTTAA ATAAATTCTG GAAGGTAATT GTTGCCATAA CACTTTTATC	5100
	TGTTGAAACC TCCATACCTT CTTTCGCTTC AATAGCTTGG TGAAGTCCAG CTTGCAACTT	5160
	AGTTCCCGGT AACATACGAC CTGTAATACG GTCAATTAAA ACAATATCAC CATTATATAC	5220
40	AAAGTAATCG ACATTAGATT CAAACAAATA TTGTGCGCGC AGTGCTAAAT TAATATTACG	5280
	CACTAGGACC ATCGCTTGTT CGCTATATAA ATCTTCAACA TTAAAGTATG ATTGTGCCGC	5340
45	TTCAATACCT TGATTTAACA GCCATATTTT TTTTTGGTC TTCTTCATTT TAAAATGCAC	5400
	GTCTTCAATC AATGTATCTA CAAACTCTTT CACAATATGA AATAGATTTG ATTGTAATCT	5460
	TGGTGACCCC GAAATAACTA ATGGTGTTTG AGCAGCATCT AAAATGATTG AATCCACTTC	5520
50	ATCAATAATA CCGTAATTTA ATTGTGGTAA AAATTTCCCT TCCGCACTAT CAGCCAAATT	5580
	ATCAATTAAA TAATCAAAC CGAGACGTCC ATTAGTTGTA TATATAATAT CATGTTTATA	5640
55	TATATTACGT TTTTCCCCTT TTTGATACTC ATAATCCACA ATATCAACAA AACCTAATGA	5700

	TAATCATTCTG TTGTAATTAA ATATGTTCTT TTTCCCGAAA GAGCATTTAA ATATAAAGGC	5820
	ATCGTTGCCG TTAATGTTTT ACCTTCGCCT GTTTCATCT CCGCAATGTT ACCTTCATGC	5880
5	AATACAATCG CTCCGATTAA CTGAACTTCT TTAGGATACA TACCTAATAC TCTCCAGCTC	5940
	GCTTCACGTG CCACTGCATA AGCTTCAGGT AACAAATGTAT CTAGTGTATC AACTCCTGAT	6000
10	GCTAAACGTT CTTTAAATTC TATTGTCTTT TGTTTTAAAG CATCATCAGA ATATGATTTA	6060
	ACTTCATCGC TCCATGTATT GaTGSgTTcA CTATTTTTCT AATCGACTTT AGTCTTAATT	6120
	CGTTTTATCGT AACATCTAGT TTATGTTTCA TTTACTTCCC CACCATTTCAG TTTCGATACA	6180
15	TCTAAGTAAT CTA AAAAATCG TACTGGATTG ATTAAACGTG ACATATAATT TAGATGTTTG	6240
	TCTTGCTCTT CTTTAAATA AACCTCGACA TTTGTATCTT TTAGTTCATG ATTTCTGGG	6300
	ACATGTTCTG TAAGCCATCC TTTTAAATCA TCATCTTCAT GGCTTGACG ATACACTTTG	6360
20	CAACCCAAAT GCTGAGCGAC ATAAGTTGCA AAAACATTTG ACTTTGACCC ATA ACTAATC	6420
	AAATTAATAG CCTTTAGGGT ATCTTGACTT TGCAAATCAT TCTTTAGTTG CTTAATATTT	6480
	CCCTCGATAT TGTCGTCCAT CCAACGTTCA ACGAGCCAAA CATGACCAA CAGTTTCAAA	6540
25	AAATCATTCTG AAATAGTTGG ATAGGTGTCA GATGGTTCTG CAATAATGAC ATTGATCATA	6600
	TCATTTCCAT ATTGGTCATC GCCTATCTTC GTCACCCGCA TGCTTTTATA CTCTAAATCA	6660
30	TATTGATGCG TCATCTCTGT GATTGTTAAA CATCTAAATA TAAGACTCGT CGATGCTGCA	6720
	TTCATCATTT TTATTTTATA AGCATAGGCK TCATCAGGAT ATTGAATCGT AATACTATTT	6780
	GACTTTACAA TCTCAGTACT TAGTTTTGTG CCATTTTTAT TATAAAAAAT GATGATAAAA	6840
35	TACTACTGAAC CAGCAGGCGT TGCATCAAAA TCAAATGCA ATTTATAATG CTGTCCTCTA	6900
	CGCAAAATTG GKAAACTTGG CGcACTTTTA TATTTTGAAA ATTGCTTTAA CATCAACCAC	6960
	TCATGAATCG GTAATCCAGA GGGCATCAAA GGATTTATAA AAGTCACTTC ACCATTTGAA	7020
40	AATGATACTT TAGAGCCATA CATAAATGTA GTTTGTGAAA TATAATTCCA AGTAACTTTA	7080
	AATGTTTTGT TTTTCAGCAT GTTGAACCTC CCCAACTTG TCTTCCAAA TAATGTTGTA	7140
	AAAATTAACA AACCAACTTG CAATGGTAGG TGAATCATCA TTATGTCGCC CAGGAATACT	7200
45	GCGATTCATC ACTCTTGCTT GGTGTGCTGT CAATACAGGT AATAGCTCTT GAAATGCATG	7260
	TGGATCATAA TCATCATGTT GCATATATGC TATGGCAAAA ACAGTTTGTG ACAATGATTY	7320
50	CTTTTGAAAT GTTTGCCAAA ATTTTIGATT TAATGCCTGT ATCGACGCTT GAGATGTATC	7380
	ACCTTCATTA GACACCAGGA CGTCTAATGC TGTACCGAAC TCTTCTGGTC TAAGTAATCG	7440
	CATATGTTCA GCAATCGTTC CAATATTAAC AAGTGGTTTA CCAACAATAA TTGCCTGAGG	7500

55

	TAATTCATGT GATTTAAAAT TCAGCTTTTC TAATGTCTCG TCAATAACAT TGATAATACC	7620
	TTGTTTCATAT TCAGATGAAC CGATATAAAA ACTACCACCT TCAACACGAG GATCGCCGAT	7680
5	AAGTAAAAAC GGTGCATTCA TACGTTTCAT CATATAATAT CCTTCGAAAC CTTCGGCTGT	7740
	TCGATAACCA CTAAAATATA CGTTTAGTGG CGGTTTCATA TCACCAGGGT GGAAATAATA	7800
	AATAAATTCC TGTCGTTGAC TATCTACGAA ACGACTACCA CCAAGTAAAA ATTGACCCAT	7860
10	GTCTAATCTA GACCATCGTT TGTGTATAGG TCCTAAATGT ACCGTCCCGT TCCCACGCGC	7920
	CTTAACAGTT ACACTTATAT AAGCATCAAA TGGTTTCGCA GGTATCTCTA AAGGACTGTC	7980
15	TAACATATCA TCAGTCAATA CGATTTGTTT AATTAATGCA CCATCAGCGC CAGTCTGAAT	8040
	CAATCTAAAT GSTATATTGCA ACTCGACCGC ACCATCAATA TCAAATTCTG GCCATATTTG	8100
	AATGACTTTA TCTTTATCGT AAACGAGATT ATTTTGCCAA GATGCGATAG GTTTAAATTC	8160
20	TTTCCCAAAT TCTCCACTCA ATGTGAGCTC TGAATTACCT TGGTAAACGA CATCTCCTTT	8220
	AAAAATTCGGA TGCACAAGTG CTAACCTAGG AGAAACCTTA TCTCCATACT GTCCTGAGAA	8280
	GCTAACTGCC TCTAATTTAT TATTACGTTT TTCAATATTC CGGTAATGTA ATGGTTGAAC	8340
25	AACGTATTTT TGGACATTTT CGTCTTGTTT ATATTCAACT GACCAAAATG ATTCATCAAC	8400
	ATACGTATTG TATGGTTTCG TATCATTTG TAATAAATTC GTTAATGTCT CCGAGTATGG	8460
30	TGCTTGAATA TAGATAAAAT CAAAGCGCCC TTCTGCTTCA ACAATCGCTT CAATAGCCTC	8520
	TACATAACCA CTATCAAATT CAAACAATCC AATATCGAAG TAATCCCAAC TCACACCTTT	8580
	TTTGTGTTGA AAAATAGGTT CTAAATCGTC TCCTCCAATT TGCAAAACTC TAAATTTACG	8640
35	TGGCATCATT TTCACCTTCT ATTAATCAT CGAGCTGATT AATAATATTC TTAGAAGCAT	8700
	ATGCATCTAT TAATTTTAAA GAATAGGCGT ACGCATAATT CCAATTTTTC AAATAAAATA	8760
	AATAATAATT TAACGCATCA TCTAATTCAT CAACTGTATT TATAATACGG CCATTGTCAT	8820
40	AATCAGAGAC GTAATCTGTT TGTTGACCAT TAATTTGTGG AATCCCAGCG CTAATTGCAC	8880
	TAATTTGTAA ATACAAGTCA GGTTCCTTTG ACATATCTAT CACAAGTCGC AACGTCCGCA	8940
	ATGCTTCTAC AACATCATGT TCAGCATGTA TCGTCTTAAC AGCAATGATG TCATCTTGAT	9000
45	CTTCAGGTGT CATTAAATGCT GAAACATTAA CATCCGCATT CTGTTTAGCT TGGTATTCCT	9060
	CATTTACCGA CGTAATACAT TCACGAAGCC ACATCGGTAT GTCATTTTGA TGGCGCGATA	9120
50	ATAAAATTAA ACGGTAATAA TCTTCCTGTG CGATATAATC CACAAGTCGT TGCATCATTT	9180
	GTTGCAAATC AGCGTCACTC ATACCATCTA TCCATACACC TATAAATGTT TCCATCAATT	9240
55	GACTACTTAT ATTAGGTGAT TGTCTCGTTT CAAATGGTGT GATTGGAATC ATTGTATTCT	9300

	TTAAATGGGC ATTCTTTACG ATAGATTGAT ATTCCTCATC TGACACAGTT TCATTTCTAT	9420
	TTTTAAAAA TGAATAACTT AATGATTTTCG CTGGAATATG ATTGGCTATT TGTCGATTGT	9480
5	GCCTAGCATC TGAAGCCACA ATCACATGAT CATCTTCATG TATTTGTTGT GCAATCATTG	9540
	CTTGAAATTT TTCTTCAATT AGTTGAGCCA TATTGTTATA TTCTGTTTGT TGATAGTGAT	9600
10	GTTGATATCT TTTTGAAACA GTGACTCTGC CATTTTTCAA ATCTTCATGA AGTACACAAT	9660
	CTCCATTAAT CGTTAAATAT TCTTGGTAAG AAGCCTCTCC CTGATCATCA AAATAACGTA	9720
	TCGCTGATAA ATAACCTCTG TCATCAAAAA TATAACGCCG TTGTAAGTGA TCTCTTTCAA	9780
15	ATTCTTCAAA CCAAATTGAA TACCCTTCTT GACTAAAATA AATATTTGTA TAGGTCTGTT	9840
	CACTCGTCAC ACATTTTAAAT AAATACGGTG TGTACACAAA CTCAACATCA TCCGGCCATT	9900
	TTAAGTGATG ATAATTAATC GCTTGTGGCG CATGGTGAAT GAATCCTTGA ATTTTCATCA	9960
20	ACACAGACGA ATACTTTGTC TCATATAAGT CATATCGATG TAAAAATGTT CTAAATTTG	10020
	GTGCATGATT GAGAACAATC AGTTGATAAT CTAAGTCATT TTCAAGGTGC ATTCCCATT	10080
	AACTAATCAT ATCGTCAAAT TCCGTCTTAT TTTGTAGTTG ATAATACGGC ACAGTCGTGT	10140
25	CTTGCCACCA TCGTTGGTCA TCGTACCAAG CTGGAATAAA GTATTTTATA ATTACCTCCT	10200
	TACCAATACT GGTTTAAAAA TGGCTTATAT TTATCAAAAT ATAAATATGT ACGAATTGTT	10260
30	TCTGCAATAT TAATACTGAT GTAAACTAAT ACAATCAGTT GTACTGAGAA ATAAATTTCA	10320
	GTAGATAAAT GCGGTACAAA CAATGTGAAA TAAAGCGGTA TACCAATAAT GACTGTAAT	10380
	AATGCCAATC CAAACCAACA TACGCGTCGT GCTTGATAAT TTAAATAACG TTCTGTATCC	10440
35	TTACCAGGTT TAACTCCTGA AAAATAATTG CCACTCTTTA AGAAATCTTT GGATTTTTGT	10500
	TTAGTATTGA TTAAAAATCT CGATAAAAAA TAACCAATA ACATTTGAAT CACTAAATAT	10560
	ACTGAAATAC CTACTGGACT ATCAAATGTC AGCATTGGCA TGTCTCTGA TATGCTTTTA	10620
40	TTAAACATAG ATAAATAAAA ATGAATGCCA CTTTTTAAGA AAACAAAAGC TGAAATACTC	10680
	ATCATTAAG TAATACTGCC TGCAGGGTTA ACTTTCCAAG ATAAATAAGA TTTCATATTT	10740
	GTTGCGGAAA CGTTCATTAA ATCGATATAT GGTATTCTCA CTTCTACTAA TTCAATAAAT	10800
45	AATAAGATAA ACAATGTGAT TATCACAAGG ATGATTAACA ACGCAATCAC AATATGACTT	10860
	GCATCTATAT ATTCCATTTT TTGATGCATC ATTGATTTAA TAATACTAAC CATTACAATC	10920
50	GGCAITGGTC CTGCGATGCC GTAGCGACTA TTTTGTGAG CTAACCAAAC TAATAACATC	10980
	GTTCCAGTAA CCAAATCAA TATTGTTAAG TAAATATTGT CTTGATGAAC ACGTTCTTTC	11040
	GAAACATATT CATGAATCAC AAAATAACTT TGAATAACAC TTAAATTTAA TGTTAAGATG	11100

55

	GAAATCAGCA TCAAGATAAT CATTGATGTT AACCACGGAC CTAAtCCTAA AGTGAAAATG	11220
	TTTAAAGTAT TAACGTCTCC ACCCATATTA GAAATAGCTA TTTTAAAAAA TGACTCATGT	11280
5	TTTACTTGCA TATCGTTaTA GGAAACGATG GAAATGTTTG TGCCTAATAT ATAAATAaAC	11340
	AAGATAAAAC ATGTGTATAG CATACGTTTA TATATAATTT TATATTGTA TTGTTGTAAA	11400
	AGTTTTAACA TGTTGCACCT CTTTTATATC AAAAACATTA AAAAGACTAA GGGTTCATCA	11460
10	CTAATTATTA AAATCCTATA TCGATTTTTT TAGTGATTGG TGCCTCAGTC TTTTTAATTT	11520
	TAGCCAGCTA TAAATTCAAT TTATGCTTGA GAATCATCTT GATCATTTTT ATCTTTCTTT	11580
15	TTCTTTCTCT TCATTAAACC TAAACCAACT AATAATGTCA TAACGCCACC TAGTAATCCA	11640
	TTTTGTTTTA TTGAGTCACC TGTATCTGGC AATCTTTTTT CACTTTGTGC TGGTGTGCCA	11700
	TTATGTTTAG TCACTTCAGA TGTTGCACIT AATGTAGACT GAGATTCACT CGTGCTCGTT	11760
20	GTTGCTTCAC TTGATAAGCG AGATGTGCTC GTGCTGTGAG TATGATGCAT ACTCATTGAG	11820
	TCTGACGGAT GCATTGAGTT AGATTCAGAT GTACTTGTTG AGCCGGACAT ACTTGTTGAT	11880
	GTTGAGTCAG AAATGCTTTG TGAACCAGAC ATAGATGTAC TCAGTGATTG GGATGTGCTT	11940
25	GTCGAATCGG ATGTGCTCAA TGACGTTGAT GTGCTTGTTG ACACTGATTG TGAGTCACTA	12000
	ATTGATGTTG AGTCGGATTT GTCTTGTCAC ATTGAAACAC TCGATGAATT AGATTCACTC	12060
	ATTGATGTTG AGTCAGATAC GCTCGTTGAA CCTGAACCAG ACGTACTTAA TGATTGAGAT	12120
30	ATGCTTGTTG AAGTTGAACC ACTTGTTGAG TCCGATGTAC TTGTGATGT CGAGTCTGAA	12180
	TCTGATGTAC TCAATGATTG TGAGTCACTG ATAGAAGTTG AATCACTTGT AGATTCTGAT	12240
35	TCTACTGTAC TTTGTGAACC ACTGATACTT ATTGAAGTAG AATCACTGAT ACTGTCTGAT	12300
	GTTGATAATG ATGTCGACAC CGATGTGCTT TGTGATGACG ATGTACTAGC ACTCATTGAC	12360
	ATTGATGTTG ATATCGATGT ACTTAAGGAA CCAGATGCAC TTGTACTTGT TGACTGGCTT	12420
40	TGTGACATTG AATCACTTAA TGATGTAGAT GTGCTTGTTG AGCTCGAGTC ACTTACACTT	12480
	GTTGAACCTG ATATTGAGTC ACTTAAACTT GTCGATGTTG AAACGTGatwC GcTTCCGCTC	12540
	ATTGAGTCAG ATGTTGAAAG TGATGTACTC GTTGAATTTG ATCCACTGAT GCTAGACGAA	12600
45	TCACTTGTA ACATTGAGTC GCTTTCTGAT GCACTGATGC TCATAGAGTC AAATTGACTA	12660
	TTACTTGTTG AGCTTGACTG CGAATCGCTC ACACTTGTTG ACGTTGATTG TGATCCACTC	12720
	ATACTTTGCG AGCTACTCAA TGATTTTGAA TCACTTAATG AATCCGAAGT GCTAAGACTT	12780
50	GTGGAACCAC TTAAAGATAT TGATCCACTT AATGAGTCGG AGTCACTTGT ACTAGTAGAA	12840
	TCACTCATTG ATATTGAATC ACTTAGcGAG GTAGACTyGc tTACGCTTTC TGAACCACTT	12900

55

TTTGAATCAC TTAATGAATC AGATTCACTC ACGCTTTCTG AACTTCTTAG TGACGTCGAT 13020

ACACTTAATG ATGACGAATC GCTTGTGCTT ACTGAATCG 13059

(2) INFORMATION FOR SEQ ID NO: 221:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10758 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

AGGGATGGCC TTACCTAAAA AACCGGGnAA ACCCTCCAAA ACCCATTAAA AGGnTGGnTA	60
CCCTTTAAAA TGGTAGCATT TAACCGCCAC CCGCCAAGGT GGGTGGTTTA TTCTTCCGTT	120
ATTTAAATTA GTACACCATG CAGATTCTGT AGTTGAGGGA TATTTTAACG AAAGCTTATT	180
AGCAACTGAT AAAAAAATAC GTCCTAAGGC ATATATTGCT TCATGGAAGG ACATCGAGCC	240
GGCTAAGAAA ATAGAATTTA AAATTAAAAA AGGTATTAAA TGGCATGATG GTAATGAATT	300
GAAAATTGAT GATTGGATTT ATTCAATTGA AGTCTTAGCT AACAAGGACT ACGAAGGTGC	360
TTATTATCCA AGTGTAGAAA ATATCCAAGG TGCGAAAGAT TATCATGAAG GAAAACTGA	420
TCATATTAGC GGATTGAAGA AAATAGATGA CTACACTATG CAGGTTACAT TTGATAAAAA	480
ACAAGAAAAT TACTTAACAG GATTTATTAC TGGACCTTTA TTAAGTAAAA AATATTTATC	540
AGATGTACCA ATTAAAGATT TAGCGAAATC AGATAAAATC CGAAAATATC CTATTGGTAT	600
TGGACCGTAT AAAGTTAAGA AAATCGTTCC AGGTGAGGCT GTTCAACTCG TTAAATTTGA	660
TGATTATTGG CAAGGTAAGC CTGCACTAGA CAAAATCAAT TTAAAAGTTA TTGATCAAGC	720
GCAAATTATT AAGGCAATGG AAAAAGGCGA TATTGATGTT GCGAATGATG CTACCGGTGC	780
AATGGCAAAA GATGCTAAGT CATCTAATGC TGGTCTCAAG GTATTATCTG CGCCAAGCTT	840
AGACTACGGT TTAATAGGtT CGTATCTCAT GATTACGATA AAAAAGCTAA TAAACTGGT	900
AAAGTGAGAC CAAAATATGA AGACAAAGAA TTACGTAAAG CAATGCTTTA TGCAATTGAT	960
AGAGAAaAAT GGATCAAAGC GTTTTTCaAT GGTTACGCTA GTGAAATCaA TAGTTTTGTA	1020
CCATCTATGC ATTGGATAGC AGCCAATCCT AAGGACCTAA ATGATTACAA ATATGATCCT	1080
GAAAAAGCTA AAAAAATCTT AGATAAGTTA GGTTATAAAG ATAGAGATGG TGACGGATTT	1140
AGAGAAGATC CTAAAGGTAA TAAATTTGAG ATTAACTTTA AACATAATTC AGGTTCTAAT	1200
CCTACTTTTG AACCAAGAAC TGCTGCGATA AAAGATTTCT GGGAAAAAGT TGGCTTGAAA	1260

	AATACGATTG	CTGTTTATAT	GCCATATATC	ACATCTTATT	TCATGACGCG	TGCTATCGGC	1380
	GACAGACCTT	TAGTCGTCCC	GCATCAATCT	CAGAACTTAG	CATTTATTGG	TAACTTTGCA	1440
5	GAAACAGAGC	GAGACACTGT	ATTTACAACA	GAATATTCCG	TTCGTACTGC	CATGGAAGCT	1500
	GTTTATCAAT	TACTAAATAT	AGATCGTGGT	ATTCCAGAAG	TCATCAATAG	TCCATTTGAT	1560
	CTTCGCGTCT	TAATGGATGC	CATATACGAA	CTGAATGACC	ACCAAGATTT	GCGTGAGATT	1620
10	ACTAAAGATT	CGAAAATGCA	AAAACCTCGA	TTAGCAGGAT	TCCTTAAAAA	GATAAAAGGT	1680
	ACGTACATTG	AGTCATTATT	AAAAGAACAC	AAATTGTTAT	AACGAAAACC	ATTAATAGAT	1740
15	TTTTATTTGG	TGATTTCAAA	TCATGAGACT	GGGACAGAAA	TGATGTTTTT	ATAAAAATTA	1800
	TTTCGTTGTT	CCACTCTCAT	GATTTTTTTG	ATGAAACATA	ATTACATGAT	TGATTGCATC	1860
	ATTTTGTTAA	ACAAGTGATT	GCAAACCTGC	CATTTTCACAC	TGAAAATTTA	CATAATAAGT	1920
20	GACGATATTT	TACAAGTCAT	ATACAAATAA	CATATATTGT	TAAATAATTT	TACCTAATCT	1980
	TAACATTAAA	TTTACAATTA	TAAGCGATAA	TCTAAATATA	AAGCTTATTT	GAGGTGAAAT	2040
	AATGGAAATG	TCGGTTACAG	AAGTCATTTT	CTCCTTTTTA	GGTGGTTTAG	GTATTTTCCT	2100
25	TTACGGCTTA	AAAATCATGG	GAGACGGGCT	TCAAGCATCA	GCAGGAGACA	GGCTACGAGA	2160
	TATTTTAAAC	AAATTTACAT	CAAATCCAGT	ATTAGGTGTT	ATTGCAGGTA	TCGTTGTAAC	2220
	TATTTTAATA	CAAAGTAGTT	CAGGTACGAC	AGTTATCACA	ATCGGACTGG	TAACAGCTGG	2280
30	ATTTATGACA	TTGAAACAAG	CCATTGGAGT	GATAATGGGT	GCTAATATCG	GAACAACGGT	2340
	AATGCAATT	ATTATCGGTA	TAGATTTAGG	CGAATATGCA	ATGCCAATTT	TAGCATTAGG	2400
35	TGCATTCTTA	ATCTTTTTCT	TTAAACGCTC	TAAAATCAAT	AACATTGGCC	GCATACTATT	2460
	CGGTTTCGGT	TCACTATTCT	TCGGTCTAGA	ATTTATGGGT	GATGCCGTTA	AACCTTTAGC	2520
	ATCATTAGAT	GGATTTAAGC	AATTAATGCT	TGATATGTCT	ACAAATCCAA	TACTCGCTGT	2580
40	CATTGTCGGC	GCAGGGTTAA	CAGCACTAGT	TCAAAGTTCA	AGTGCGACGA	TTGGTATTTT	2640
	ACAAGAATTT	TATCAACAAG	ATTTAATTAG	CTTAAACGCA	GCAATCCCTG	TGTTACTAGG	2700
	CGATAACATT	GGTACCACGA	TTACAGCTAT	CTTAGCTAGT	TTAGCCGGCT	CAATCGCTGC	2760
45	AAAACGTGCG	GCGCTTGATC	ACGTCATCTT	TAACCTAATC	GGGGTAATTA	TCTTCACAAT	2820
	TTTCTTGCCA	GTTGTGATTC	ATTTGATTAG	TTTGTTACAA	GATTTATGGC	ACTTAAACC	2880
	AGCGATGACG	ATTGCAGTAT	CACATGGTAT	CTTCAACATA	ACAAATACTT	TGATTCAATT	2940
50	ACCATTTGTA	GCAGGTTTAG	CATGGATTGT	TACAAAGCTT	GTCCCAGGTA	AAGATATTGC	3000
	TGATGACTAT	AAACCTCAGC	ACTTAAACAA	AGATCTTGTT	TATCACGCAC	CTGGTGTTGC	3060

55

	AGACATTTCGC GAAATTACAA AAGACGATAA AAAATTGATC AAAAAGCTTG AACAAAAGCA	3180
	TCAAGCTGTT GAAACAATCA ATGATAGCAT TCGAAATTAT TTAGTTAGAA TTTCTACAAA	3240
5	AGCCATTACG AAGGCAGACG TTGAGCGTTT AGCAGTTATG TTTGATGTCA ATCGCTCTAT	3300
	TTTAAAAGTA GCAGAGCTAA CAGAAGAGTA TGTCGCTCAA TTAAAACGCC AACATGATGA	3360
10	AGATATTTCGC ATTACAGAAG ATGCACAACG CGGTATGGAT AAATTATTCA ACCATGTTGC	3420
	TGAGTCATTT GATAAAGCCA TCGACATGTT AGATGTTTAT GACAAAACGA AAAAAGATGA	3480
	AATTGTAGAA CGTAGTAGAG AATCATTTAA TATTGAACAT AAACACGCA AAGGTCATAT	3540
15	TAAACGCCCTT AATCGTGGTG AATGTACAAC AAAAGGCGGA TTAATATATA TCGATATGAT	3600
	TGGTGTTCCTT GAACGTATCG GTTATCATTC ACGAAATGTT TCTGAAGCAC TTGTTGGCCT	3660
	TAACGATGAT GTACCTACAG ATGAAGAAAT TGCAACAAC TAAATTTAAT TTTTACTGTC	3720
20	TTATTTATAT TCATATTTTT TTAAAATTAG AGATTCAGAT GCATGTAAAA AGCCAATCCA	3780
	ACATTCATGG GTTGGCTTTT TTGTTTAGCA AAATTTATTA TCTTAAATCG GCTATAAACA	3840
	CTGATATAAT AATGCTTCAT TAGTATGCGG TAAGCATGAC GGACACTGTT CTCGGAGTCT	3900
25	GACCCCGAAA CGTTTAATAT AACTTTTTAC ACGTCGCCCTT CATTGAAGCG AATTGCCATA	3960
	ACCTTCACAT TATATATAGT TCTTTCCATA TAAATGTCCA AATTTTTAGA ACAACGCAAT	4020
30	AAATAACCAT CCACCTAACT TATCAAAAAT TTAAGTGGAT GGTTTTTCAT TTTCATTTAT	4080
	ATTIATATTA GTGTTAATCC AATCATAGAT TTATCTATAT GCACTGCTCT ATACATTTCC	4140
	TCATTTAATT TGCTTTACTT TCATTTATAT CATTATCAAA AACTTTGGCG TGTCATCGTT	4200
35	ATTATTTTCGC ATCTTTGACA CGTTTATCAT CATTAGGAAT CGCGAATAAA ATTGCGATAA	4260
	ATGCCATGAT TCCCATTAAT ACGTTAACCC AAAGTGCAAT CATCGCACCT GTATGAATGC	4320
	TCGTTGCAGC AACTGCACCA GCATATACAG CACCACTAAT TCGGACACCG AATGCGCCAC	4380
40	CAAGTGATGA AGCCATTTTA TAAATACCTG AAGCAACGCC AACTTTATCT AACGGTGCAT	4440
	TCGAAATAGC TGTATCTGTA GAAGGTGTTG CATAAATACC TAAGCCTAGT CCGAAACATA	4500
	AATATCCTAC GACACAACCTG ATAACATAAA ATATGCCTGG TAAGAATACT AATGAAATAA	4560
45	GTGCAATACC AATGACCACA ATGAATGTAC CTAATAACAT TGGTCGCTTA GAACCCATTT	4620
	TTTGTAAATA TTTTTCACCA ACTCGAATCA TCAATAACAC CATGATTAAA TAAGTAATTG	4680
50	ATAAGTATCC TGCCTGCAAT GCTGTATAAC CTAAACCTTG TTGCACGAAT GTATTGCTTA	4740
	CAATTAATGT ACCTGCAAAA CCGTTTAATA AGAAGTTCGA AATCGTTGCA CCTGTATATG	4800
55	GTTTATTTTC AAATAATTTA AAATCAATAA GTGGATTATC TACTTTTTTC TCAACATTTA	4860

	AACCAAGTGC	TGCACCTTTA	GTAATGACAA	CGTTTAAACT	TAGCAACATA	ACTACTAGAA	4980
	CAATTAGCCC	TGCAACGTCA	AATTTATGTG	TATTGGTAAT	TTCTGATTTC	GTTTCAGGCG	5040
5	TCCCTTTGAT	GAGTAACATT	GAAAGTACGG	CAACGATAAT	TGAGAAGATG	AAAATCCATC	5100
	TCCAACCCAT	AGTTGTGCGA	ACTGCACCAC	CGAAGAGTGA	ACAGATACCA	CTGCCACCCC	5160
	AAGAACCGAT	AGACCAATAA	CTTAAGGCAC	GCTGACGTTT	AGCACCCCTGA	TAATAAGTTT	5220
10	TCATAATGGC	CAATGTAGAA	GGCATAATAC	ACGCTGCTGA	TACACCTTGT	ATAACACGAC	5280
	CTAAATTA	TAATGCCGGT	AAATTCGTAA	TAATAATTA	TGCTGAACCA	ATAATACTTA	5340
15	ATAATAAACC	GATATTCGTC	ATTTTCACGC	GCCCAATTTT	ATCTGCCAGA	CCACCTGCTC	5400
	CAACAACAAA	CATGCCTGAA	AATAGTGCAG	TTAGACTGAC	CGCAATACTA	ATTGTCCCCA	5460
	TGTCTGTACC	AAAACCTTGT	TGTAAATTCG	GTACAACATT	TACAAGTGAT	TGTGCAAACA	5520
20	ACCAAAATGT	AATAACACCT	AATACAATAC	CTAAGATTAA	CTTGTTGCCC	CCGCGATACG	5580
	TTTCATTCAT	GTTAGTTATC	TCCTTTAAGG	TAATCTAAAA	CAACTGTCCC	TACTGCTTCT	5640
	GCAGAAATAA	GTAATGATTT	TTCTGAAATG	TTAAATTTAG	GATGATGATG	TGGGTAAATT	5700
25	TCACCATTTT	CCACCGCTGC	ACCTGTATAA	ATAAAGGCAC	TTGGGCGTTC	TTTAGCATAA	5760
	TATGCAAAGT	CTTCTGAAGG	TGGTTGTGGT	TCACACATTT	CAACACCAAA	ATCAAGGTTT	5820
	GCTTCTTTCA	ACGTCTTAGC	CACGTACTCA	GTAAACTCTG	GATCATTATA	TAATGCTGGA	5880
30	TAATCATCGT	TATATTCTAA	GGTGCAAGT	ACACCATACA	TATCCTCTAA	TCCTTTTGAT	5940
	AAACGTTTAA	TTTCTTTTTC	AATTGTTGCT	TTTGTAGCAT	CTGTTAATCC	ACGTACATCA	6000
35	CCTTCAATTT	CAACAACATC	TTTAATGACA	TTGAATTGAC	CTTTACCGTC	AAATGAACCG	6060
	ATTGTGACAA	CACCGGTTTC	AAATGGACTT	AGTCGTCTAG	ATACAACTGT	TTGTAACGCT	6120
	GTGACGAAGT	AGCTACCTGC	AACAATGGCA	TCATTGGCCA	TATGTGGTGA	TGAACCATGA	6180
40	CCACCTTTAC	CTTGAACCTT	CAATTTGAAG	AATGCGCGTC	CTGTTTGAAC	ATAACCAGGT	6240
	CTGTAATACA	CTTTACCTGT	TTTCATTGTG	CTCATGACGT	GTACACCTAA	TACATGATCA	6300
	ACACCGTCTA	ATACACCATT	TTCAATCATT	GTTTTAGCAC	CACCTGGTGG	TACTTCTTCA	6360
45	GCTGGTTGAT	GTATCACAAC	GACTTTTCCT	GTAACACTAT	CTTTCATTTC	AGCAAGCGTC	6420
	TCTGCTAATA	CAAGCATGTA	TGCTGTATGT	GCATCGTGAC	CACATGCGTG	CATAACACCT	6480
	TTATTTTGTG	ATGCAAAAGA	TAATCCTGTA	TCTTCAGTAA	TGGGTAATGC	GTCAAAGTCT	6540
50	GCACGGATTG	CTAATGTTTT	ACCAGGTTTC	CCTGAATCAA	TCGTTACTTT	AATTCCACGT	6600
	GGTCCGACAT	TCGTTTCTAC	TTCCACATCT	TTACCTTGT	AAAATTCAGC	GATGTATTTT	6660
55							

	ATCATTITGC CTTCTTTAGA TTTTAAAGTT TCAATTAATT GTTGATTGAT ATCCTTCATC	6780
	TCCTTAGTTA CATCATAAAT GATTAATCAT TATTTATATT GCCAACACA GAGATGTTAA	6840
5	CCATTAAATTT TTTGCAATTT TAGCTTTGAA TATAAAAAAT CACAAATTAT GTATATCAAA	6900
	ATTTGTGATT TGTGATCATT TTATGAACTT GGGTAACGTT TTAATTCAAT TAAGTGAATC	6960
10	CCATTGCTAA TCATTTTAAT GTTTAATGCC AGTGTGTCCG TGATATCTAT ATCATATACT	7020
	TCTAATTTG GAAAACATCAT TCGATTAAAG TAATCTATAG AGTCCTTGTC CATGCCATGT	7080
	ATCGTATGAT GTTTGCGCCA AAGATTAAAT AACGCACCAT TTTCTTTATC TAAGGTAAAA	7140
15	TGTTTAATCT TATACATACC TTCTTCCAGG GCATTAAATGT TCAAATGAAT CATTTCCGTC	7200
	GCACGCATAT TCATTTGATT GTCCAACGCT AAGTACGGAT TAAAATGCTT TGCATCATAT	7260
	AACAATATTT GAAAATTTGA ATCAGTCCCC GTGACAATAC ATGTATCATC AGAATACAAA	7320
20	ATATTGCTTG TTAATTTATT AAATAGCAAT GCCGTGAAAT AGACCGGACG TTTTCCATTA	7380
	TATTGATGAA ATAGTTCAAT AGAATTCATA TAATCCCGTT CATTTTTACA ATGACTGACG	7440
	TGCAAATCAT AATTCAACCA ATACCCGATA CCCTCTACTT TAGAACTTAA TTTTAATAAT	7500
25	TGCTCAATGA TGATACCACC TCTAAAATAT TCGCCGTTTG TAATAAATGT ATCACCCGTC	7560
	AATGTATTCC AATTGAGTAA AATGAGTGA CGCTTTAGGC GATGACGATG CATTAAAGTCG	7620
30	ATAAGGTAAT TCGTTTTATT AATAATCATT TGAATCGCGG TTTTAAATTC ATCATCATTC	7680
	ATTTTATTAA AATCAACAGC GTCATTTGAA TTGGCATCAA ATACAAAATG GTCGATGTGT	7740
	GGCTCAAGTC GTTCAATAA TGCTAGATGT CTTTCCGTAG CTTGATCTAA GTGAATGTAC	7800
35	AAGCCACCAT TAGGGAATAA TGCTTTAAAA TAATCAATCA TTTCAATCAA AGACGTGTGC	7860
	AATGTCGTCA CATAAAGTT GAACCTCAAA TCTTTTCTAT GACTGACATG CAGGGCAACG	7920
	TGATGGATAA AAATTTTAAA TGCATCGATA TAATCACGTG AGTCATACTG ATCCAAATGC	7980
40	ATGGTCAAAC TAAAGTTATG ATCTAATAAA AAGTCTAAAC ACAAATCAAT ATCATAAAAT	8040
	ATATTCGAAA TTTCTGCATC ATACGTGAAT GCGCATTTGA GCTTTTTTCAT GATATATGGA	8100
	ATCACATCAT ATGCTAATAC TTCATTGACT TGAAAATCAT GATGACATGT AAGCAACTGT	8160
45	GATTGATACT GTGTATTGAG CAAATTCCTC AAATAGCCCA CTTGAATAAT ATGATTAAAT	8220
	TGATTTAGTT GGTGATTGGT TGGTTGAAAG GCAATCTCTT TATAGTTCAT CTTTTCAATA	8280
50	TCTTCAATAA AATGATTCAT TTCTTCAATG TAGTCATTTA AAAGTAATAT CAATTCACGG	8340
	TCGTGATAAT CATGTTGTGC CGATTGCTGG TTTTCAGTGA TTGCTGGACG ATCACCTCGA	8400
	TATTGTTTAG GTGTTTGATG CGTAAAGTGT TTAAATGTTT TCGCAAAGCT CGCTGCACCT	8460

55

	TTCCGATGCT CAATTCGCGT CGTATTTAAG AAATGATGGA ATCCTACACC TAGCGATTCT	8580
	GTAAACTTTT TAGACAGATG GCTCTCTGAC CACCCACGT ATTCGCTTAA TTCTGAAAGG	8640
5	CTTAAATCTT CATGAAAATG TAACTCGATA TAGTCGCATA CTTGATTCAC TTTATCATCA	8700
	TTTAAGATAC TTTGGTTCGA ATGATATGTA CCGGGACAT AATGAATCAT ATGCATAAGC	8760
	AACTGAATCA CAAGTTGTTG CTCAGTCAAT TTAGACAAC CATTATGTCG GATATGTGTT	8820
10	GAAACCAGTC TTGCCATTAT ATTTCTCAGT TGATGTATAT TCTTTGTTGT GGTGCGATCT	8880
	GTTAAGTGAA AATATAGACA ATGCACATCA TCAAACCTGT CTGCTAAATA TTTCATTTGG	8940
15	AATGGATAT AACATATGAT GCCATCTTGT TGAAGTTGAA ATCGATACAA GTCGCGGTGG	9000
	TTAATGATGA AAATGTCGCC ACTGTTGCAT TCGGTCATAT TATTTTCATC ATAAATGTGT	9060
	GCCTCnCTT TAATAACAAA ACCAATCATT AAACCTATTGA GCCTTTTGAA ATCTGACATA	9120
20	CTCTCAGTTT CTACTCGAAT TAAATAATCA CGTTGCATAC TATCCCTCAA TTCAGTAATA	9180
	TGAATACGTT TATTTTACAT TATTTTACAG CAACATATTT GAATTTTATA TTGAATCGTG	9240
	TGTGTGGATG ATTATTTATC CTCACTCGGT TCAAGATGTA GACTATCAGT AAAAAAGTA	9300
25	TTTTACCTT TTTTCTCCAC AAAAGTAAAT TCAATGTCTT TATATCCAAC TGrTGaACCT	9360
	TTTAAGTCTC CCgAACCTTT CaACaATAAC TTTGGTGCTT TATTCGTTGG TATTTTATAT	9420
	CTTTTTCGTA ATTGTTTTAC ATTATAGTCA TCATTAGTTA ATTGATATTT TGCTGAATAA	9480
30	CTCGGTACCT CTGGATTATA TGATATATCG CCGTCTTTGT ACTTCGACAA ATCTTTAAAG	9540
	CTGCCATATT GCGCGAAGAA CTTAAATTC TCGATTCTT TTTTTATATT TTCGTCTTTG	9600
35	ATACCTTTAG TTGGAATGAT TTTATTGTCT ACCATTTTAA CGGGATATTC TTTATCTTTA	9660
	CTCTTAGGTC TACCATCTTC ATCATGAAGT GTTTCACCTA CTATATACTT CCCGGTTGTA	9720
	GTCTTAGTGT TTCTATTQAT ATATAGAACC ATACCTTTTG ATTTCATACG TTCCCCTTTA	9780
40	GGTTGAACAA CCATTTTCTT ATCGTATAAA TCCTCTAGAT TTTAATCGG ATACATACTC	9840
	TCATCACGAT AACCTTCTTT ATCGTATAAA TCCTCTAGAT TTTAATCGG ATACATACTC	9900
	AATGTTTTTT CAAAGCTTTT CTTAACTTCC GCTTCTTTAC CTATGCCACA ACCAGCAGTG	9960
45	AAACTAATGA CTAATATCAA AAAACTAATA TACAATACCA ATTTGTTTAA TCGTTTCATA	10020
	ATTTACAAAT CCTATTCTTC TTATTATCTT TCCTGGATTG ATTTCATATT TTGATCGAGT	10080
	CATGATTATT TATCCTCACT TGGTTTAAAA ATTAACCCAT CACTAAAGTA AATGTTCTCT	10140
50	TCTTTTTTCT CTACAAACGT AAATTCATG TCTTTATATC CAACTGATGA ACCTTTTAAA	10200
	TTCCCTGTAC CyTTCAACAA CaCTTCGGy GCTTTATTG TTGGTATGTC ATATCTTTTA	10260
55		

ACCTCTGGAT TATATGATAT ATCTCCATCT TTATAATTCA TTAAATCTTT AAAATTGCTA 10380
 TATTGCGCAA AAAACTTAAA GTTTTCGATT TCTTTTTTTA TGTtTTCTTC TTTAACTTCC 10440
 5 TCAGTAGAAA TGAATTTATT ATTAATCATT TTAAGTGGAT ATTTTTTTTG ATTATCCTGA 10500
 GCTACTTCGT ATTTCTCCGT CTTTAtTTCA TTAGTATAGT AAAAtCCTTT TGCACCTCTT 10560
 GTATTTCTAT CTATCTTCAA AAGCATGCCT TTTATTTTTA GAGCTTCTCC TTTATTTTGA 10620
 10 ATTGCCATTT GAGAATTTAC AATCCATGTT CCCTTATCAT TTTTATCAAA TTGATCATCA 10680
 CGATATCCTT CTTTATCGTA TAAATCCTCT AGATTTTTAA TCGGATACAT ACTCAATGTT 10740
 15 TTTTCAAAAC TTTTCTTT 10758

(2) INFORMATION FOR SEQ ID NO: 222:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1109 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

nTaTcAACTT TGGaATTTAA AgTCAATAAC TTTTTTAAAA ACTTTTTGTG TTCACAACCC 60
 30 GCTTCTTTTT CAACGCGTTT ATTGCTTAAC ACAAGAACTT ATTTTACCAG CATTCCAAAA 120
 CAAATCAACA TAAAAACGTA CAAAATAAAA GTAATTTTGT ACGTTTAGCA TATATTATAC 180
 CTATTTATTT GTAGCAGCTA TAACTTTTTG TGCAATCGAG CTATAAATTT TACCTAGACG 240
 35 ATCATCTGAT TGATATATTG ACGGTGCAAA ATCTTTTGGG TTCCAAGATG GTTGCTCTAA 300
 AGGTAATTCC CCAAGTAATT GAGTATTAAG TTCATCAGCT AACTTAGTAC CGCCACCTTT 360
 GCCAAGACA TATTCTTTAT TACCCGTCTC TTTACTTTCA AAATAACTCA TGTTTTCAAT 420
 40 TACGCCAAGA ATAGAATGAT CCGTATGTTT TGCCATCGCA CCTGCGCGAg CTGCAACAAA 480
 TGCTGCTGTA GGATGAGGTG TCGTTACAAT AATTTCTTCA CTGAAGGTA ACATCGTATG 540
 AACATCTAAA GCTACATCTC CTGTTCCAGG TGGAAGATCG AGTATTAAAT ATTCAATGTC 600
 45 TCCCCATTTA ACTTCTGTAA AGAAATTCGT CAACATTTTA CCTAACATTG GCCCTCTCCA 660
 TATAACTGGC GCATTTTCTT CCACAAAAAA GGCCATTGAT ATAACTTTAA CGCCATGACG 720
 50 TTCAACTGGA ATTACTTCCT TCCCTTTAAT TCCAGGCTTT TCATCAATAC CCATCATATC 780
 TGGTACACTA AATCCATATA TATCGGCATC TACTAATCCG ACTTTTTTCC CTTACGAGC 840
 TAAGGCAACG GCTAAATTTA CTGCAACAGT AGATTTACCG ACACCACCTT TACCGGAGGC 900

ATTTTCTTCT TTTGGTTTAA ATTGATTAC TTTTCTTCC GGCAATGTTT CAAATCGTAT 1020
 ACCGACCGTT TTCGCACCGT TTTCTTTTAA TGCAATTAACA ACAGCCATCT GTAAATCTAA 1080
 5 aTTGCGtGCA CCACCTAATT GTGCCATTG 1109

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3997 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

TCTTTATTTA AAAAAATGAT TGTCTAGTTT GTATCTCTCT GAAGATTGG CAATAAATAA 60
 20 AAGCCGATAA CCGTATAATG ATTATCGACT TAAAGTTTAT GTGGCATTIT TTACTTTTGT 120
 AATTTCAAGT GAGTTAGATG ATTATTATCA GATAGATTAT TGCTTATAAT CATATGATGT 180
 TTGAATGATA TCTTTGATTT CACTGATTAG TGCTTCTTTA GGATTAGCAG TTGTACATTG 240
 25 ATCTTCAAAT GCGAGCTCTG CCATTCTATC AATTGACTCA TTTAATTCTT CTTCAGACAC 300
 ACCTTGATGAT TTCAAATTCA TTTCAATTCC GACTGATTGA CCTAATTCGT AGACAGCTTT 360
 AgCTAATGAT TCTACGAGTG CTTCTGTCGT ATTACCTTTT AATCCTAAGA ATTTGGCAAT 420
 30 ATCTGCATAA TCTGTATCTG CTCTGAAGAA CTCATATTTA GGAATAATG CATGTTTTTG 480
 CGGGTCTTTG GCATTATAAC GGATAATATG CGGTAGTAAT ATCGCATTCT CTCTACCATG 540
 CGGAATACCA TATTCGCCAC CAATTTTATG CGCAATTGAG TGTGCAATGC CTAAGAATGC 600
 35 ATTTGCAAAT GCCATACCAG CCAAAGTTGA TGCATTATGC ATTTTCTCTC TTGAAACTTT 660
 ATCAcCCTTT TCAACAGATG ATTTTAAATA TTCGAACGTC AATTTAATCG CTTGTAGACT 720
 40 CAAACCTCTT GTGTAGTCTG AAGCCATTAC AGATACATAT GATTCCATTG CATGCGTTAG 780
 TACATCCATT CCTGTATCTG CTGTAACGCT TTTTGGCACA CTCATCACA ATTGAGGGTC 840
 AATAATTGCA ACGTCAGGTG TTAAAGCAAA ATCAGCCAAC GGATATTTTA CATTTGTTTC 900
 45 ACTATCTGTG ATAAGTCAA ATGGTGTTAC TTCTGAACCT GTACCTGATG TCGTAGGGAT 960
 ACAAATGAAC GTCGCATTTT CAGGCATGCC TATTTTATAA GTACGTTTAC CGATGTCTAG 1020
 GAACTTTTGT TTAGCACCGA AGAATGATGT CTCAGGGTGT TCAAAGAACA TCCACATTGC 1080
 50 TTTTGCAGCA TCCATCGCTG AACCACCACC AAGTGCAATG ATTGTATCCG GTTGGAAATC 1140
 AACCATCATT TCCAGACCTT TATATACTGT ATTAGTTGAT GGGTTCCGTT CGACTTCGCT 1200

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATAACCGAAT TCTACCATAC CAGGGTCACA GACAATCATC ACTTTTTCAA TCTTGTCCAT 1320
 TGTGTGTTAGA CTCATGATTG CATTTTCTTC AAAATAAATT TGAGCAGGCA CCTTGAAAAAT 1380
 TTGAGTATTA TTACGTCGTT TAGCAATCGT TTTAATGTTT AATAAATCTG TCGCACTAAC 1440
 ATTATGTGAA ATTGAGTTTC TACCGTAGaA CCACAACCTA ATGTTAAAGA CGGAATCAAT 1500
 TCGTTATACA TATCACCAAT ACCTCCAACC GCTGATGGTG TATTTACAAG TACACGACAA 1560
 GCTTTCATTG TTAGTCCAAA ATCTTTTTGT AATGTTTCAT CTTCTGTATG GATAACGGCT 1620
 GTGTGTCCTA ATCCACCAAA ATGTAGTGTG TCTTCACAAA TTTGAAATGC TTGTTTTGTA 1680
 GATTGGGCTT TTAATAAGGC TAATACTGGA GATAATTTTT CACGAGATAA CGGATAGTCT 1740
 GAACCTACAC CGCTAATTTT GGCTATGATA AGTTTTGTAT TTTCGGGGAC AGGTATACCT 1800
 GCTAATTCAG CTATTTCAAC TGCAGATTTA CCGACAATAT CAGGCTTAAT ACCTGTTTTT 1860
 TGTTCAATCa TAATTGCATT TTCTAAGCGT TGTAATTCAT CTTTTTTAAC AAAGTATGCT 1920
 TGATGTGCTT TAAATTCATT AGTAACATCT TTATAAATTT CTTTATCAAT GACTACAAC 1980
 TGTTCAGAAG CACAAATCAT ACCATTATCA AATGTTTTTG AACCAATGAT ATCATTTACT 2040
 GCACGTTTAA TGTGTGCTGT TTTTCAATG TAAGACGGCA CGTTACCTGG TCCCACACCT 2100
 AATGCCGGTT TGCCAGTTGA ATATGCAGAC TTAACCATGC CCGAACCACC TGTTGCTAGA 2160
 ACTAATGCAA TACCTTTGTG ATTCATTAAT TGTTTTGTG CTTGATAGA AGGCACTTCA 2220
 ATCCACTGAA TAATATCTTT AGGTGCACCT GCCTTCATTG CCGCTTCTAA TACAACCTCT 2280
 GCTGCACGCT TCGACGATTC TTGTGCACTT GGATGGAATG CAAAATGAT TGGATTTCT 2340
 GTCTTAATTG CAATCATCGC TTTAAAAATA GTTGTGACG TAGGATTTGT TGTTGGCGTA 2400
 ACACCACAAA TAACACCAAT TGGTTCCGCT ACATACGTTA ATCCTTTTTT TTTATCTTCA 2460
 CCAATAATCC CTACTGTCTT ATTGTCTTTT ATTGAATTC ATATATATTC AGAAGCGTAT 2520
 AAATTTTAA TCGCTTTATC TTCGTATATA CCTCTTCCAG TTTCTTCATG TGCTAATTTT 2580
 GCTAGCACCA TATGTTGATC AACAGCTGCT AAGCTCaTTT GATGAACAAT ATGATCAATT 2640
 TCTTCTTGTG ACTTTTTAGA TAATGCTTCT AATGCTTTTT TCCCTTTGTC AGCTAGAGCA 2700
 TCAATCATAA TTGCCACTTC TTGTTCTTTC GATCCACGAT TTTCTTTTTT AGGTATAGTT 2760
 AACATATACA ACCACTCCTT TATACTTTGT GAATTATTTT ACAACATTA TAGTACATGT 2820
 CTCTCAGGAT ATAAAGAAAA TTCTATACAA AAAAGTTTAA TTTCGAATAT TATTTGAACA 2880
 AATATCAAAT TTTAAAATAA ATGTTTTTCAT GAAATCATTG TTATTTCCGT GTTTTTAGAA 2940
 TGATTTTATA ATCATAATTT TTTCAATGAC ATAATTTATT CATAATTATA TATTTAATTC 3000

TCCTTGTCGA TACCTATCAA CAGATGTTAC AAATAAAAAC CaCCCGTGTG AACGGGTGGT 3120
 TTGTTCTGCG gCTATAAGCC TTCCTTACTG GCCaGCCCTA AAAGGGCACT GACAAGTCAG 3180
 5 CCAACTGCAC TACTATTCCA GCAATCCTAA AGGTTTACTC TTTTTCCTT CTTTTCCTTAT 3240
 TTTTCTCTCC AGTGAAAGGA TCTAAATATT CTTCATTGA AATTTTCATCT GCAACGATAT 3300
 CCTCTTGTA TTGATTACGA ATATAATTTT CAATCACTTT TTTATTTCTA CCTACTGTAT 3360
 10 CCACATAAAA TCCTTTACAC CAAAACCTTC TATTTCCATA TCTATACTTT AAGTTAGCAT 3420
 GTCTATCAAA TATCATTAAA TTACTTTTCC TTTTAAATAG CCAACAAATG ATGATACCCC 3480
 AAGTTTGGGT GGTATACTTA CTAACATATG GATATGATCT TTACATACCT CTGCTTCAAT 3540
 TATCTCTACA CCTTTTCTTT CATATAATTG ACGTAATATA ATCCCTATAT CTTTTTTTAT 3600
 TTTTCCATAT ATCGCTTGTC TTCGTATTT AGGTGCAAAG ACAATATGTT ACTTACAATT 3660
 20 CCATTTGCGT TGTGCTAAAC TGTTTGTGTC AGATGACATT AAATAGCATC TCCTCGTGT 3720
 GATTATTTTG GTTGGCTGAC CAATATTTAC TCTAACATGT AGAGATGCAT TTTTTTGACA 3780
 ATGGTAGAAC CTTTTCTGGG GAGTGGGACA GAAATGATAT TTTCGCAAAA TTTATTTGCG 3840
 25 CGTCCCAACC CAACTTGCAT TGTCTGTAGA AATTGGGAAT CCAATTTCTC TTTGTTGGGG 3900
 CCCC GCCCA ACTCGCATTG CCTGTAGAAT TTCTTTTCGA AATTCTCTGT GTTGGGGCCC 3960
 CTGACTaGAA TTGAAAAAAG CTTaTTaCAA GCGCATT 3997

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1391 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

GnGCGAGACA AACACACtTA TTGGTGCCAT TATmCcTAGA ATGaATTCaT ATGCAGTAGA 60
 TGAaaCAATC AAAGGATTGG CAAAACAATG CCAAAAATAT GAATCaCAAT TAATTTTAAA 120
 45 TTACACAGGT TTAAATATCG AAGCAGAAAT ACAAGCGCTT GAAACATTAG CACGCaGTAA 180
 AGTAGATGGT ATTGTTTTAA TGGCTACAGA CATAACAGAG AGACATATTG AAGTCATTAA 240
 TAAAATGAAT GTACCAATCG TTATTGTTGG TCAACAACAT GAACAACCTC ATAGTATTGT 300
 50 GCATGATGAT TATAAAGCAG GTCAAATTAT AGGCGAATGG ATTGGTCAAC AGGGATATCA 360
 ACAAGTTGAA GTGTTTAGTG TAAGTGAAAA AGATATTGCA GTTGGTATAC ATAGAAAACG 420

TACTTATGTG GAAGCACAAA AAGATGTTGC AAATGTTTTG GAAAATGTGG AGCAAGTAGA 540
 TCGCGTTGTT GGAGCAACTG ATACGATTGC ATTAGCTGCC TATAAATATT ATTCTGATAA 600
 5 AAAAGATGTT ATGAAACCAC ATCAAATATA TGGTTTTGGT GGTGACCCAA TGACACAATT 660
 AGTGTCTCCA TCGATAAAAA CAATTCATTA TAATTATTTT GAAGCTGGCC AATGCGCGAT 720
 GGaAGAGATA CAACAGATGC TTAAAAAGCA AGATATGCCA TATAGCGTCA CAGTAGATGT 780
 10 TAATATTTAG ACGCTGTATT TTTTAAAATA AATGTGGAAC CGATACCATA TAACTATAAA 840
 TGGATAGGTT AAAAGTTAAA GAACGTAGGT AAAATTTGCT ATAATAGAAT ATAAATTGTT 900
 AACAGCATAA ATTATAAAAG GAGGACTGGG TAAATATTAT GACCGAATGG ACTAGAGAAG 960
 15 AACGTTATCA ACGAATCGAG GACGTTGATA CTGAGTATTT TAAACATTA AAACAACAAAG 1020
 TTGATCAATC AAAATTTTGT CAACAATTTT ATATACAACC AGAAACAGGC TTATTAAATG 1080
 20 ACCCCAACGG ACTTATTTTT TATAAAGGGA AGTATTATGT TTCACATCAA TGGTTCCCAT 1140
 TAGGCGCACT ACATGGCTTA AAGTATTGGT ATAACACAC GAGTGATGAC TTAATAAACT 1200
 TTAAAGCTGA AGGGCCAATT TTAAATCCAG ATACTAAATA TGACAGCCAT GGTGTATATA 1260
 25 GCGGTAGCGC TTTTGAATAT AACGGGCATT TATATTATAT GTACACAGGA AATCATCGAG 1320
 ATAATCATTG GCAACGACAT GCGAGTACAG ATGATCGCAC GATTGAAAGA AGACGGTTnC 1380
 AGTTGGnAAA A 1391
 30

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 930 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

ATTTATTTTA ATGTTTATAT TTTCTAACAC TTTTATATGA TCATAGTAGT AATTGACATT 60
 TTTCAATTCA AAGACTGGTG TCATCGTATC TCACCTCGCA TTCAACTATA CAACTCCTAG 120
 45 TAACATATGT AAACAGTAAT GTTTACGACT CAAAATTAGA CAAAATAAAG AGATATGCCC 180
 CCTTCAAGTT TTATTTATCG CATTTCTTGA AGAGAGCATT ATCATTTTAT TGTTCATAA 240
 CCTTATTTTT TAATTCTGGG TCAAATTGCT GTTGTTTTAA CATTTCAATT TCAAGTTTAT 300
 50 ATGGCGGTTT TTTATTTTTC TTATCTTCAC CAACATAAGG TGTTTCTAAG ATTTTCGGAA 360
 TATCTTTAAA ACTATCATGA TGCACAATGT AATTTAATGC ATCAAAACCA ATGTAACCGA 420
 55

GAACAACTTT GATTCTGTCG ACTCCAATGA TTTTATCAAA TTCATTTAAT ACGCCATCAA 540
 AGTCCTCTTT AACATTATAT CCAGCATCAT GCGTATGACA TGTATCAAAA CATACTGATA 600
 5 AACGTTTCGTT ATTATGAACT CCATCAATAA TACGTGCTAA CTCTTCAAAT GAGCGACCAA 660
 TCTCTGTACC TTTACCTGCC ATCGTTTCAA GCGCAATACG TACATTATTG TCATTTCGTTA 720
 AAAC TTCATT TAATCCTTCA ATAATCTTAT TAATTCCGGC ATCAACACCA GCTCCAACAT 780
 10 GCGCACCTGG ATGTAATACA ATATCTTTAG CCCCTATAGC TTGCGTTCTk TCaATTTCTT 840
 GTTGCAAGAA ATCTACACCA AGATTAAACG TTTCTGGTTT GGTTGTAttG CAATaTTaAT 900
 15 GATGTATGGT GGCATGAACA ACAATATTAG 930

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1984 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

TGACGCACCA ATTTATAACG CAATTGACAA AACAATTAGA TATACCTGTG AAATTTGTAC 60
 CTGGAAACCA TGATTTATGG GAAGTTGAAA GTATGACTAC GCAAGACATT TGGATAAATT 120
 30 ATAAGAGTAT GTCACAGTGC TTGGTAGGAA AACCATTAT AGTAAATGAA GAATGGGCAA 180
 TCATAGGACA TACTGGCTGG TATGATTATA GCTTTCAGC ACAACGATTT TCATTAGATG 240
 35 AGTTACAAAA AGGAAACAT TATGCTGCGA CTTGGCAAGA TAAAGAACGA ATATCTTGGG 300
 GCATATCAGA TCAAAATTTA TCTAAAATAG CGGCTGAACA AGTGAAGAAA GATATATTAG 360
 AAGTAGGAAA TAGACGAGTG ATTTTAGTCA CACATGTTGT GACGCACCCT GATTTCAATTG 420
 40 TTCCTATGCC GCATCGTATA TTCGATTTTT ATAATGCATT TATTGGGACA AGTGATTTCa 480
 ATCCTTTGTA TGCGATGTTT GATATACCAT ATAGTATTAT GGGCCATGTT CATTTTCGTA 540
 AAAGTGTGAT AGATGATGGC AGATGTTATC TCTGTCCGTG TCTAGGCTAT CCAAGACAAT 600
 45 GCGGTTTCAAG AGATATTTAC CAGGAAATAA ATGAGACGAT ACAAATAATA GAAATTTAAA 660
 ATGCGCAAAC CTGACCCAGT TTGCGCATTT TATGTTTTAC ACACGCGAGT AATGTGTTTA 720
 CTTACGTGTG TTTATTTTGT TGCTGATTTT CAATTGTATA TGAATGTGGT TGCACATAAA 780
 50 TGCACTTTCT TCCTGGTGAA TTAAAGCTGT ATTCCATTTT CTCTTTACGA ATTTTAATAA 840
 TTTGTTTGCG ATTTGGAATG ATGGCAGGTA AACTAGGCC ACGACGAATA TGACTCCAAA 900

TTGAAACTTG TTTCGCTGGC TTGTTATCAA AGCGGAAAAC ACGTAGTAAT GGTTTAGAAC 1020
 CAAGATTAGT ATGGTATATT AACACAGGTT GACCTTGATC GATAATACCT TTAAGATCTT 1080
 5 CTAACGATTT ACCAGTGCCG TCTACGATAT TAGGATTGTA TTTTGTAAA AATGGTACAT 1140
 ATGCTTCTGG AAATATCGTT TGATGATAAT TGCCAAGCTT AATGAATAAG TGATGTCCAA 1200
 CATAACCTTT ATGTGGATTG TTCGGATGTG TCGGCCAATG TCTCATAATT TCTGTAGCAG 1260
 10 GGATATGTTG GTTGTGTAT TGCAACATCA TGGCTGCGGA AACACCTTCA CACCCCATGA 1320
 CCATAGGGAT AGGAAATAGC TGACTGATAG GTTTAACTGG TAATATTTTT CGGTTCAATA 1380
 TATAGTCCTC GCATTGATTG AATAAATATT TAATATAATT ATATAGCGTC AATGCAAAAT 1440
 15 GTCCTAAACA TATGTTTTAC ATGAGTGAAT AAAATTAATG GAGTGATAAA ATGGAATATC 1500
 AATTACAACA ACTTGCGTCG TTAACGTTAG TAGGTATTAA AGAAACGTAT GAAATGGAC 1560
 20 GACAGGCTCA GCAACATATA GCAGGGTTTT GGCAAGATG TTATCAAGAG GGAGTAATTG 1620
 CGGATTTACA GTTAAAAAAT AATGGTGATT TAGCCGGGAT ACTTGGCTTA TGTATACCTG 1680
 AATTAGACGG TAAGATGTCA TATATGATTG CAGTTACCGG AGATAATAGT GCTGATATTG 1740
 25 AAAAATATGA TGTACATAACA TTAGCAAGTT CAAAGTATAT GGTATTTGAA GCACAGGGCG 1800
 CAGTACCTAA AGCAGTTCAA CAAAAAATGG AAGAGGTTCA TCACTACATA CATCAATATC 1860
 AAGCAGATAC GGTAAAATCA GCACCATTTT TTGAGTTGTA TCAGGATGGT GATACTACAA 1920
 30 GTGGAAAATT AATATTACCA GAAATTTGGG ATnCTGTTA AAGGGGTGAT TGAAATAnGA 1980
 AnTG 1984

35 (2) INFORMATION FOR SEQ ID NO: 227:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6373 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

45 GATTCCACGT GTGTTAAAAG AAGTTACAcC TTCAATGATG GTATTTACTA ATTTCTTTAG 60
 AGATCAAATG GATCGCTTCG GTGAAATTGA TATTATGGTT AATAACATTG CAGAGACAAT 120
 TAGTAATAAA GGCATCAAAT TATTGCTAAA TGCTGATGAT CCATTTGTGA GTCGTTTGAA 180
 50 AATCGCAAGT GATACGATTG TGTACTATGG TATGAAAGCA CATGCCCATG AATTTGAACA 240
 AAGTACGATG AATGAAAGTA GATATTGTCC AAAGTGTGGT CGCTTATTGC AATACGATTA 300

55

	AAAAATATGAA ATATCAAGTT TTGATGTGGC ACCGTTTTTA TATTTAAATA TCAATGATGA	420
	AAAAATATGAT ATGAAAATTG CAGGTGACTT TAACGCTTAT AACGCGTTAC AGCATATACT	480
5	GTTTTAAGAG AGCTAGGGTT AAATGAACAA ACAATTAAAA ATGGCTTTGA AACGTATACA	540
	TCAGACAATG GTCGTATGCA GTACTTTAAA AAAGAACGAA AAGAAGCGAT GATCAATTTA	600
10	GCTAAAAATC CTGCAGGAAT GAATGCAAGT TTATCAGTTG GTGAACAATT AGAAGGCGAA	660
	AAAGTGTATG TTATTTGCT AAATGATAAC GCTGCAGATG GTCGAGATAC TTCATGGATT	720
	TATGATGCAG ATTTTGAAAA ATTATCTAAG CAACAAATTG AAGCTATCAT CGTGACAGGT	780
15	ACACGAGCAG AAGAACTTCA ATTGCGATTG AAGTTAGCAG AGGTTGAAGT ACCAATTATA	840
	GTTGAGCGTG ATATTTATAA AGCAACGGCA AAGACTATGG ATTATAAAGG TTCACAGTT	900
	GCAATACCAA ACTATACATC ATTAGCGCCT ATGCTTGAAC AATTAAACCG TTCGTTTGAA	960
20	GGAGGTCAAT CATAATATGC ATGAATTGAC TATTTATCAT TTTATGTCAG ATAAATTGAA	1020
	TTTATACAST GATATAGGAA ATATTATTGC TTTAAGACAA CGTGCTAAAA AACGAAATAT	1080
	TAAAGTTAAT GTCGTAGAAA TCAATGAAAC AGAAGGTATT ACCTTTGATG AATGTGATAT	1140
25	TTTCTTTATC GGTGGTGGAA GTGATAGAGA ACAAGCATTG GCAACAAAAG AATTAAGTAA	1200
	AATTAAGACA CCACTTAAAG AAGCGATTGA AGATGGTATG CCGGGATTAA CGATTTGTGG	1260
30	AGGCTATCAA TTTTLAGGGA AAAAATATAT CACGCCTGAT GGTACAGAAT TAGAAGGGTT	1320
	AGGTATTTTA GATTTTATA CTGAATCAA GACAAACCGA TTAACAGGAG ATATTGTTAT	1380
	CGAAAGTGAT ACTTTTGAA CTATTGTAGG TTTTGAAAAT CACGGTGGTA GAACATATCA	1440
35	TGATTTCCGT ACACTTGGTC ATGTTACTTT TGGTTATGGT AATAATGATG AAGATAAAAA	1500
	AGAAGGCATT CATTATAAAA ATTTATTAGG TACTTATTTA CATGGACCAA TTTTACCTAA	1560
	AAATACGAA ATCACTGATT ATCTGTTAGA AAAAGCTTGT GAACGTAAGG GTATTCCGTT	1620
40	TGAGCCTAAA GAAATAGATA ATGAAGCGGA AATACAAGCG AAACAAGTAT TAATAGACAG	1680
	AGCAAATAGA CAGAAGAAAT CTCGTTAACT CTGAACATCG CATCAATGGA TTTAATATTG	1740
45	ATAACGATG AAGTTTAGTA ATTAATCATA TATGTATAAA CACACACATT ATTTTGGATG	1800
	GAAACAACCA AATTGATGTG TGTTTTTTTG TTCTAGTGAA TAATTATTAT ACAATGAGTA	1860
	TCTATCCTAG AATTATCAAT AGTAATGGTG ATTATGCAAC ATGAAAAAAT GAATGATGAA	1920
50	AGGAATTTGA CGATGAAGCC TACTAAAGTG ATATTAAAAG ATGCATCTTA TTTACATAGC	1980
	AAAACATCGA TAACATTTAT TTTAAAAGAT GTAGTTATCG AAGAAGATAA TAAAATTTAT	2040
55	TATTTGACA CTAGTGCACT TTCGAAGATC AAGAAGTTAA ATTTGAATTT GCACTCTTTG	2100

TTATAGAACC TGATTTACAT TTTACAATTA TTGATTTTAA TCAAGAACTG CTTTGTATTT 2220
 ATATTGATTT TGATTCTGGT TTAAGGCATT CAAACATGGC AACAGAATCT GGTATTTTCAT 2280
 5 TAAGGATAAA TGTGCTAAA TCAGATTTTA CTAAATTTAT TAATGAATTA GCCTCTTTAC 2340
 ATTAATGATT TAAATCTGAT ATGTAATTAC AATCAAAAAA GACAGCCACA TCCCTCCGTA 2400
 GTTTAGGCGT GTGGCTATAT TTGAGTCTGA ATATTTATGC TTGTAATTTT AAAAAGGGAC 2460
 10 ATGCTATATA CGATAAAAAG AGGCGGGGAC ATAAATCAAT GTTCTATGCT CTACGAAGTT 2520
 ATATTGGCAG TAGTTGACTG AACGAAAATG CGCTTGTAAC AAGCTTTTTT CAATTCTAGT 2580
 CAGGGGCCCC AACAAAGAGA AATTGGATTG CCAATTTCTA CAGACAATGC AAGTTGGGGT 2640
 15 GTGGGCCCCA ACACAGAGAA TTTGCGAAAAG AAATTCTACA GGCAAAGCGA GTTGGGGTGG 2700
 GACGACGAAA TAAATTTTAT GAAAATATCA TTTCTGTCCC ACTCCCATGG TGCCAATTAG 2760
 20 CATAAGGTAC TAAATTAAG CATATCTGCT GTCTAGCAGT CGATAAATCA TTAGAACTTC 2820
 GTATAGTATA TGACTTTTAA TTTGATTTTC ACCACTAATT TCAAGTGCTT TTATAGTCGA 2880
 ACGTAAAGTT TCTACAGAAT CATCTTCTCT CTAAAAGAA CCATCATAAA ATATATCTTT 2940
 25 GATGCTACTA CTAATTTTTA GCAATGCCAT TTTTTCGTCA CCTGAAAAGT TAACACGAGT 3000
 ATTTTTAGGC AAGTAAATGA TATTTGATAA ATGAGTGATA AACAAACGAT TCGTATATGC 3060
 ACGTTTAGTT AATTGATTGA GTAATTTCCA ATCACATTCT TTTTCTTAT GATAGCTTAA 3120
 30 TTCATCACGT TGATAACTTA TTAACGTTTC AACTTGATTA TTTAAATTGA AAATATTTTT 3180
 ATATGCTTTT TCGCTTTTAT CAGATTGCAG TCTTGATAAG ATAAGTTCTT GGCAGCGATT 3240
 GTAAAATAAT TTATACATCA AGGCATCTGT CTTACTTAAT TTTTCTTCGA CCTGACCATA 3300
 35 ATACTTAGGT GGAAACACCA TGAAGTTAAT TAAACCTGAT GTCACGAGTC CAATAATTGC 3360
 TGTCAATGTT CGAGACAAAA AGTTGAATAT GTAGGCATCA TGAATACCTG GAATCATAGC 3420
 40 TAATGATGTT AGTACAGCGA CATTGCTACC AACTTGCAAT TTGAGTTTTG TACAGAATAA 3480
 AATCGTGAAC GTTGCACTCA ATGCATATGT AAAAGGTGAT TGATCGCCGA ATAAATATGT 3540
 AAATAATACT GCAAAGCCTG CACCAATTAC CGTAGCAGGT AATCTACGAT AACCTTTAAT 3600
 45 AAGTGATGCC TTGGCAGTTG GTTCAATTGT GACTACAGCT GTTAAATGG CATAGATGGG 3660
 TGTTAAATCT AGTGCCATAC AAAAGACAGC TGTTAAAAAA ATGGCAATAC CAGTTTTAAT 3720
 TGTTCTGGCA CCAATTAAAT GTTTATACCA TTGATCGTTC ATTTTTTAAC CTCTAATCAT 3780
 50 CGTAAATCT TAGCGAGCGC TTTATAATAA TAGTATCGTA CATTGGAAAA GTTCATGTAT 3840
 GTAAAATATT TGAAATAATC ATACATAAGC ATTACTTTGA TTTTCATATA CATTAATCAA 3900

55

	CAAGCATT TTT TCAATTATAG TCCGGGGCCCC CAACATAGAG AATTTCAAAA AAGAAATTCT	4020
	ACAGGCAATG CAGGTTGGCG GGGCCCCAAC ACAGAAGCTG ACGAAAAGTC AGCTTACgAT	4080
5	AATGTGCAGG TTGGCGGGGC CCCAACATAG AGAAATTGGA TCTACAATTT CTACAGGCAA	4140
	TGCAAGTTGG GGTACAACGA TAAAGAAATA TTTTCTTT ATCACACTAT GTCTCACTCA	4200
	CTTTCCAAA TACTAAAGTA ACATCTTTAG TATATCAAAG AATTTTGTGCT ATAATAAGTT	4260
10	ATAATTATAT AAAAAAGGAA CGGGATAAAA TGATTGTAAA AACAGAAGAA GAATTACAAG	4320
	CGTTAAAAGA AATTGGATAC ATATGCGCTA AAGTGCACAA TACAATGCAA GCTGCAACCA	4380
15	AACCAGGTAT CACTACGAAA GAGCTTGATA ATATTGCGAA AGAGTTATTT GAAGAATACG	4440
	GTGCTATTTT TCGGCCAATT CATGATGAAA ATTTTCCTGG TCAAACGTGT ATTAGTGTCA	4500
	ATGAAGAGGT GGCACATGGG ATTCCAAGTA AGCGTGTCT TCGTGAAGGA GATTTAGTAA	4560
20	ATATTGATGT ATCGGCTTTG AAGAATGGCT ATTATGCAGA TACAGGCATT TCATTTGTCTG	4620
	TTGGAGAATC AGATGATCCA ATGAAACAAA AAGTATGTGA CGTAGCAACG ATGGCATTG	4680
	AGAATGCAAT TGCAAAAGTA AAACCGGTA CTAAGTTAAG TAACATTGGT AAAGCGGTGC	4740
25	ATAATACAGC TAGACAAAAT GATTTGAAAG TCATTAAAAA CTTAACAGGT CATGGTGTG	4800
	GTTTATCATT ACATGAAGCA CCAGCACATG TACTTAATTA CTTTGATCCA AAAGACAAA	4860
	CATTATTAAC TGAAGGTATG GTATTAGCTA TTGAACCGTT TATCTCATCA AATGCATCAT	4920
30	TTGTTACAGA AGGTAAAAAT GAATGGGCTT TTGAAACGAG CGATAAAAGT TTTGTGCTC	4980
	AAATTGAGCA TACGGTTATC GTGACTAAGG ATGGTCCGAT TTTAACGACA AAGATTGAAG	5040
35	AAGAATAGTT CAACATATAC TAAGACTAAA GTATGAACAT CATTTAGTTC CGGAGCCTAT	5100
	TCATATTGGT TTCGGAAC TG TTTTATAATA ATTAAGAACA CAATCAATGC GTCATTTCAA	5160
	AAATATGTTG TAACAAAGTA GTTTTAAAGC AAACATATCA TCGACATCAA CGAAGATACA	5220
40	TAGCGCATTT GGTATTTTAA AACTTATTAT AAAAGGTGAT AGTTATGAAC TATGTTGAAC	5280
	GTTATATTGA ACAGTTTTTG AGAGCAACAG TAAGAAATAA TATCAAGCAC TACCTTTTAA	5340
	TGCTAGATGA AAAAAAGAAA AATTAGATG ATTATATGCG TTATTTAATT ACTAAAAAG	5400
45	AACAACCTAG CAAGTTAATT GACAGTCTAA TGCTAACATT AGAAAATAAA TATATTGATA	5460
	TTGCTGAAGC ATTTCAAATT CAATGTGCAA GAGAAATCAA TAATCAAGAA ATTGAAAATA	5520
	TTAAATCAGA GTTGAATAAA GTTGAAGCAT ATTATGCACA AATTGAAACT CAAATTCAAC	5580
50	AAACTTCAAC TGAAAAATA GCAACAGAAA AAACATCGTA TCTAATAAAT TATATGAACG	5640
	CTGTGGCATA GAAAGGCGGC GAAACATGAC ACACAAATAT ATATCAACGC AAATGTTGAT	5700

55

CTTTTTACTC GTTCTATTAT TGGGATGTGT ATTAGTTTAT GTAGGATATC TTTATTTTCA 5820
 TAAATACGT GGCCTTTTGG CGTTTTGGAT AGGCGCGCTA TTAATTGCAT TCACATTATT 5880
 5 GTCTAATAAG TATACAATCA TCATCTTGTT CGTCTTTTTA TTATTACTTA TTGTGCGTTA 5940
 TTTAATACAC AAGTTTAAAC CAAAAAAGT AGTTGCGACG GATGAGGTTA TGACTTCACC 6000
 ATCTTTTATT AAACAAAAGT GGTTCGGTGA GCAACGTACA CCAGTTTATG TATATAAGTG 6060
 10 GGAAGATGTA CAAATTCAAC ATGGAATTGG CGACCTACAT ATTGACTTAA CAAAAGCTGC 6120
 AAATATTAAG GAAATAATA CCATTGTTGT TAGACACATT TTAGGTAAAG TGCAGGTTAT 6180
 15 ATTGCCGGTT AATTACAATA TTAATTTACA TGTAGCTGCT TTTTATGGAA GTACTTACGT 6240
 GAATGAAAAA TCATATAAAG TTGAAAATAA CAATATTCAT ATTGAAGAAA TGATGAAACC 6300
 GGATAACTAT ACAGTTAATA TCTACGTATC AACGTTTATC GGAGACGTAG AGGTGATTyA 6360
 20 TCGATGAAYC ACT 6373

(2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4488 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

ATAGnGAAAG CGTTTTACAC TTAATAACTC CCTCTTAAAT GCATCCAGGT TCTATGTAGT 60
 35 AAATCATGAA nATAACATAT AAATnTAGAG GAGATTTACC TTTGAATACA GAGAACAACA 120
 AGAATCAAAA CCAATCTGTT AAAAATTCTG AAAGaCGCGG CATGTTAAAA GGATGCGCGG 180
 GTTGcCTTAT TTCTTTTATT TTATTAATAA TCTTATTATC AGCCTGTTCA ATGATGTTTA 240
 40 GTAATAATGA CAATTCCTACT AATAATCAAT CATCAAAAAC GCAATTAACT CAAAAGATG 300
 AAAATAAAAA TGAAGATAAG CCTGAGGAAA AATCAGAAAAC AGCAACAGAT GAGGATTTAC 360
 AATCAACCGA AGAAGTACCT GCAAATGAAA ATACTGAAAA TAATCAACAT GAAATTGATG 420
 45 AAATAACAAC AAAAGATCAA TCAGACGATG ATATTAACAC ACCAAACGTT GCAGAAGATA 480
 AATCACAAGA CGACTTGAAA GATGATTTAA AAGAAAAGCA ACAATCAAGT AACCATCATC 540
 AATCCACGCA ACCTAAGACC TCACCATCAA CTGAAACAAA CACGCAACAA TCATTTGCTA 600
 50 ATTGTAAGCA ACTTAGACAA GTATATCCGA ATGGTGTAC TGCCGATCAT CCAGCATATC 660
 GACCACATTT AGATAGAGAT AAAGATAAAC GTGCATGTGA ACCTGATAAA TATTAAACAA 720

	GGGAGATTTT TTAGGCATGA GCAATCAATT CAAAAGCGAA GAAGAGCGAA GACAATGGGA	840
	ACAATTCCAA GCTTTCCAAA ATCAACAAAA CCAACAGAAC CAGCAATACG GACAAAAGAA	900
5	ATCTAAAAAA GGATGGTTCT GGGGCTGTGG TGGTTGTCTA GTATTATTTA TTTTAATTAT	960
	CATCGGTATT TCAGCTTGTA CAGCTGGTAT TACAGGTAAC CTTGGCGGAA ATAGTTCTAA	1020
	AGAAACGAAC AAAACCCATA AAATCGGTG AACTGTTAAA AATGGCGACC TTGAAGTCAC	1080
10	TGTAAATTCA GTGGAACTA TGAAATCTGT AGGACCATCT CTTGCACCAA CAAACGCTAA	1140
	AGGTATATTT GTCGTTGCTG ATGTGACGAT TAAAAACAAA GGTAAAGAAG CGTTAACAAT	1200
	TGATAGTTCA ATGTTTAAAGC TAAAATCCGG TGATAAAACA TTTGAAGCAG ATAATACAGG	1260
15	TTCAATGTCT GCTAATCAAA GTGACAATGG TAGTATAGAA AATTCATTTT TCTTACAGCG	1320
	TATAATCCA GATAGCACTG CTCAAGGTAA AATTGTTTcG ATGTGTCAGA AAACATAGCC	1380
20	AACGCAAAaG ATAAAAAATT AGAAGTTATT TCTAGTTTAT TTAGCGTCAA GAAGATTACA	1440
	TTTGATTAT CCGATGCTAA AAAAACATCA AAAGCTAAAA AAGACAAGCA AGATACAGAA	1500
	GTAGCTGTTG CGAGTTCAAA TAGCGATAAT GTAAGTTATG AAGCTTCGGC TACTACACCT	1560
25	GcTACAACTT CTAGTGCGGA TACTGATTCT GAAGATAGCG AAAAGTCTAG TAAAGATGAG	1620
	GATAAGCAGA ATGCGTCTAA AAgTGATAAA TCTAGTGTAG AAAAAAGTGA ATCTAATGAG	1680
	GAAACTGCTC CTGTAGAGCC CATGCCCCAT AGCAAACCTA CCACTAGTGA aGCACCACCT	1740
30	AGCCAAAATA TTCACaATGa AGATAGCmTG TACGACGCTT CAACAGAATA AAATtnyCAG	1800
	cAGCTCGGCT ACCCTTCTTT TACGGAAAAA TTAATTATAC ATAATCaAT CaAGGAGATA	1860
35	AAAAAATGAA ATTCAAAGCT ATCGTTGCAA TCACATTATC aTTGTCACTA TTAAGTGCCT	1920
	GTGGTGCTAA TCAACATAAA GAAAATAGTA GTAAATCAAA TGACACTAAT AAAAAGACGC	1980
	AACAACTGA CAACACTACA CAGTCAAATA CAGAAAAGCA AATGACACCA CAAGAAGCCG	2040
40	AAGATATAGT TCGAAACGAT TACAAAGCAA GAGGCGTTAA TGAATATCAA ACATTAAATT	2100
	ATAAAACAAA TCTTGAACGA AGCAATGAAC ATGAATATTA TGTGACAT CTAGTCCGCG	2160
	ATGCAGTTGG CACACCATTA AACGTTGTG CTATTGTTAA TCGACACAAT GGCACAATTA	2220
45	TTAATATTTT TGATGATATG TCAGAAAAAG ACAAAGAAGA ATTTGAAGCA TTTAAAAAGA	2280
	GAAGCCCTAA ATACAATCCA GGTATGAACA ATCATGATGA AACAGATGGT GAGTCAGAAG	2340
	ACATTCAACA TCATGACATT GATAATAACA AAGCAATTCA AAATGACATA CCAGATCAAA	2400
50	AAGTCGACGA TAAAAATGAT AAAAATGCTG TTAATAAAGA AGAAAAACAT GATAATGGGG	2460
	CAAATAATTC TGAAGAACT AAAGTTAAAT AATGGCATAC TTTGATTAAT CGTAATTTTT	2520
55		

	ACTATGCATG GTCTTTTAA TCAACTTAAA CTCGGCATT TTTCAATCGA AAACGCAGAG	2640
	CATACGCTTT TTACACCTTA TATGTTGGAA ACGCTCTCTT CCCTAGGCGT GAAAGACAGC	2700
5	ATTGTGGAAT TAATTCATAA AGGGACTGAA TTAGAAGACT TTGCGGCATT TAATTTATCA	2760
	ATTGAAGACA CAGTTACAGT CTGTTTACAA AGAACTGAAG AACTATTAAA ACAATACAAA	2820
10	AATGTGGAAT TCAATGACAA AATATTAATC AATTGGCGTA TTATACAAGA GAAATAGACA	2880
	TATAAAAGTC GAATGTAACt ACGTGAGTAT TGATTTTATT CTTTGTAAT TACAAGCATT	2940
	TCATATTATA AAGTTTGAAA AGAGGTATAT TGAAATGGAG AAAAATGAAT ATATAGCTAA	3000
15	ATATAATGAA TATAGTCAAT TATTAGACGC TACATACTCG CAAGCTGTAG CATmCCTTTT	3060
	AAGtAAATat GCGCTGTAA CCGATGATTA TTATAAGaA AAATCATACA CGCGATTTTT	3120
	AAAtGGAGnA ATCAAAAGTA TTTCAAAAGG AAAATACACT AGAGCTAGCG AaGGATTATA	3180
20	TTGCCATCAT ATAAGCGAGG ACAAATTCCA AAATCTATCT GATCTAAGAT TCATtTCCAA	3240
	ATTTAAGTAC TCATACGACG TTCAAAAGAA AGAAACTTA GTGTACTGTG ATCTAATCGA	3300
	GCATTTAATT TTACATGCAA TTATTACAAA AGAATCCCAT GGCCAATTTG GTGTAGCTGG	3360
25	ATTATGTCAA ATGATCAAAC CAACAGTCAT TGATTGGTAC ATTGGCGAAT ATAATCCAAA	3420
	ACCAGCATGG ATGCAAGCCA CCAAAGCAGC TGCCTATTTG CCTGGAATAT TAGTAGAGAA	3480
30	ATTACTCAIT AAAATTGACG ATATGTTAAA AGGAATAGAA ATATAAGATT TCCTTGAGTC	3540
	TAGATAAATG ATTAATGTAG ATTTATTTTT TGCTGTTGAG ATTTTGTTAT AGATGTTTAA	3600
	ACCTGTAATT AAATATATTT TATAAAATAG ACCACGCATA CCTATCTATA AACGGrCAAT	3660
35	GTTTATAAAT GAGTTTGCAAT GGtCTTGAAT TGTATTAAAT TTCTTTTGGT TTTAATAAAT	3720
	CGACTAGATT TTCACAATAT TTATCAAATA TGTATTCTTA AATTATACAG CCTTAATCCA	3780
	GCAgCTACTT TCGAAACTTC CAACTTAGTT GATATAAGGT TCAATAGTTT GTTTCGTTCT	3840
40	TTTTTCAGATA AACCAGAACT TAAATTGATA TTATTGACTT CATAAAAATT ATAGACTAAT	3900
	GCCTCTATTT GCTTTTTTAGG CATAAGTAAG TCGACTGAAA ACTGATTTAC GTCGCTTTCA	3960
	TAAATCATTT CATGTAAAT CTTTAGACTA TTATCGTTAC TATCTCTCAT TAAGTCTGTA	4020
45	TTTTTAAATA AATAACGGCC CAATTCACGA GCTATTGAAA ATCTTGTATT ATTAATCGAG	4080
	TGATTATTAT TAATATAGAT TGTTCTTCCA CTAAATAAC CCGAAGTATT ACCCTCCATT	4140
50	TTAATATATC TAACATTTAA ATTAAGTTGA AATAATAGCT TGTCTATGTC AATAGCAAAG	4200
	TGTTCAGAAG TAATAAAAAG TTGATCCATT TTGTCCTTTA TAAATGCCTG AAATAATCGA	4260
	ACTATTTTTG ATTCTAAAAT ATCTTCATAA TGAACTTTCT CAATAACTTT CAATTGATTC	4320

55

AACATTATTA AAATAAAACC CCTCTACTAC TATATGTAAC GAAGGGACAT GATTTCAAAA 4440
 TAAAATACCT TTTTATATAA TnTATTATAA TATCCCCCAC TATACnAC 4488

(2) INFORMATION FOR SEQ ID NO: 229:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 846 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

TATGGCGCCA TATTAGTTGT AACTGGTTTA AGAGGTCCaA GAAaTATCaA ATAAAGTTGT 60
 tCCTGGGCTT GGTACTGTTA TCTCaATATT GmWTGCaTTT GGTGGTCTAG CTTTTAATAT 120
 TGGTAATATT GCTGGTGCCG GTTTAGGTTT AAATGCAATT TTTGGATTAG ATGTAAATG 180
 GGGCGCAGCT ATTACTGCAA TCTTTGCAAT ATTAATCTTT GTAAGTAAAA GTGGCCAAAA 240
 AATTATGGAC GTTGTITCAA TGATTCTTGG TATTGTGATG ATTTTIGTTG TGGCATATGT 300
 GATGTTTGTT TCTAATCCAC CTTATGGTGA TGCTTTTGTG CATACATTTG CGCCAGAACA 360
 TCCAATGAAA TTAGTCTTGC CCATCATTAC GTTAGTTGGT GGAActGTar GTGGTTATAT 420
 TACCTTTGCA GGTGCACATC GTATATTAGA CTCTGGCATT AAAGGTAAGC AATATTTACC 480
 ATTTGTAAAT CAATCAGCAA TTGCTGGTAT TTAACTACA GGTATTATGA GAACGTTACT 540
 ATTCCTAGCG GTATTAGGAG TTGTTGTAAC AGGTGTGACA CTAAGTTCTG AAAATCCACC 600
 AGCGTCAGTT TTTGAACACG CAATTGGACC AATTGGAAAG AATATTTTTG GTATTGTGTT 660
 ATTTGCTGCA GCTATGTCAT CAGTAATTGG CTCAGCATAC ACAAGCGCAA CATTTTTAAA 720
 AACACTTCAT AAATCACTTA ACGAAAGAAG TAATTTAATT GTGATTGTGT TTATCGTTAT 780
 TTCAACAATG ATTTTCTTAT TTATTGGAAA ACCAATCAGC CTTTTAATTA TAGCAGGCGC 840
 GATAAA 846

(2) INFORMATION FOR SEQ ID NO: 230:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2072 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

	TCTTTTAAA AGGTACTAAT ATTTCTTTAG TGAAAATTGA ATCACGGTCG TTTATTGGTG	120
	CCTTGAGTAT ATTATTATAG ACGGAATCTG ATCTAATAAT ATTGATTTTA TACATGATAA	180
5	ACCTCCTTAT GTTGTCAGCA TAAAGGATAA CGTAACGTGA TTTTCAAGCA GTAATTGTAA	240
	CTAATTGAmA AAAATTAAGA AAAGTATGTG AGTGTTCTTA AwTAATATGa TTAAAATGAT	300
	GGCGAATAAG TGTCTaAAAG CATCTTAAAG GGACATTGTA TAGGGTAAAT CACTTCATAA	360
10	ATAAGGGaAA ATCCTTATGT TCACTTTTTTC ACAATCATra TAAAATATAT ATGTAGTCAA	420
	TACTTTGTCT ATATTGAATG TTTTCATATA AATGAAAGCA TTTTAAATA ACATTGACCT	480
15	CTAATATATA GGCAGAGTAT TGATATCTAT TAAAAAATAA ATGATTTTGA TGAAGGTGAA	540
	ACGTATGTAC AAAACAAAAG GTGGCTTTCA ACTTACATTA CAAACATTAA GTTTAGTGGT	600
	TGGGTTTATG GCTTGGAGTA TAATTGCGCC ATTAATGCCC TTTATTAAAC AAGATGTCAA	660
20	TGTTACTGAA GGTCAAATAT CAATCATTTT AGCGATACCA GTTATTTTGG GATCGGTGCT	720
	CCGTGTGCCA TTTGGTTATT TAACAAACAT TGTGGCGCT AAATGGGTAT TCTTTACTAG	780
	TTTTATCGTA TTGTTATTCC CGATATTTTT CTTAAGCCAA GCACAAACAC CGGGTATGTT	840
25	AATGGCTTCA GGATTTTTTC TTGGTGTAGG TGGTGCAATT TTCTCAGTTG GTGTTACATC	900
	AGTTCCTAAA TATTTCCCTA AAGAAAAAGT AGGTCTAGCA AATGGTATTT ATGGTATGGG	960
30	AAATATCGGT ACAGCAGTTT CTTCATTTTT AGCACCACCG ATAGCGGGTA TTATTGGTTG	1020
	GCAAACAACA GTTAGAAGTT ACTTAATTAT TATCGCTTTA TTTGCATTAA TTATGTTTAT	1080
	TTTTGGTGAC ACACAAGAAC GTAAAATTAA AGTACCATTA ATGGCtCAAA TGAAAmCATT	1140
35	ATCTAAAAAC TACAAATTAT ATTACTTAAG TTATTGGTAT TTTATTACTT TTGGTGCTTT	1200
	TGTAGCATT TGGTATTTCT TACCTAACTA CTTAGTAAAT CATTTTGGAA TTGACAAAGT	1260
	AGAT ⁻ CTGGT ATTCGATCAG GTGTATTCAT TCGCTGGCA ACATTCTTAA GACCAATAGG	1320
40	TGGCATTTTA GGTGATAAAT TTAATGCAGT TAAAGTATTG ATGATTGATT TTGTTGTTAT	1380
	GATTATCGGT GCCATTATTT TAGGTATTTT AGACCATATC GCATTATTCA CTGTAGGCTG	1440
	TTTAACAATA AGTATTTGTG CAGGTATTGG TAACGGCTTA ATCTTCAAAT TAGTACCATC	1500
45	ATACTTCTTA AATGAAGCGG GATCCGCAAA TGGTATCGTA TCAATGATGG GTGGTTTAGG	1560
	AGGATTCTTC CCACCACTAG TAATCACGTA CGTAGCTAAT TTAACAGGAT CAAGTCATTT	1620
	AGCATTATTT TTCTTAGCGG TATTnGGAnG TATTGCATTA TTTACCATGC GTCATTTATA	1680
50	CCAAAAAGAA TATGGCTCAT TGAAaAACGG TTGATATGTA ATACATGCCA TTcATTAGT	1740
	TAAATACAAA GCCTT ^a ATAT CATGCGCAAT ATTCGTAGCA TGACATTAAG GCTTTAGTAG	1800

55

CTTGGTTTGA TTTTAGGCAA GGTAATGGTT AATAACCCAT TTTCAAAACT AGCAGTAATA 1920
 TGTGCTTAT CAACAGCTTC AAAATCAAAT TGACGCATTA ATGATTGCGAA GTTACGCTCA 1980
 5 TCTAAATGA GTTGTTCAGA TTTGTATTTT GCGCTTCTAG TAGCTTGAAT AGTGAGCGWA 2040
 TTAttATTGA AATCgATACT AATAtCTccc TG 2072

(2) INFORMATION FOR SEQ ID NO: 231:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3159 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

20 CGTCTTCTCT TGGTTATTCT CTGTGTTTTG TCTTGTTC AATTCGATTT TTTGTTGTTT 60
 GAATGAATGT AATACTTTAT TTTTTTTTGG CACATAATCC AAGTnATTTT TAGGAATTAA 120
 TATACGGTCT TTAATGCTT CTTGTATTTT GCTCACAATC AATTGGTATA GTTGCTCTTC 180
 25 TTTTGATAAA CGCACTTCTA GTTTTGTGG ATGAACATTT ACGTCTACTA AGATTGGATC 240
 CATTTCAATA TTAATATAAC AAATCGGGAA CCTACCTATT GTTAAGAGTG TATGATAGCC 300
 TTCTAAATC GCTTTATTTA GCATAAAGTT TTTAATGTAT CGTCCATTAA TAAAAATAGA 360
 30 AATATAGTGC TTATTACTTC TAGAATGTTT AGGCTTTGCA ACAAACCTT CAATGTGATA 420
 ATCACTTGTA TCTCCAGATA TATGTACTAA ATCTCGTGCA ACTTTCATCC CATAAATCTC 480
 35 TGCCATCACT TCATTAGTTC GTCCTGAACC ATTTGTACTT AACATTGTTT TGCCATCTGA 540
 AATGAGTGCT ATTCGAATGT CCGGATGGCT CATTGCCATT CTGTTGACAA TATCTGTTAT 600
 TTTACCTAGT TCAGTGTATA AACTTTTAAT ATATTTTAAA CGTGCTGGTG TATTATAAAA 660
 40 TAATGATTCT ACAAGTATAT CTGTTCCCTT TTTGCTTTT GCAGGCTTAT GATTTAATAT 720
 TTCACCATT TCTACATATA TTTCAATTCC ATTAGCATT TCCGTGCAAG TCTTCAATGT 780
 TACTTTAGCA ACTGATGAAA TACTGGCTAG TGCTTCACCA CGGAATCCTA ATGTCTAAT 840
 45 ATGAAATAAA TCTTCATCTT GATCTAATTT ACTAGTCGCA TGTCTATGAA ATACTAATCC 900
 TAAGTCTTCC GCTTCAATTC CGCTTCCATT ATCGACTACG CGAATAGATT GGACGCCAGA 960
 50 CTCTCTACT TCAATGCTTA TTTCTGTAGC GCCTGCATCT ATAGCATTTT CCAACAGTTC 1020
 TTTCAACA GAACTCGGTC TTTCAACTAC TTCACCTGCT GCGATTTTAT TTGCTAATGA 1080
 GGTGTTGAGT TCTTTAATTT TCCCCATTTT GCAACACCTC TATTTTAATT GATTTTGTAA 1140

	TTGTAGTTCA ATCTCGCTTT TTTGATCATT TTCAAACAAA TcAAATGATG CyTGTTCAAA	1260
	GTCTTTTTGA GATAAAGTAT CaGTTGTTTC TTCaACACTT aAGTTTAAAT TTTCTTGATT	1320
5	AATTTCCAGG TTCATTTTCG ACCATTTTTTA AATTTGATAT CGATGATtTT TCACCAGCAG	1380
	ACGCTTCAAA CTCGCTTAGA ATCACTTG TG CTCTGCTAAT AACTTTTTCA GGTAAATCAG	1440
10	CTAATTTTCG AACTTGAATA CCATAAtATC GTCAACTGCA CCATCTTTGA CTTTATGCAA	1500
	GAATATAAGT TCACCTTTAT ATTCAATTAGC AGCGACGTGA ACATTTTTTTA GACTTGGTAA	1560
	TGCTTGATCT AATGTTGTCA ATTCATGATA ATGTGTTGAA AATAACGTTT TAGCATGTGA	1620
15	TGTTTCAGCT ACATACTCTA TCATTGCCTG CGCTAAAGCT AAACCGTCAT ACGTTGAAGT	1680
	ACCACGTCCA ATTTTCATCGA AAATAATCAA ACTATCCTCT GTTGCCATAAG TTAATGCCTT	1740
	TTGTGCTTCT AGCATTCTTA CCATAAACGT ACTCTTACCT GAAACCAAAT CATCTGCCGC	1800
20	ACCTATTCTA GTGAATATTT GATCAAATAT AGGTAACACT GCCTCTTTAC AAGGGACATA	1860
	AGCTCCCAT TGGGCCATTA TACTAATTAT GGCAACTTGT CTCATATATG TCGATTTACC	1920
	AGACATATTC GGACCTGTAA TTAAATATAT AAATGTTTCA TTATCTAATC GACAATTATT	1980
25	AGGCACATAG TCATTATAAT CCATTACTCT TTCCACTACT GGGTGCCTAG ATTCCACTAA	2040
	TTCTAATGTT TTATTTTCAC TAAATGAAGG CCTAGTGTA TATATTTTTT GAGCAATTTT	2100
30	TGCAAAGCTC TGTAACAAT CTAGCTCTGA AATAATTTTA GCTTGTTGTT GTAAACGTTT	2160
	AGTATATTTT TTAACCTCTT CACGTAGCTG AACAAATAAT TGATATTCTA ATTCGATGGC	2220
	TTTGTCTTCC GCACCTAAAA TGATATCTTC TTTTCTTTA AGTTCATCAG TTATAAAACG	2280
35	TTCAGCATTC GATAACGTTT GCTTCCTCAT ATAACCAAAT TCACTTGGTT CAAAATTTTG	2340
	CAAGTTGGCA CGTGTTATTT CTATAAAATA ACCAAACACT TTATTAAGC TTATTTTCAA	2400
	TGATTTTATT CCTGTACGTT GTCTTTCTTT GGCTTGTAAT TCTGCTAACC ATGTTTTTCC	2460
40	GTTTTTTGAA GCTTCAAGAT ATTCATCTAA TTGCGTATTA AAACCAACTT TGAATAGTCC	2520
	GCCATCTTTA ACTGAAATTG GTGGTTCTTC TACTAACTC TGTTCTAATA TATCAAGTAA	2580
	ATCATCaAGG GGTTC TAGTT GATTAACCTG TACAAGAGTA TTCTGATTCA TAGAATTTAG	2640
45	TAATGCTTTA ATATTCGGTA TTTCAGAAAT GGAATGTTTA AGTTGAATTA AATCTCTCGC	2700
	ATTAACATTT CCGTAACTAA CACGCCCAAC AAGACGTTCA ATATCATACA CTTGATTAA	2760
50	ATATGTTCTT AAGGTGTCTC TTTCTATGAA ATGAGCACTA AATTCATCAA CGATATCTAA	2820
	TCGTGCTTCA ATTTGTTCTT TACTTATTAG TGGTCTATCT ATCCATTGTT TTAAGCGGCG	2880
	TGCTCCCAT GGTGTTTTCG TTTCGTCCAT TAGCCAAAGT AGCGTTCCTT TTTTGTATT	2940

55

ATCTATAGCT GCATATTGAA CAACATCCTC GATATGCGAT AAATCACGTT TTTGTGTATG 3060
 ATGAATATAA TCTAGCAATA ATTGTGTGCG TTAGATACaTT AATTTATGTT CAGTTTGATT 3120
 5 CACACTATAG ATTtCTGATG ATAACGTTTC CCTGACTGT 3159

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1238 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

TAAGCGAGAA GCAATTGGTG TTATGTATGC TAGTGATAAA CCAACAGGTG AAAGTACAAG 60
 20 GTCATTGCT GTTTATTCT CTCCTGAAAT TAAGAAATTT ATTGCAGATA ATTTAGATAA 120
 ATAAATCATC CATCCATACA TTGATAAATG ATTTTyAGAA ATTAACAACA AAATCAACAA 180
 TTTTAAACAT CTCTGTGATT CTATTTATTC GAAATGATTT AAAAAATAAA ACTTCAAAAA 240
 25 CCTAACCTTA TATTTATACG AATACTTAGA GGAGCACAAA AATGAATAAA AATATAATCA 300
 TCAAAAGTAT TGCAGCATTG ACGATTTTAA CATCAGTGAC TGGCGTCGGC ACAACAGTGG 360
 TTGAGGGTAT TCAACAAACG GCTAAAGCTG AACATAATGT GAAACTAATC AAAAACTA 420
 30 ATGTAGCACC ATACAATGGT GTCGTTTCGA TAGGATCTGG AACAGGTTTC ATTGTCGGTA 480
 AAAATACAAT TGTTACCAAC AAGCATGTCTG TTGCAGGTAT GGAAATTGGT GCACATATTA 540
 TAGCGCATCC CAATGGTGAA TATAATAATG GCGGATTTTA TAAAGTTAAA AAAATTGTCC 600
 35 GTTATTCAGG TCAAGAAGAT ATTGCCATTC TACATGTGGA AGATAAAGCT GTTCATCCAA 660
 AAAACAGGAA TTTTAAAGAT TACACAGGCA TTTTAAAAAT AGCATCAGAA GCTAAAGAAA 720
 40 ATGAACGCAT TTCAATTGTT GGCTATCCAG AACCATATAT AAATAAATTT CAAATGTATG 780
 AGTCAACAGG AAAAGTGCTG TCAGTTAAAG GCAACATGAT TATTACTGAT GCTTTCGTAG 840
 AACCAGGCAA CTCAGGTTCA GCTGTATTTA ACAGTAAATA CGAaGTtGTA GGTGTTCACT 900
 45 TTGGTGGAAG CGGCCCTGGA AATAAAAGTA CAAAAGGATA TGGTGTTTAT TTCTCTCCTG 960
 AAATTAAGAA ATTCATTGCA GATAACACAG ATAAATAAAT CCTTACATAG ATAAATGATT 1020
 TTAAAAATTA ACAACAAACT CAACaATTCA AATCATCTCT GTGATTCCAT TTATTCGAAA 1080
 50 TGATTAAAAA AAATAAAACT TCAAAAAGCT AACATTATAA TTATACAAAT ACTTAGAGGA 1140
 GCAGAAAAAT GaATAAAAAAT ATAATCATCA AAAGTATTGC AGCATTGGAC GATTTTTTAAC 1200

(2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6444 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

	TGATAAGTCA TTAAATTGT CACCTATTGA CATGACTTCT TTCATTTCAA TCCCTAATCT	60
5	TTCGGCAATT GTTTCTAGCG CAATACCTTT TTGTGCATCT GAATGCGTTA TTTCTATATT	120
	TCCTCTCGAA GATGATGATA TAGCTAAATT CGGAGAKTCA GCTAAAATTT TACTAGCTTT	180
	GTCAATTTTT TCTAAATTTT CATCAAATGC TAATATTTTC ATAATTAAAT CACCAGGTAT	240
10	GTTTTCAATA GCATCATAAT TATCAACAAC TyTCAACGTA CCATTATCTA TCGCTCTTTG	300
	AATACCATTT TTAATACGCT CAACGTTTGC ATGTTGACCT GCACGCTCAG CAATATCTAT	360
	GTAAATGTCT AAATCTCTTT GTGGATCTTC AGTATAAATC GCACGACTCG TGTATACTTG	420
15	ATAATAAATA CCTGCATCTT TTAAACATT TGTAATTTTG TGTACTAACG ATTTATTAAG	480
	GTGTGAAGTG CTCATTACAT TGAAAGTTTC ATCACGTA CTAGCACCAT TCAAACAAAT	540
20	ATATGGTACT GTTAAATCTG TGTCAGCAAC TGGTGCTTGK GCTTCATAAA ATGCTCGACC	600
	TGTCGCGATA ACAACCGTTA TCCCTTGTTT TTGAGCGTAT TTAATCGCAT CAATATTAGG	660
	TTGAGAAATT TCATGTGCTG CATTAAAGTAG CGTGCCATCC ATATCAGTGG CTATTAGTTT	720
25	TATCATTATG TnACCTCGTT TCGTAAATnT AAAATCTTGT TCTTAAATAA GrATATATAC	780
	TCAGCGCACA TACTTTTCTA TTAmCATTTA TATkGTCATT aATTTATCAT ATAATGTAAT	840
	TCTaACAAAT nTTAAATAGT ATGTACTATC GTCTAAITGG TGGATTTCTT ATTGGCTCTT	900
30	AAgTTTTTAA AAAATGTTGT TAATAATGTG CTACATGCTT CTTTAAGTAC ACCTTTATCA	960
	ACAATTGCAC GATGATTAAA ATTAGATTGT TGCAATAAAT TCATTAAACT GcCACTACAA	1020
	CCACCTTTAG GATCATCTGC GCCATAGACG ACTCTTGGA TGGGACTCAT TACAATTGTT	1080
35	CCTGCGCACA TGACACATGG TTCTAAGGTT ACATATAATG TGCAACCTTC TAAACGCCAA	1140
	CTACCTAACA CTTTGGCTGC ACGTTCAATT GCAATATGTT CAGCATGCGC CGTTGGTTGT	1200
40	TGTAGTGT TTCTTaAATT ATGTGCTCTA GCGATAACTT CATCATCTTT AGTGATGATA	1260
	GCACCTATAG GtACTTCGCC TAGTTGAGCT GCTTTTTTAG CTTCTTCAAT CGCTAATGTC	1320
45	ATAAAATATA TATCATTGTG CATTtATGTC CAGATACCTC ACTTATGGTA CAATACTCAA	1380

	CTATTGGCGT AGGTAAATCT TCACTTGCAC ACAAATTAAG TCAAACCTTTA GATTTTTATG	1500
	AAGAAAAAGA AATCATCACA GAAATCCAT TTTTATCAGA CTTTTATGAA GATATCTCTA	1560
5	AATGGAGTTT TCAAAC TGAA ATGTTCTTTT TATGCAATAG ATATAAGCAA TTTCAAGATG	1620
	TAACACAAC TAAATCAAGGT GTAGTTAGTG ATTATCATAT ACATAAAAAAT AAGATATTTG	1680
	CTAAAAATAC TTTGAGTTCT GTTGAATTC AGAAATTCAG TAAAATTTAT GATATTTTAA	1740
10	CTGAAGATAT GATTATGCCG AATATGATTA TCTTTTTAGA TGCAGACCTT GATGTGTTAA	1800
	AATCTAGAAT TGCTAAACGT AACCGTAGTT TTGAGCATCA AATAGAAAGT AtActGTaAg	1860
15	TTAAAAAAG ATTATCGTGA GTATTATGAG TCCTTACAAA GTAATGGTTC AAATGTAGTT	1920
	TTAATCGATA CnACTTCTAT TGATTTTCTT AAAAATGAAC AAGATTACGA AGATATATTA	1980
	CATATTATAT TACCTATGAT AGGAGATATT ACCAATGAAT AATTACGGTA TTCCACAAAA	2040
20	TGCCATTATA ACCATTGCAG GTACAGTTGG TGTTGGAAAA TCAACACTAA SGCAAGCACT	2100
	TGCAGATAAA TTAAACTTTA AAACGCTTTT TGAAAATGTC GAACATAATC CATATTTAGA	2160
	TAAATTTTAC AGCGATTTTG AACGATGGAG TTTCCATTTG CAAATTTACT TCTTAGCTGA	2220
25	ACGTTTTTAAA GAACAAAAGC GTATGTTTGA ATATGGTGGT GGCTTTGTCC AAGATCGATC	2280
	AATTTATGAA GATGTTGATA TTTTTCGAAA AATGCATGAA GAAGAAGGCA CAATGAGTAA	2340
	AGAAGATTTT AAAACATATT CAGACTTATT TAATGCCATG GTCATGACAC CTTATTTTCC	2400
30	TAAACCTGAT GTAATGATTT ATTTAgAATG TAACTATGAT GAGGTCATTG ATCGTATTAT	2460
	TGAACGTGGT CGCGAAATGG AAATTAATAC AGACCCTGAA TACTGGAAAA AGCTATTTAA	2520
35	ACGCTATGAC GATTGGATTA ATAGCTTTAA TGCAATGTCCA GTTGTTACGTA TCAATATTA	2580
	TGAATATGAT ATCCATAAGG ACCCCGAATC TTTAAATCCT ATGATAAACA AAATTGCTCG	2640
	AATTATTCAA ACATATCGAC AAGTAGATAC ACGATAAAAG ACTAAAGACA TAGCGTATAT	2700
40	GTTTATATTC AATGTATATT CCATAGATAT TATCGATTAT TTTATCAATT CTATCGAATA	2760
	CATTAATTCA CATATACACT ATGTCTTTCT TTTAATTTA AAGCTTCTAA AATATCTGCC	2820
	GCACTATTTA AAATAATATC AGCTTCATGT AATCTTCTT TTGTTGCAAT ACCTGTTAAT	2880
45	ACACCTATTG CCATACCTAA ATTTGCATTA CTTGCTGTCT TCATATCATT AGCAGTGTCT	2940
	CCTACTATAG CTACTTTCTG AGGATCTACA TTATATTGCT CAAATAAAGG CGATAATACT	3000
	TTAGGATTTG GCTTCTCATA GGCATCCGCT TCGGTAGAAA TGATCAAATC GAACAACGAG	3060
50	GTAGCATTGG TATGTGCTAA AAATTGTTCT ACACCTTTTT TAGTATCACT CGTAACAATA	3120
	CCAAGTTGAT AGCCTTTTGC TTTCAAATCG ATAAGTGCTT CTTTAACACC TTCTACCCAA	3180

55

	GTATCTTGTC CCGTCACATC ATTAAATGCC TGGATAATTT GTTGTAAGA TCCTGAACCC	3300
	ATCACTGATT TTGGATCAAT AGATTCTTTA ATGACACCGA GTTGTCTTAA AGCAGCTTCT	3360
5	TTATTATGTA CTGGGAAAGT CTCAAGCAAT GATTGTACAA ATCGTACCCC TATTTTTTCC	3420
	CAACTTCTAT CAAATTCAAT TAACGTACCA TCTTTATCAA ATAATATCCA TTCCATTGaT	3480
10	ATCAATACTC CTATTTATTT ATTTCTGTAAT ATGCTGATTC TATGATATTC GTTATCCCCT	3540
	GAAAATGAAC TCGTAGTATT GTTCTATTTA AATATTGaAT TAAATATAAT AATAAGTGAA	3600
	ATCCCCTTCA ATACTTAACA ATAAACATTG TAAACTTAAT TTATTACCAT GCTTCGCTTC	3660
15	ATTGAAAGGG ATTTTAGTCA TGATTAACTT TTGCATATTG TTTTCATGAT TATATTCAAT	3720
	TTTTATTAAT ATTTTGGTAC AACGACTCTC CAACCATTTT TATCTTCTAA AGTACCATTT	3780
	TGAATACCAG TATAGACGTC GTATAATTTT TGAGTAATTT CACCAGTCTC ATTATTATTA	3840
20	ATAACGATTT CACGATCTTC GTATCTCAAT GTACCCACAG GTGAAATAAC TGCTGCAGTA	3900
	CCACTACCAA ATACTTCTGT TAACTCACCT TTATCATATG ATTCGAATAA TTCATCGATT	3960
	GAAACGCGGC GCTCTTCGAC TTCATATCCT AAGTTTTTAG CTAATTCGAT AATAGATTTA	4020
25	CGTGTAATAC CAGGTAAAAT ACTGCCATTC AACTCTGGTG TAATTACTTT GCCATTTTCA	4080
	ACGAAGAAAA TGTTTCATGCT ACCAACTTCT TCGATATATT TCTGTTCAAC ACCATCAAGC	4140
30	CATAATACTT GGTCAATAAC TAATTTATTT GCATTAGTTT GTGCTAATAA ACTTGcCGCA	4200
	TAGTTACCTG CAACTTTTGC AAAGCCTACA CCGCCACGaA CAGCACGCAC ATATTTCATCT	4260
	TCTACATAGA TTTTAGTTGG TTTTAAAGTT TCACCACCAT AATATGCACC TGAAGGAGAT	4320
35	AAAATAATTA ATAATTTATA CTGATGTGAT GCACCAACGC CAAGTGCCCC TTCTGTTGCA	4380
	AAAACAAATG GACGAATATA TAATGATTGA CCTTCCCTT CAGGAATCCA ATCTCTTTCA	4440
	ATATEAACTA ATTGTTTTAG CCCCTCTAAC AATTCTGCTT CGTCTACTTG AGGCATTTCT	4500
40	AATCGTGCTA ACGAGTTATT AAGACGCTTA AAATTTTCTT CAGGACGGAA AAGTGCAACT	4560
	TCCCCATCTC TTTTATATGC TTTTAATCCT TCGAATACCG ATTGACCATA ATGAACACCT	4620
	TGTGCAGCAG GTGAAATTTT AATAGGACCA TAAGGTACTA TCTTCAAATC ATGCCATCCT	4680
45	TTATCTGCAT CATAATCATA ACTCAACATA TAATCAGTAA AATATTTACC AAAACCTAGT	4740
	TGAGATGTAT TTGGTTTTTG TTTTAATGTT TCTCGTCGTT CAACTTTAAC TGCTTGTGAC	4800
50	ATGGTGATTG CCTCCTAATA ATATTGTATA AGAATTTGTT TAACTTAAAT TATAACAATC	4860
	CaTATTTTGC TGTTCAACAA ATTTTCTAAA AATTCAAAT TAATTAACAG ATTTCTAGAA	4920
	AGACTATATC TTTTAGTATA AACGTATTAA TTTACACAGG ACAAGTAATC TGTGTTTTAC	4980

55

TAAGTATAAT GAATAATATT AGAATTCATG CACTAGTTTA TTAAAATAAA GAGTAATTTA 5100
 AAATATCATT CCGTGTATTA AAGTGAATGG AAATGATTAG TTATTATTTT TAACAGTATC 5160
 5 TTTTGTTC AATAGCTTCTA ACATTAATTT AGTCATGCTC GCTAAATCAT ATTTAGGATC 5220
 AAATCCCAT TCGCCACGTG AACAGCTTGT ATCAATAGAA TCCGGCCAAC TATTAGCGAT 5280
 ACCTTGTCTA ATAGGATCAA CATCGTAATC TAATGTAAAA TTGGGATAGT ATTCTTGAAT 5340
 10 TGCTTCTTTT ACCATCTCTG GATCAAACT CATTGCGCTC AAATTATAAC CATTTCTAGT 5400
 TTCTAATTTA GCGTCGTCTG CTTCCATAAG TTTAATAATT GCTTCAATTG CATCATCCAT 5460
 ATACATCATA TCCATATACG TGCCTTTATC TATGAAGCTT GTATAATGAC CCTCTCTTAC 5520
 15 TGCTTTGAAG TATATTTCAA CAGCATAGTC TGTAGTACCG CCACCTGGCT CTTTAACATG 5580
 CGAGATTAAA CCTGGGAATC TAACACTTCT TGTATCTACA CAAAACGTT TGAAATAGTA 5640
 20 TTGACACAAT AATTCTCCAG CTACTTTATT TACACCATAC ATTGTCGTAG GTTGCTGAAT 5700
 CGTTACTTGT GCGGTATTAA CTTTAGGAGT TGAGTCTCCA AATGCACCAA TTGAACCTGG 5760
 TGTGAAAAAG TGCAAATTAT AAGTTCTTGC AGCTTCTAAT GCATTCATTA ATCCACCCAT 5820
 25 ATTTAAATCC CAAGCTAGAA TTGGATTTT CTCAGCAGTT GCTGATAATA ATGCTGCCAT 5880
 ATGCATTAGA CTATCCGCTT CAAAGTCCCT AACTAATCA AACATACGGT CACGATCTGT 5940
 TACGTCTAAG ATTTCAAATG GTCCATTTTG TACAGGTGAG TCTGCTTCAG GTTCCCTAAT 6000
 30 ATCTGTAGCA AGAACATTAT CTGTCCATA AATTTCTCTG CACTTAACAA CTAATTCTGT 6060
 ACCAATTTGT CCTAATGCAC CAGTAATCAT AATTTTTTTC ATAGAAATAT CTCCTTTGtC 6120
 TCTTCTATAT AGCTATAGTC CATCACAAGC GgAcATAATA TTCATTTTCA TAATAATTAT 6180
 35 AATATAAAAG CGCTTTCTTG TATATATGAC ATGTACATGT TGCTGATATk TCTGTAAATG 6240
 GAAATTCTAG TTGTATTAAT TGATTTTAGT AATTTATAGC GTTTATTATT GCTAATTACT 6300
 40 GtCAAATTAA ATTTTTTATC CCTCAACTCT TAAACTCTGG ATATCTTTCA TTATATTAGC 6360
 TTTTTTATAA CCATGGATAT CATGTAAAGC CTTATAAGCn TTAAATAATG TTTCATACCT 6420
 TTGTACTTnT TCCGCTTCTG GATT 6444

45 (2) INFORMATION FOR SEQ ID NO: 234:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4721 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

55

	GCCCATGAGA CAATTTTACT TGCTTTTCCC ATTGGTTATC ACGTTCTTAT TACATAGATT	60
	TAAACCGAGA AATATTATTC AAACGCTATT TATTGTATCG TTGATTTCTT TAGGACTTAT	120
5	GATAGTGATT CATTTTCATCA CTGGAGATAA TTCACGTGTG TATTTTGGA CAGATACACG	180
	ACTGCAAACT TTATTGCTTG GTTGTATATT AGCATTTAATT TGGCCTCCGT TTGCTTTGAA	240
	AAAAGATATT TCTAAAAAGA TTGTCGTATC ATTAGATATT ATAGGGATAT CTGGTTTTGC	300
10	GGTTCTAATG ACTTTGTTCT TTATAGTTGG AGACCAAGAT CAATGGATCT ATAATGGAGG	360
	ATTTTACATT ATATCATTTG CAACTTTAATT CATTATTGCA ATTGCGGTAC ATCCTTCTAG	420
	TTTATTTGCT AAATTTTTAA GTATGAAACC TTTACTAATT ATAGGTAAAC GATCATATAG	480
15	CTTATACTTA TGGCATTATC CTATCATTGT TTTTGTGAAC AGTTATTACG TACAAGGACA	540
	AATACCGGTA TACGTTTATA TTATAGAAAT TTTGTTAACA GCGTTAATGG CTGAAATTC	600
20	GTATCGCTTT ATTGAAACAC CTATACGTAA AAAAGGATTT AAAGCTTTTG CATTTTTACC	660
	TAAAAAGAAG GGGCAATTTG CTAGAACAGT GTTAGTTATC CTATTATTGG TTCCGTCTAT	720
	CGTTGTGCTC AGTGGACAGT TTGATGCACT TGGCAAACAA CATGAAGCCG AGAAGAAAGA	780
25	GAAGAAGACG GAATTTAAAA CAACGAAGAA AAAAGTCGTT AAAAAAGATA AGCAAGAGGA	840
	TAAGCAGACA GCGAATAGCA AAGAGGATAT TAAAAAGTCA TCACCACTAT TAATTGGTGA	900
	CTCGGTCATG GTGGATATTG GTAATGTCTT TACTAAGAAA ATACCAAATG CACAAATTGA	960
30	TGGTAAAGTT GGACGGCAAC TCGTTGATGC TACACCAATT GTGAAATCGC AATATAAAGA	1020
	CTATGCTAAA AAAGGTCAAA AAGTTGTAGT AGAGCTTGGT ACAAATGGGG CATTTACGAA	1080
	AGATCAATTA AATGAACAT TGGATAGTTT TGGAAAAGCA GACATATATT TAGTTTCTAT	1140
35	TAGAGTACCT AGAGATTATG AAGGTAGAAT AAATAAATTA ATTTATGAGG CAGCTGAAAA	1200
	GCGCTCTAAT GTACATCTAG TCGATTGGTA TAAAGCTTCT GCAGGTCATC CGGAATACTT	1260
40	TGCATATGAC GGTATTCACT TAGAATATGC AGGTAGTAAA GCGCTGACTG ATTTGATTGT	1320
	AAAAACGATG GAAACACATG CTACAAATAA GAAATAATTT GATGCACTAA ACTTTTGAAA	1380
	TATTACATTA CTTCTGATAT TTATTATCAA AAATGATGTA TTTCAITAAA AGTTTAGTGC	1440
45	TTTTTTATTT TCAAATCCCA TAGTAACGGT GCAGAAAAAG TGTTGTAAAC ATTCTAATTG	1500
	GTATAITACA TTCAATGAAG CTTTATTAGG AACAGATTAC ATTATGATAA CAAAGCCCGC	1560
	AAGACACCTA ATCTCTGTTA TAGTTTGTTT TGTCGCAAAA CTATAAAAGT TATAATTGTT	1620
50	TGCATACTAA AAAAATAAAA AATATAAAAT TTAAATAAT TGAGTCGCTA ATGACTATAT	1680
	GTATAGAGTG TTTTGATTAT TGGGAGGATA TTTAATTATG AAAAAAATCG TTACAGCTAC	1740

55

	ACAAAATAAC AATGGATATA ATTCTAATGA CGCTCAATCA TACAGCTATA CGTATACAAT	1860
	TGATGCACAA GGTAATTATC ATTACACTTG GACAGGAAAT TGGAATCCAA GTCAATTAAAC	1920
5	GCAAAACAAC ACATACTACT ACAACAATA CAATACTTAT AGTTATAACA ATGCATCTTA	1980
	CAATAACTAC TATAATCATT CATATCAATA CAATAACTAT ACAAACAATA GTCAAACAGC	2040
	AACAAATAAC TATTATACTG GTGGTTCAGG TGCAAGTTAT AGCACAACAA GTAATAATGT	2100
10	TCATGTGACT ACAACTGCAG CGCCATCTTC AAATGGTCGT TCAATTTCTA ATGGTTATGC	2160
	ATCAGGAAGT AACTTATATA CTTCAGGACA ATGTACTTAT TATGTATTTG ATCGTGTGG	2220
	TGGGAAAATT GGTTCACAT GGGGTAACGC AAGTAATTGG GCTAACGCAG CTGCATCATC	2280
15	TGGCTATACA GTGAACAATA CACCAAAAGT TGGTGCTATC ATGCAAACAA CACAAGGCTA	2340
	TTACGGTCAT GTTGCTTACG TTGAAGGCGT TAACAGCAAC GGTTCGTTC GTGTTTCAGA	2400
20	AATGAACTAT GGACATGGTG CTGGTGTGGT TACGTCTCGT ACAATTTTCTAG CAAACCAAGC	2460
	AGGTTTCATAT AATTTTCATTC ATTAATCAAA TGTAATCAAA ATGACGTCAA TATTCTCTAA	2520
	CATGAGAGTA TTGGCGTTTT TGTTTTATAT AAATATAAAT GAGAGCGGTT TATTCCTGA	2580
25	TCTTTAGGGA ACTAAGTAAT AAAGTGATAA TTTATACTAT GTCAGTATGA TTGAGAGTGA	2640
	TTCAATTTAG ATGAAAACCA TGAAAAAATA TATTAAAACA GCATTTTTTTT GTAGTATGTA	2700
	TTGGTTAATT GTTCAACTAA ATATAGCAAA TTTAGGTACA AGAATTCCTG ATAAGTATTT	2760
30	TCGTCAGAAG TACATAATAT TTAAATCATT TAACTTTGAG AAGCATGGAA AATTTTGGAA	2820
	CAAATGGTTT TACGTAAGAA AATGGAAACA TAAGATTTTA GATGGTCATC AGCTTAATCA	2880
	AAATATATAT GATCAGCGTC ATTTAATGAC AATCAATACT GATGAAATTG AAAAAATGAT	2940
35	TATAGAGACA AAGAGGGCAG AGTTGATTCA TTGGATATCG ATACTTCCAG TCATCATATT	3000
	CAATAAAGGC CCTCGTTTAC TAAAGTATAT AAATATTTTC TATGCAATGA TAGCTAATGT	3060
40	TCCAATCATT ATTGTGCAAC GCTATAATCG ACCGAGATTA ACGCAGTTAC TACGCATATT	3120
	AAAACGAAGA GGTGAACGTC ATGACTAAAC ATATCATCGT TATTGGTGGT GGCTTAGGTG	3180
	GGATTTCTGC AGCAATTCGA ATGGCACAAA GTGGCTATTC GGTCTCATTA TATGAACAAA	3240
45	ATAATCATAT AGGAGGCCAA GTGAATCGTC ATGAATCAGA TGGCTTTGGC TTTGATTTAG	3300
	GTCCATCTAT TTTAACGATG CCTTATATTT TTGAAAAATT ATTCGAATAT AGCAAGAAGC	3360
	AAATGTCAGA CTACGTTACA ATCAAGCGAT TGCCACATCA ATGGCGTAGC TTTTTCCTAG	3420
50	ATGGAACGAC TATCGATTTG TATGAAGSTA TTAAAGAAAC AGGTCAGCAT AATGCGATAT	3480
	TGTCGAAACA GGATATAGAG GAACTGCAAA ATTATTTGAA TTATACAAGA CGAATCGATC	3540

55

5 TTCATGGGCC ATTAAATGCT CTTATTAATT ATGATTATGT ACATACTATG CAACAGGCCA 3660
 TAGACAAGCG TATCTCGAAT CCATACTTGC GACAAATGTT AGGCTATTTT ATCAATATG 3720
 TAGGTTCTTC ATCATACGAT GCGCCAgCTG TATTATCTAT GTTATTCCAT ATGCAACAAG 3780
 AGCAAGGCCT TTGGTATGTA GAAGGTGGAA TCCATCATT AGCCAATGCC TTGGAAAAGc 3840
 TAGCGCGTGA AGAAGGTGTC ACAATTCATA CAGGTGCACG TGTGGACAAT ATTAAAACAT 3900
 10 ATCAAAGACG TGTGACGGGT GTCAGATTAG ATACAGGTGA GTTTGTAAAG GCAGATTATA 3960
 TTATTTCAAA TATGGAAGTC ATACCTACTT ATAAATATTT AATTCACCTT GATACTCAAC 4020
 GATTAAACAA ATTAGAGAGG GAATTTGAGC CGGCAAGCTC AGGATATGTG ATGCATTAG 4080
 15 GTGTTGCTTG CCAATACCCG CAATTAGCAC ATCATAATTT CTTTTTTACG GAAAATGCTT 4140
 ATCTCAATTA TcAACAAGTT TTTcATGAAA AGGTATTGCC AGATGATCCG ACCATTTATC 4200
 TAGTAAATAC GAATAAAACT GATCACACAC AAGCGCCAGT AGGTTATGAA AATATCAAAG 4260
 20 TCTTACCACA TATTCCATAT ATTCAAGATC AGCCTTTTAC CACTGAAGAT TATGCGAAGT 4320
 TTAGGGATAA AATTTTGGAT AAATTAGAAA AAATGGGACT TACTGATTTA AGAAAACACA 4380
 25 TTATTTATGA AGATGTTTGG ACACCGGAGg ATATTGAAAA AAATTATCGT TCTAATCGTG 4440
 GTGCAATATA TGGTGTGTA GCAGATAAAA AGAAAAACAA AGGATTTAAA TTTCTTAAAG 4500
 AAAGTCAGTA TTTTGAAAAC TTGTACTTTG TAGGTGGATC AGTAAATCCT GGTGGTGGCA 4560
 30 TGCCAATGGT TACATTAAGT GGGCAACAAG TCGCAGCAAg ATAAACGCGC GAGAAGCGAA 4620
 GAATAGGAAG TGATATCTAT GAAATGGTTA TCACGAATAT TAACAGTAAT AGTGACCATG 4680
 TCTATGGCGT GTGGTGCaTT GATATTTaAT CgTAGACATC A 4721

(2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3516 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

TATTCGTGCG CAATGGGCTA AATTAGGTCT AGGTTTAGAT TATAGTAGAG AACGTTTAC 60
 TTTAGATGAA GGTTTAAGTA AAGCAGTTAA AAAAGTTTTT GTTGATTAT ACAATAAAGG 120
 50 AATTATTTAT CGTGGCGAAC GTATTATAAA TtGGGATCCn AAAGCACGTA CAGCTTTATC 180
 TGaTATTGAA GTAATACATG AAGATGTTCA AGGTGCGTTT TATCATTTTA AATATCCTTA 240

	TACAGCGATT GTTGTTAACC CTAATGACGA ACGATACAAA GATGTAATCG GTAAACTGT	360
	TATATTACCA ATCGTAGGAC GCGAACTGCC TATTTTAGCA GATGAGTATG TTGATATAGA	420
5	CTTCGGTTCT GGTGCTATGA AAGTGACACC AGCACATGAC CCTAATGATT TTGAAATTGG	480
	TCAAAGACAT CAATTAGAAA ATATTATCGT TATGGATGAA AATGGTAAAA TGAACGACAA	540
	AGCGGGTAAA TATGAAGGTA TGGACCGTTT TGATTGTCGT AAACAGCTAG TTAAAGATTT	600
10	AAAAGAACAA GATTTAGTTA TCAAGATTGA AGATCATGTT CATTCTGTAG GTCATTGAGA	660
	ACGATCTGGC GCTGTTGTTG AACCATATTT ATCAACACAA TGGTTTGTGC GCATGGAAGA	720
	CTTAGCGAAA CGTTCATTAG ATAACCAAAA AACAGATGAT CGTATTGATT TTTATCCGCA	780
15	ACGTTTCGAA CATACTTTA ACCAATGGAT GGAAATATT AGAGATTGGA CGATTTCAG	840
	ACAATTATGG TGGGGTCATC AAATTCCGGC TTGGTATCAT AAAGAAACAG GCGAAATATA	900
20	TGTTGGAGAA GAAGCGCCAA CTGATATTGA AAATTGGCAA CAAGATGAAG ATGTATTAGA	960
	TACGTGGTTC TCaAGTGCTT TATGGCCTTT CTCyACGTTA GGTGGCCTG ATTTAGAAAG	1020
	TGAAGACTTT AAACGATACT ACCCAACAAA TGCCTTAGTT ACAGGTTACG ATATTATCTT	1080
25	TTTCTGGGTA GCACGCATGA TATTCCAAGG CTTAGAATTT ACAGATCGTC GTCCATTAA	1140
	TGATGTATTA TTACACGGT TAGTTCGTGC TGAAGACGGG CGTAAGATGA GTAAATCATT	1200
	AGGTAATGGT GTGGATCCAA TGGATGTTAT TGACGAATAC GGTGCTGATA GCTTGCGTTA	1260
30	CTTCTTAGCA ACAGGTTTAT CTCCAGGACA TGATTAAAGA TACTCAACTG AAAAAAGTTGA	1320
	GTCAGTGTGG AACTTTATCA ATAAAATCTG GAATGGGGCA CGTTTCAGTT TAATGAATAT	1380
	CGGTGAAGAC TTTAAAGTTG AAGATATCGA TTTAAGTGGT AACTTATCAT TAGCAGATAA	1440
35	ATGGATTCTA ACACGTTTAA ATGAAACGAT TGCAACAGTT ACTGATTTAA GTGACAAATA	1500
	TGAATTCGGC GAAGTTGGAC GTGCATTATA TAATTTTATT TGGGATGATT TCTGTGATTG	1560
40	GTACATTGAA ATGAGTAAAA TTCCAATGAA TAGTAATGAT GAAGAACAAA AACAAAGTTAC	1620
	ACGTTTCACTA TTGAGTTATA CTTTAGACAA TATTATGAGA ATGCTACATC CATTATGACC	1680
	ATTTGTAACA GAGAAAATAT GGCAAAGTTT ACCACATGAA GGTGACACAA TTGTTAAAGC	1740
45	TTCATGGCCA GAAGTGCGTG AATCATTGAT TTTTGAAGAA AGTAAACAAA CAATGCAACA	1800
	ACTTGTGTA ATCATTAAAT CTGTAAGACA ATCACGTGTA GAAGTAAATA CGCCATTGTC	1860
	TAAAGAAATA CCTATTTTAA TTCAAGCTAA AGATAAGAA ATTGAAACAA CTTTATCACA	1920
50	AAACAAAGAT TATTTAATCA AATTCTGTAA TCCTAGTACC TTAAATATTA gCtGACGTGG	1980
	AAAwTCCTGA GAAAGCAATG ACaTCAGTTG TAATTGCAGG TAAAGTGGTA TTACCATTAG	2040

55

AAAGCGAATT AGATAGAGTA GATAAAAAGC TCTCTAATGA AAACTTTGTA AGTAAAGCAC 2160
 CTGAAAAGGT TATAAATGAA GAAAAACGTA AAAACAAGA TTATCAAGAA AAATATGATG 2220
 5 GTGTGAAGGC AAGAAATTGAA CAATTAAAAG CATAGGAGTT AGTAACAATG AATTACCTAG 2280
 AGAGCTTGTA TTGGATACAC GAAAGAACTA AATTGGCAT CAAACCAGGT GTTAAACGTA 2340
 TGGAATGGAT GCTAGCACAA TTTAATAATC CTCAAAATAA CATTAAGGGT ATTCATGTAG 2400
 10 GTGGCACAAA TGGTAAAGGC TCTACAGTTG CTTACCTTAG AACAGCTTTA GTTGAAAATG 2460
 GTTATGAAGT AGGTACATTT ACGTCGCCGT TTATTGAAAC ATTTAATGAA CGAATTAGTC 2520
 TAAATGGTGT GCCAATATCA AATGACGCTA TTGTAGAATT AGTATCACGT ATTAAACCAG 2580
 15 TAAGTGAAAT GATGGAACGT GAAACAGATT TAGGTGTTGC AACTGAATTC GAAATAATCA 2640
 CAGCGATGAT GTTTTTATAT TTTGGTGAAA TACATCCTGT TGATTTTGTC ATTGTTGAGG 2700
 20 CTGGATTGGG TATAAAGAAC GATTGACAA ATGCTTTTAC ACCGGTTTTA TCAATCTTAA 2760
 CTAGTATCGG TCTAGACCAT ACAGATATTT TAGGTGGTAC TTATCTAGAT ATTGCTAGGG 2820
 ATAAAGGCGC GATTATAAAG CCTAACGTTT CAGTGATATA TGCTGTTAAA AATGAAGATG 2880
 25 CATTAAATA TGTTCTGTGA CGCGCAATTG AACACATGC AAAGCCAATT GAATTAGATA 2940
 GAGAAATTGT TGTGTATCG CAAAATGATG AATTTACTTA CCGTTATAAA GATTATGAAT 3000
 TAGAAACAAT CATTTTAAGC ATGTTAGGTG AACATCAGAA ACAAATGCT GCATTAGCCA 3060
 30 TAACAGCTCT TATTGAATTA AATGAACAAG GATTAATTGA ATTAGATTTC AATAAGATGA 3120
 TAGACGGTAT TGAATCAGTT CGTTGGACTG GACGTATTGA GCAGGTGCAT GACAAACCTT 3180
 TAATCATTTT GGATGGCGCA CATAATTCAG AGAGTATAGA TGCTCTAATT GATACAATTA 3240
 35 AACAGTACCA TGATAAGAA AAAGTAGATA TTTGTTCTC AGCAATAAAC GGAAAACCGA 3300
 TTAACGAGAT GGTCAAACAT TTAAGTTTAA TTGCGCATAC GTTTTATGCA ACTGAATTTG 3360
 40 ATTTTCCGAA AGCGTTACGC AAAGAAGAAA TTGTAGGTAG TATTGAAAAT GATGAAATAC 3420
 AATTAGTAGA TGACTACGTT GAATTTATAA AAAATTATCA AGGTGATACA TTAGTAATTA 3480
 CCGGTAGTCT GTATTTTATA AGTGAAGTTA AATCAA 3516

45 (2) INFORMATION FOR SEQ ID NO: 236:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7481 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

55

	TGAGTGATAG AATCAAAAAA GCCATCTCAA AAATTAATCA AGCAAACAAC ATTCCAAACA	60
	ATGsTCGCAA ATCACCAATG TATCACTCTC CAATTACGTA ACTATGATTT AATTTAAGCA	120
5	TAGTTATTGA GGTTTTGTGA TATATAGTAT AAAATTAATG AGAATTAAAT TTAATAATGT	180
	AAAATTCATm TTCgGGGTCG GGTGTAATTC CCAACCGGCA GTAAATAAAG CCTGCGACCT	240
10	GCTAGTATGT ATCATATTAG TGGCTGATCT AGTGAGATTC TAGAGCCGAC AGTATAGTCT	300
	GGATGGGAGA AGATGGAGGT TTTTGTGTGT GCAATAATCC TCCTATTCTT ACGAGATGAA	360
	TGGAAGGAGA AAATTGAATA TGCAACAAAA TAAACGTCTT ATCACAATAA GTATGTTGAG	420
15	CGCGATTGCG TTTGTGTAA CTTTTATCAA GTTTCCTATA CCATTTTTC CACCATACTT	480
	AACTTTAGAT TTTAGTGATG TACCGTCACT ACTAGCTACA TTTACGTTTG GACCAGTTGC	540
	CGGTATCGTA GTTGCACTGG TTAATAATTT ATTGAACTAC TTATTTAGTA TGGGCGATCC	600
20	AGTTGGACCA TTTGCTAACT TTTTAGCAGG CGCAAGTTTC TTATTAAGTG CTTACGCCAT	660
	CTATAAAAAT AAACGTTCAA CAAAATCTTT GATTACTGGA TTAATCATTG CAACAATCGT	720
	TATGACTATC GTGTTGAGTA TTTTGAACATA TTTGCTTCTA TTACCTTTGT ACGGTATGAT	780
25	ATTTAACTTA GCTGATATCG CAAATAATCT TAAAGTAATC ATTGTTTCAG GAATTATACC	840
	ATTCAATATT ATTAAAGGTA TCGTTATTTT TATTGTATTT ATTTTACTAT ATAGAAGGCT	900
	TGCGAATTTT TTGAAAAGAA TTTAATCAAA TTAAAGCAAA ATAATATACA CATAATAATA	960
30	AAAAGCAGGT GACTATCAAT AAACGATAGC TTGCCTGCTT TTTCTATAGA ACATTTGTCT	1020
	AAAAATCAA TTATTCaAT TTTAAAGCGT CCCCATCAAA TGATTGCTCT GCAATTTTAA	1080
35	TAGAATCTGt AGGGCATCCa TCAATTGCAT CTTCCATATC TTCATATAAT TCCtCAGGTA	1140
	CTTCTGCAGT ACCTTGGTTA TCGTCAAGGA TTACGAAAGC AATACCTTCG TCGTCGTAAT	1200
	CATATATATC TGGCGCTGCT GCACCGcATG CACCACATGC AATACAAGTA TCCATATCAA	1260
40	CGATTGTATA TTTTGCCAAT GTCTTCGCCT CCTTTGATAA AAATGCTAAA ATAGTAATGT	1320
	GACTAAAATT TTAGACAGCA TCATTTTAT TTTCAAATTA TCCGTTTAC AGAGTGAGGG	1380
	TTAAATTTGC AACACATTAT AAAACAGCA TTACAACAAA CATTTAACTA TAAAACAAAT	1440
45	AAAAGTATTT ACAATATCTT AGTTGGTAAG AAATCTCACC AAACCTTTTT TGACGCTTGT	1500
	AGTCAACAAC AGTTGTCATT ATATCACAGT TTACCACTAT TAAAATATCC GTCTTTTGAG	1560
	CTATTTtTAG AAAAAATCAA TGAATTTAAT GCTGAAATGG AAATCATGTT GCATCCTAGA	1620
50	TATACATTTG aAAGCATGGG GCAAACATTT CAAGCAATTC AACTATTAGT GCAAaCCATG	1680
	TCTAATACCA AACAACATGT TTTTCATTTT GTACCAATCT CTCAAAATAA TAAGATACAm	1740

55

	AATGAACTAC	ATAACTTGTT	TAAAGCAATA	ACTTTAAAAG	GGCCATGTTA	CTTACATTAT	1860
	TATTTGCAAG	GCTATGATGA	ACCAATGTAT	ACGAGACAGC	AAGTTAGTTT	AATAGAAAAG	1920
5	CTATCTCAAC	AGCAATTGTT	TGAATACGAA	ATGAATAATT	TAGTGACAAT	GATGTTTGAA	1980
	TTAGAAAGTG	GAGAATATAC	TATTTTATCA	AAAATAATAA	TGAAACCTAC	ATTATTAAAT	2040
10	CAAACTTATA	TTACTTATAC	AAAATTGCTT	GAACAATTCA	CGATGGAAGA	TATAGCGGCT	2100
	CAACAACAAG	TTAAAATCAA	TACTATCGAA	GATCATGTAC	TTGAAATCTT	AATCAAAGGT	2160
	TACATGTCTA	ATTACGATGA	TTATGTTGAA	CTAGAAGATC	AACTCCAGTT	TTGAATTTT	2220
15	TATCAACAGC	ATCGTGGCGA	ACGATTAAAA	TTTTACAAAG	AACAATTTGA	CACGTTATCA	2280
	TATTTTCAAT	TAAAAGTATT	AATCGTTGGA	TTTGAAAGAG	GTGATCTGAA	TGTTGCATGA	2340
	TATTTTACGA	AACAAATTTG	GATTCGAGAG	TTTTAAACCG	GGACAACAGG	AAATTATAGA	2400
20	AAGTATAATG	TCTCAACAAC	ACACTCTAGG	TATACTTCCA	ACTGGAAGTG	GAAAGAGTTT	2460
	GTGTTATCAA	ATACCTACGT	ATTTATCAGG	TAAGCCGACA	TTAATTATCT	CACCGTTAAT	2520
	ATCTTTAATG	GATGACCAAG	TTATGCAGTT	GAAAATAAAT	GGAGAAAAAC	GTGTAACATG	2580
25	TATTCACTCT	GGTATGGATG	AAATTGAGAA	AAAGCATAAT	ATTAAATGTT	TACGACATAG	2640
	CCGCTTCATC	TTTCTAAGTC	CAGAATTTCT	CCTGCAACCG	TCAAATTTTA	AATTAATATC	2700
30	TATGATAGAC	TTTGGCATGA	TTGTTCTAGA	TGAAGCACAT	TGCCTATCTG	AATGGGGATA	2760
	TGATTTCAGA	CCACATTATG	CTCTAATAGG	AAAAGTAACA	AAGCATTTTA	AAGAAGCGGT	2820
	TGTCTTAGCA	TTGACAGCAA	CTGCACCACC	GCATTTACAA	GATGATTTGA	CGGAAATGTT	2880
35	AGCGATTCAA	TTCAATGTTA	TTAAAACCTAC	AATGAATCGC	CCAAATATAA	GCTTTAAGCA	2940
	TCTTAATTTT	CATGATGATG	AAGATAAAAT	TGAATGGTTG	CTGCCGTTTC	TACAACAGTC	3000
	GGGAGCAACG	ATTATTTATG	TCTCATCGAA	AAAGATGTGT	CTGAATTTAG	CGCAACTTAT	3060
40	TTATGATTCA	GGTTTTCTTA	CAGGTATTTA	TCATGGTGAT	ATGAATTATC	AAGAGCGACA	3120
	CACAGTTCAA	CAACAATTTT	TAAATAATGA	TATTCCGATT	ATAGTCGCAA	CGAGTGCTTT	3180
	TGGTATGGGA	ATTAATAAAA	AAGATATTCG	CACAATCATT	CACTTTTCATC	TTTCAACAAG	3240
45	TCCTTCTAAC	TACATTCAAG	AAATTGGCCG	TGCGGGTCGC	GATGGTGAAC	TAAGTCAGGC	3300
	AATTAGTTTA	TTCCAACCGG	ACGATAAATA	TATTTTAGAA	ACGTTATTAT	TTGCAGATAT	3360
50	GATAACAGAA	GAAGATGTAC	AAAATTTTCGA	AATAGGAGAA	TTTTTAGCTC	CCGATAAACA	3420
	AGCCGTTTTG	ACAACGTTGc	AATCATTCTA	TAGTATCGGC	GCCTTGaAAC	AGATATTTAA	3480
55	GCAATCATTT	AAACGAAAGC	AATTAGGATT	CTTTCGCATG	ATTGGCTATT	GCAAATTGGA	3540

	ATGTTGTGAC	AATGATTCTA	ATATAACTGA	TATCGCAATT	TTAAATAAGA	AGAAGGTAAT	3660
	TAGAAGTATT	GGATTTGATG	AAAAGTTGCA	AAATTTATTT	CTCAGATAGT	ATTACTTTAC	3720
5	TAAAAGAAAA	TTGACAAGCT	ATAATTAGTG	TATACACAAT	TGAAAAATGA	TTGAAATAAT	3780
	TTTGAAAAAT	ATACATAAAC	ATATGTCATG	TGGGTATATT	TTATGTAAAA	TCATTGTAAT	3840
10	AGAATAGAAA	GGAAGATGGC	TATGTCTAAT	AATTTTAAAG	ATGACTTTGA	AAAAAATCGT	3900
	CAATCGATAG	ACACAAATTC	ACATCAAGAC	CATACGGAAG	ATGTTGAAAA	AGACCAATCA	3960
	GAATTAGAAC	ATCAGGATAC	AATAGAGAAT	ACGGAGCAAC	AGTTTCCGCC	AAGAAATGCC	4020
15	CAAAGAAGAA	AAAGACGCCG	TGATTTAGCA	ACGAATCATA	ATAAACAAAGT	TCACAATGAA	4080
	TCACAAACAT	CTGAAGACAA	TGTTCAAAAT	GAGGCTGGCA	CAATAGATGA	TCGTCAAGTC	4140
	GAATCATCAC	ACAGTACTGA	AaGTCAAGAA	CCTAGCCATC	AAGACAGTAC	ACCTCAACAT	4200
20	GAAGAGGAAT	ATTATAATAA	GAATGCTTTT	GCAATGGATA	AATCACATCC	AGAACCAATC	4260
	GAAGACAATG	ATAAACACGA	TACTATTAAA	AATGCAGAAA	ATAACACTGA	GCATTCAACA	4320
	GTTTCTGATA	AGAGTGAAGC	TGAACAATCT	CAGCAACCTA	AACCATATTT	TACAACAGGT	4380
25	GCTAACCAAT	CAGAAACATC	AAAAAATGAA	CATGATAATG	ATTCTGTAAA	ACAAGATCAA	4440
	GATGAACCTA	AAGAACATCA	TAATGGTAAA	AAAGCAGCAG	CTATTGGTGC	TGGAACAGCA	4500
	GGTGTTCGAG	GTGCGAGCTG	TGCAATGGCT	GCTTCTAAAG	CTAAGAAACA	TTCAAATGAC	4560
30	GCTCAAAACA	AAAGTAATTC	TGGCAAGGCG	AATAACTCGA	CTGAGGATAA	AGCGTCTCAA	4620
	GATAAGTCTA	AAGATCATCA	TAATGGCAAA	AAAGGTGCAG	CGATCGGTGC	TGGAACAGCA	4680
35	GGTTTGGCTG	GAGGCGCagC	AAGTAAAAGT	GCTTCTGCCG	CTTCAAAACC	ACATGCCTCT	4740
	AATAATGCAA	GCCAAAACCA	TGATGAACAT	GACAATCATG	ACAGAGATAA	AGAACGTAAA	4800
	AAAGGTGGCA	TGGCCAAAGT	ATTGTTACCA	TTAATTGCAG	CTGTACTAAT	TATCGGTGCA	4860
40	TTAGCGATAT	TTGGAGGCAT	GGCATTAAAC	AATCATAATA	ATGGTACAAA	AGAAAATAAA	4920
	ATCGCGAATA	CAAATAAAAA	TAATGCTGAT	GAAAGTAAAG	ACAAAGACAC	ATCTAAAGAC	4980
	GCTTCTAAAG	ATAAATCAAA	ATCTACAGAC	AGTGATAAAT	CAAAAGAGGA	TCAAGACAAA	5040
45	GCGACTAAAG	ATGAATCTGA	TAATGATCAA	AACAACGCTA	ATCAAGCGAA	CAATCAAGCA	5100
	CAAAATAATC	AAAATCAACA	ACAAGCTAAT	CAAAATCAAC	AACAGCAACA	ACAACGTCAA	5160
	GGTGGTGGCC	AAAGACATAC	AGTGAATGGT	CAAGAAAAC	TATACCGTAT	CGCAATTCAA	5220
50	TACTACGGTT	CAGGTTCCAC	GGAAAATGTT	GAAAAAATTA	GACGTGCCAA	TGGTTTAAGT	5280
	GGTAACAATA	TTAGAAACGG	TCAACAAATC	GTTATTCCAT	AATATAACTA	TATAAATTGT	5340

55

TAAATTGCGC TTATAAGTAT GTAGCGGTTT TTTCATTTTT CAAAGTTTGT TATTTAACAA 5460
 GGTCTTGTCT CGAATATTGG CATATCAATT TAACTTTTAA AATAGTCATC AAAAAGATAA 5520
 5 AACACCACAA TCAACAAATT TAACGAGGAA GAATAAAAAA TAAATCAACA TATTTAAATTG 5580
 TAGTGTTATT CAACTCCGTA GCTAACAATT CTCTATTAC ATTAAACAAA TTGTCAAAAA 5640
 TATATCATAA ATCTTCAAGC ACAGACTTAG CGCATCAATC ACTGAACTGT TATAATAGTT 5700
 10 TGGGATTAAA GGAGGCCGAA ACAATGCAAA AAGTTGAAAG TATCATAATT GGTGGAGGGC 5760
 CATGCGGATT AAGTGGCGCT ATTGAACAAA AAAGAAAAGG TATTGATACC TTAATTATTG 5820
 AAAAGGGTAA TGTCGTTGAA TCAATCTACA ATTATCTAC TCACCAAACA TTTTCTCAT 5880
 15 CAAGTGATAA ATTAAGTATT GGGcGagTAC CGTTTATCGT TGAAGAAAGT AAACCAAGAC 5940
 GTAATCAAGC GCTAGTTTAT TACCGAGAAG TTGTAAAACA TCATCAATTA AAAGTAAATG 6000
 20 CATTTGAAGA AGTATTAACT GTTAAAAAAA TGAATAATAA ATTTACTATT ACTACGACGA 6060
 AAGATGTTTA TGAATGTCGA TTTTAAACAA TCGCGACAGG CTATTATGGT CAGCATAATA 6120
 CATTAGAAGT TGAAGGTGCG GATTACCTA AAGTGTTCCA TTATTTTAAA GAGGCACATC 6180
 25 CGTATTTTGA TCAAGATGTT GTAATTATCG GTGGTAAGAA TTCGGCTATC GATGCTGCTT 6240
 TGGAGTTGGA AAAAGCTGGT GCTAACGTGA CGTTCTATA TCGTGGTGA GATTATTCGC 6300
 CTTCAATTAA ACCGTGGATA CTTCCAAATT TCACAGCATT AGTAAATCAT GAAAAAATTG 6360
 30 ACATGGAATT TAATGCTAAT GTTACCCAAA TAACTGAAGA TACTGTGACT TATGAAGTAA 6420
 ATGGTGAAAG TAAAACGATA CACAATGATT ATGTATTTGC GATGATTGGT TATCATCCCG 6480
 ATTATGAATT TTTAAATCT GTAGGCATTC AAATTAATAC AAATGAATTT GGAACAGCGC 6540
 35 CTATGTATAA TAAAGAAACA TACGAAACAA ATATCGAAAA TTGCTATATT GCAGGTGTAA 6600
 TTGCTGCAGG GAACGATGCG AATACCATT TTATTGAAAA TGGTAAATTC CACGGGGGCA 6660
 40 TTATTGCTCA AAGCATGCTA GCTAAGAAAC AAACGCCCTT AGAATCATAA AAATAAAGGT 6720
 CTATGTACTA AATAACTTAG TTTTACAACG ACTGACATTC ATGATATGTC AAATGAGGTT 6780
 GATGACTATT GATTGTACAT AGACCTTTTT ATGTTACGTA TTCATTATAA TTCAAAATAT 6840
 45 GATTTGATTT CAGCTTTATC TAAATTGTTG CTTAACGCGA CTAATAATTT TAATCTTGCT 6900
 TTTGGACCAT TCAAGCCGTT AGAAAAAATA AAACCTTGTT GTGCGAGTTG GTAACCACCA 6960
 CCATCGTATG CGTAAGTTGG ACTCACAATA CCATTAAAGG AACGTGAAAC TAGCACAATA 7020
 50 GGTATATTTA AAGATACTAA TTGTTGAATG CCTTCTAATG CGCTTGAGG TATGTTGCCT 7080
 TGTCTAACG CTTCAATAAC CATACCATCG ATACCTTCTC GACTATAAAA ACTAAAAATG 7140

55

TCCAATGCTT GTTGGCGATA TGGCATATGA TGGAAATTGCA CACGATCCTT TGTCAATACA 7260
 CCTAGCGGAC CATGATTTGG ACTTTGAAAT GTGTTTGTAT TAGACGTATG TGTTTTGGTA 7320
 5 ACATTACGCG CCGTATGAAT TTCATCATTa AaTACAACCA TCACGCCTTT ATGACGGGCC 7380
 TTTTCATCAG AGGCAACGCG AATAGCGGAA ATATAATTAT ATAATCCGTC AGAACCAATT 7440
 TCATTAGACG AGCGCATTGG CCAAGAAnta ACAACAGGTT G 7481

10 (2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6346 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

ATGAATTATA CCTTCCACTA GAAGTGTGCG TATTTATGTG CCTGGTGGTA AAGCAAGTTA 60
 TCCTTCAACG GTATTAATGA CAGCGACGTT AGCACAAGTA GCAGGTGTGG AAAATATTGT 120
 25 CGTTGTGACA CCACCTCAAC CTAATGGTGT GTCTCAAGAA GTATTAGCAG CATGCTATAT 180
 TACGCAAGTT GATCAAGTAT TTCAAGTTGG TGGTGCTCAA AGTATTGCTG CGTTAACTTA 240
 TGGCACAGAA ACGATACCTA AAGTTGATAA GATTGTAGGT CCAGGCAATC AATTTGTTGC 300
 30 ATATGCCAAG AAATATTTAT TTGGACAAGT AGGTATTGAC CAAATAGCTG GACCAACAGA 360
 AATAGCACTT ATTATTGATG AAACAGCAGA TTTAGATGCC ATCGTATATG ATGTTTTTGC 420
 ACAAGCAGAA CATGATGAAT TAGCACGTAC ATATGCCATT AGTGAAGATG CGCAAGTCCT 480
 35 TAAAGATTTA GAATCACGTA TTGCTAAAGC ATTGCCTAAT GTGGACAGAT ACGACATTGT 540
 TTCTAAAAGT ATCGCTAATC AACACTACCT TATCCATGCT AGTAATTTTG ATGAAGCATG 600
 CCATGTCATG AATACAATCG CGCCTGAACA TGGCTCGATT CAAACAGTAA ATCCTCAACC 660
 40 ATATATTGAA AAAGTGAAAT ATGTGGGTGC ATTGTTTATT GGACATTATT CGCCAGAGGT 720
 CATAGGAGAT TACGTTGCAG GTCCAAGTCA TGTATTACCT ACAAATAGAA CAGCTAGATT 780
 45 TACCAATGGG TTATCGGTCA ATGATTTCTT AACACGGAAC ACGGTCATCC ATTTATCAAA 840
 AGATACGTTT GAACAAATTG CTGATTACAGC ACAACATATT GTCATGTTG AAGCATTATA 900
 CAATCACCAG CAGTCTATTT TAATACGTCA GTCTTAGGGG AGTGTAATTG AAATGATTTA 960
 50 TATTGATAAA AATGAAAGTC CAGTTACGCC GTTGGATGAA AAAACAATGA CGTCTATTAT 1020
 TAGTGCAACg CnATATAATT TATATCCTGA TGCAGCATAT GAACAATTCA AGGAAGCTTA 1080

	GATTCAAAAG TTAATGCTGA TCATGCCAGA AGGTCCGGCA TTAACGCTAA ATCCTGATTT	1200
	TTTTATGTAT CAAGCATATG CGGCACAAGT AAATCGTGAA ATTGCATTG TAGATGCAGG	1260
5	ATCAGATTTA ACGTTTGATT TGGAAACCAT TTAAACGAAA ATCGATGAAG TACAACCATC	1320
	ATTTTTTATT ATGAGTAATC CACATAACCC TTCAGGCAAG CAATTTGATA CGGCATTTTT	1380
	AACAGCTATT GCAGATAAGA TGAAAGCATT AAACGGATAC TTTGTCATTG ATGAAGCATA	1440
10	TTTAGATTAT GGTACGGCAT ATGACGTGGA ACTGGCACCA CACATCTTAA GAATGCGTAC	1500
	ATTATCAAAG GCGTTTGGA TGGCCGGCTT AAGATTAGGT GTCTTAATTA GTACTGCTGG	1560
	AACGATAAAG CATATTCAAA AAATAGAACA TCCATATCCA TTAAATGTAT TTACGCTAAA	1620
15	TATTGCGACT TATATTTTTA GACATAGAGA AGAGACAAGA CAATTTTTAA CGATGCAACG	1680
	ACAGTTAGCT GAGCAGTTAA AACAAATATT TGATACACAT GTTGCGAGATA AAATGTCAGT	1740
20	GTTCCCATCA AATGCTAATT TTGTACTTAC TAAAGGCTCA GCAGCGCAAC AATTAGGACA	1800
	ATACGTATAT GAACAAGGAT TTAAACCTCG CTTTTATGAT GAGCCGGTGA TGAAGGGCTA	1860
	TGTAAGATAC TCAATTGCAA CAGCATCACA GTTAAAGCAA TTAGAAGAAA TTGTTAAAGA	1920
25	ATGGAGTGCA AAATATGATT TATCAAAAAC AACGAAACAC AGCTGAAACG CAACTAAATA	1980
	TTTCAATATC TGATGATCAG TCACCATCGC ATATTAATAC AGGTGTGGGC TTTTAAATC	2040
	ATATGTTAAC CTTGTTTACA TTTCATAGCG GTCTGTCTATT AAACATTGAG GCACAAGGTG	2100
30	ATATTGACGT AGATGATCAC CACGTAACGT AAGATATCGG CATTGTCATT GGCCAATTGT	2160
	TACTTGAAAT GATTAAAGAT AAAAAGCATT TCGTTCTGTTA TGGAACGATG TACATTCCAA	2220
	TGGATGAAAC ATTAGCACGT GTCGTTGTGG ATATAAGTGG GCGCCCATAC CTATCATTCA	2280
35	ATGCATCATT AAGTAAAGAA AAAGTTGGTA CGTTTGATAC GGAGTTAGTA GAAGAATTTT	2340
	TTAGAGCGGT CGTAATCAAT GCAAGATTAA CAACGCATAT TGATTTAATT CGTGGAGGTA	2400
40	ATACACACCA TGAAATTGAA GCTATATTCA AAGCGTTTTC CCGTGCATTA GGCATAGCGC	2460
	TAAGTGCAAC TGATGATCAG CGTGTGCCGT CATCGAAAGG TGTGATTGAA TGATTGTCAT	2520
	CGTTGATTAT GGATTAGGGA ATATTAGTAA TGTAACACGC GCTATTGAAC ATTTAGGGTA	2580
45	TGAGGTGGTT GTCTCAAATA CCTCAAAAAT AATCGATCAA GCAGAAACAA TCATATTGCC	2640
	CGGTGTCGGC CATTTTAAAG ATGCGATGTC AGAGATAAAA CGATTAAATC TCAATGCAAT	2700
	ATTGGCTAAG AATACTGATA AGAAGATGAT TGGTATTTGT TTAGGCATGC AATTAATGTA	2760
50	TGAGCATAGT GATGAAGGCG ATGCATCTGG ATTAGGGTTT ATCCCAGGAA ATATTTCCGC	2820
	TATCCAAACA GAATACCCAG TGCCACACTT AGGCTGGAAT AATTTAGTGA GTAAGCACCC	2880

55

	AATTGCATAT GCGCAGTATG GGGCTGATAT TCCGGCAATT GTTCAATTTA ACAATTATAT	3000
	TGGTATTCAA TTCCATCCTG AAAAAAGCGG TACATATGGG TTACAAATTT TCGCTCAGGC	3060
5	AATACAAGGG GGATTTATAA ATGATTGAAT TATGGCCAGC GATTGATTTG ATTGGGTCAA	3120
	CAAGTGTGAG GTTAACAGAG GGTAAATATG ATAGTGAAGA AAAAATGTCA CGCTCGGCTG	3180
	AAGAAAGTAT TGCTTACTAT AGTCAATTG AATGTGTGAA TCGTATTCAT ATCGTCGACT	3240
10	TGATAGGTGC TAAGGCACAG CATGCCCGAG AGTTTGATTA TATTAAGTCA TTGAGGAGAT	3300
	TAACAACCAA AGATATTGAA GTAGGTGGTG GCATTCTGAC GAAGTCACAA ATCATGGACT	3360
15	ACTTTGCCGC AGGGATTAAT TATTGCATAG TTGGAACGAA AGGTATTCAA GATACTGATT	3420
	GGCTTAAAGA GATGGCACAT ACATTTCCAG GTCGCATTTA TTTATCTGTT GATGCCTATG	3480
	GAGAAGATAT TAAAGTGAAC GGATGGGAAG AGGACACAGA GTTAAATTTA TTTAGTTTTG	3540
20	TGAGACGGTT ATCGGATATA CCTCTTGGCG GCATTATATA TACTGATATT GCTAAAGATG	3600
	GCAAAATGTC CGGACCTAAC TTTGAATTAA CTGGTCAATT AGTAAAGGCA ACGACGATTC	3660
	CTGTCAATTGC TTCCGGTGGT ATTAGACATC AGCAAGATAT TCAACGATTA GCGTCGCTAA	3720
25	ATGTTACGC TGCTATTATA GGAAAGGCTG CACATCAAGC ATCTTTTGG GAGGGGCTAA	3780
	AATGATTAA AAACGTATCA TTCCATGTTT AGATGTCAA GATGGTCGTG TCGTTAAAGG	3840
	GATTCAATTT AAAGGATTAA GGGATATTGG GAATCCTGTT GATTTAGCAA TGTATTACAA	3900
30	TGAAGCGGTT GCTGATGAAT TAGTATTTT AGACATCTCT AAGACGGAAG AGGGTCATAG	3960
	CTTAATGCTA GAAGTGATTG AACAGACAGC GTCACGCTTG TTTATCCCTC TTACTGTAGG	4020
	GGGTGGGATT CAAAGTCTCG ATGATATTAC CCAATTGCTA AATCATGGTG CAGATAAAGT	4080
35	ATCATTAAAT TCAAGTGCTT TAAAAATCC ACAGCTCATT AAACAAGCGA GTGATAAATT	4140
	CGGTAGACAA TGCATCTGCA TAGCAATTGA TAGCTATTAT GATCCTGAAA GAAAAGCACA	4200
	TTATTGTTGT ACGACTGGTG GTAAAAAAT GACAAATATT AAAGTATATG ACTGGGTACA	4260
40	GCAAGTAGAA CAGTTAGGTG CAGGTGAGCT CCTCGTTACA AGTATGGGAC ATGATGGTAT	4320
	GAAACAAGGC TTTGATATTG AACACCTAGC AAATATTAAG TCTCTGTAA ATATTCCAAT	4380
45	CATTGCTTCT GGTGGTGGTG GCAATGCACA ACACTTTGTA GAATTATTTG ATCAGACGGA	4440
	TGTTTCTGCA GGTTTAGCTG CAAGTATATT ACATGATCGA GAAACGACGG TTCAATCTAT	4500
	TAAAGAAGTG ATACGGCAAG GGGGTATAGC AGTAAGATGA CCAAATATAA AATTGATTTT	4560
50	AGCAAAGGTT TAGTGCCAGC AATTTTACAA GATAATCAAA CAAAACAAGT ATTGATGTTG	4620
	GGTTATATGA ACCAAGAAGC TTTTGATAAA ACGATAGAAG ATGGTGTGGT ATGTTTCTAT	4680

55

AAAGATATTC ATGTAGATTG CGACAATGAC ACTATTTTAA TTGATGTCAT ACCAAATGGA 4800
 CCAACATGTC ATACAGGCAG TCAAAGTTGT TTCAACACAG AAGTTCCATT TTCAGTGCAA 4860
 5 ACATTAGCGC AGACAGTTCA AGATAGTGCC CAATCCAATA ATGAAAAGTC ATATACAAAA 4920
 TATTTATTAA CAGAAGGTAT AGAAAAGATT ACAAAAAAAT ACGGTGAAGA AGCTTTTGAA 4980
 GTCGTAATTG AAGCAATTAA AGGTGACAAA AAAGCATTGT TAAGTGAAGT AGCAGATGAA 5040
 10 CTTTATCATT TATTTGTCTT GATGCATGCG CTTGGCGTCG ATTTTTCAGA AATTGAGGCG 5100
 GAATTAGCGC GTAGACATCA TAAGCGCAAT AACTTTAAAG GTGAACGACA AAATATCGAA 5160
 CAGTGGTAAA GCAAGTATGG ACTAAGATAT AAGGAAAAGG ATCATGGCTT ATACACTTAC 5220
 15 AAATATTGTG GAAAACGTGA CATTTTCAAG TTTAAAATAC GACACCAACA TATTTTAACT 5280
 ATGAATGCTG TGATGGTACT AAAGTTGCGA ACTCGTTATA GATAAGTAGT GGATAATCAC 5340
 20 AATACGAAAT CAAAATAAT TATAAAAAGT AAATTGAGCA ACTCAGGAAT AGATGTCACT 5400
 GTTAAAGATG TCGAAAAGTA TATGAATCGA TATAATGAAG TTATGAAGGG AAAAAATGGC 5460
 GAAAAAGCTA AAGAGTTATG TTTGTCGTTA CTACCTATTA ATATCATAGT TGTCTTTACA 5520
 25 TTCITTGTAT TTATACTATA AATACAAATA TATCTAGCCT GAAATAGAAA TGTCATAGCC 5580
 TATTTAAAAG ACAATCTCCA TTAGAATAA GATATGCATC CCGAAAGTTA GACTAAAAAA 5640
 CTAACTTTAT GGGATGTATT TTTATGCTAA TCATCATAAA TTCGAGATTA AGTTAAAGGT 5700
 30 AGTTCAAGAG TAATTAAACA ATAAATWAAA AATAGTAGGA TACTTACTTT GAGGGAAGAA 5760
 AATTAACTGT ATATATTTAG TTTAGGAACA AGTATTACGG TTTATCCTGA TACAATTATT 5820
 GTGGATGGGA TGATATTTTT AGGTTTAAAA TACGACACCA GCAAACATAA TAACTGTAAT 5880
 35 AGCTCATAAA TCTCCCCATA TAGCTAATCT AAAAAAATAA TACATCATTG GAATTAAGCC 5940
 CCAAGCATGT AAATATTAAA AATCAAAAma GATATmTGTA AAAaAGTTAC AATTtGCATA 6000
 ATTAAATTGT GTCTAATTAT TGACTAATTA AATTTTGCCA AATATAATAT TAATTAATAA 6060
 40 TTTGaAATGA TTAGCGTATA CACTTTAAAT TCTCTTTGGA GAATATATTT TTTAAATACA 6120
 AATGTAAACG CTTTCTCGTC AAATTAAACA ATAGAAAGGA TGGTCATTAT GAGTGCTTGG 6180
 TTAAGTAAAT TATTTGAGTT TATTCCTCGA ATAATTATCA ATTTGTTTAT CTAAAATAAA 6240
 45 AAAATAGAGG TGCTGACAAT GATGAAAAGT CAAAATAAGT ATAGTATTCG TAAATTTAGT 6300
 GTAGGTGCAT CTTCCAATTTT AATAGCTACA TTAATTTTTT TAAGTG 6346

50 (2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3775 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:

	TGATAATGAT TCTCATTGTC ATACATCAGC AAGGAGGCTA ATTAGTCAAT GAATAAAGTA	60
10	ATTAAAATGC TTGTTGTTAC GCTTGCTTTC CTACTTGTTT TAGCAGGATG TAGTGGGAAT	120
	TCAAATAAAC AATCATCTGA TAACAAAGAT AAGGAAACAA CTTCAATTAA ACATGCAATG	180
	GGtACAACCTG AAATTAAAGG GAAACCAAAG CGTGTTGTTA CGCTATATCA AGGTGCCACT	240
15	GACGTCGCTG TATCTTTAGG TGTAAACCT GTAGGTGCTG TAGAATCATG GACACAAAAA	300
	CCGAAATTCG AATACATAAA AATGATTTA AAAGATACTA AGATTGTAGG TCAAGAACCT	360
	GCACCTAACT TAGAGGAAAT CTCTAAATTA AAACCGGACT TAATTGTCGC GTCAAAAGTT	420
20	AGAAATGAAA AAGTTTACGA TCAATTATCT AAAATCGCAC CAACAGTTTC TACTGATACA	480
	GTTTTCAAAT TCAAAGATAC AACTAAGTTA ATGGGGAAAG CTTTAGGGAA AGAAAAAGAA	540
	GCTGAAGATT TACTTAAAAA GTACGATGAT AAAGTAGCTG CATTCCAAAA AGATGCAAAA	600
25	GCAAAGTATA AAGATGCATG GCCATTGAAA GCTTCAGTTG TTAACCTCCG TGCTGATCAT	660
	ACAAGAATTT ATGCTGGTGG ATATGCTGGT GAAATCTTAA ATGATTTAGG ATTCAAACGT	720
	AATAAGACT TACAAAAACA AGTTGATAAT GGTAAAGATA TTATCCAAC TACATCTAAA	780
30	GAAAGCATT CATTAAATGAA CGCTGATCAT ATTTTTGTAG TAAAATCAGA TCCAAATGCG	840
	AAAGATGCTG CATTAGTTAA AAAGACTGAA AGCGAATGGA CTTCAAGTAA AGAGTGGAAG	900
35	AATTTAGACG CaGTTAAAAA CAACCAAGTA TCTGATGATT TAGATGAAAT CACTTGGAAC	960
	TTAGCTGGCG GATATAAATC TTCATTAAAA CTTATTGACG ATTTATATGA AAAGTTAAAT	1020
	ATTGAAAAAC AATCAAAATA ATTAAGGAGT TTTACGATGC TACTTAAACC AAAATACCAA	1080
40	ATCGTTATTG CTGGTTTATG TCTTGCAATA GTAGCTATCT TAAGTTTAAAT GATTGGAAAT	1140
	ACGCTTGTGT CACCAGGTAC GGTGATACAG GCGTTATTCA ACTTTGATAG TGAAAACGAT	1200
	TTACATGATG TTGTCACTGG TGCACGGGCG TCGAGAACAA TCATTGCGTT ATTGACTGGT	1260
45	GCTGCCCTTG CTGTCTCAGG TTTGTTGATG CAAGCACTTA CACGAAACCC AATAGCCTCA	1320
	CCAGGGCTTT TCGGTGTCAA TGCAGGCGCA GTATTTTTTG TCATTTTTAG TATTACATTT	1380
	ATCCAAATTC AATCTTTTAA AATGATTGTA GTTATTGCAT TTTTGGGGGC TATTGTTGTT	1440
50	ACTGTATTAG TTGTTGCACT AGGTATGTTT AGACAAACAC TATTCTCACC TCACCGTGTC	1500
	ATTTTGCGAG GTGCTGCGAT TGCATGCTA TTTACAGCCT TTAACAAGG CATACTTATT	1560

55

	AATATTTGGG ATATCCCATG GATTATTCCG CTTGTATTGA TACTTATTTT AATTGCATTT	1680
	AGCATGGCTG CACACATCAA CATCTTGATG ACAAGTGACG ACATTGCAAC CGGCCTCGGT	1740
5	CAAAACATAA AATTAAATCAA ATGGATGATT ATTATGCTCA TCAGTATGTT AGCCGGTATT	1800
	TCGGTAGCCG TAGCTGGATC AATCGTCTTT GTGGGTCTTA TCGTACCGAA TATTAGCAAA	1860
	CGATTATTAC CACCAAATA TAAGTATTTA ATTCTTTTA CTGCATTAGC TGGAGCAATC	1920
10	CTAATGATCA TTTCAGACAT TGTTGCTCGT ATAATAATTA AGCCACTAGA GTTGCCTATC	1980
	GGTGTCTGTTA CCGCTGTCTT TGGCGCTATT GTCTTAATCT ATATTATGAA GAAAGGACGT	2040
	CAACGCTTAT GACCGAAAAG ATTAATAAAA AAGACAATTA CCATCTCATC TTCGCGTTAA	2100
15	TCTTTTTAGC CATCGTTTCA GTGGTAAGTA TGATGATTGG TTCAAGCTTT ATACCATTAC	2160
	AACGCGTACT GATGTACTTT ATAAATCCAA ATGACAGTAT GGATCAATTC ACTTTAGAAG	2220
20	TATTACGCTT ACCTCGCATT ACACTTGCGA TTTTAGCAGG TGCCGCACTA GGAATGAGTG	2280
	GTTTAATGTT GCAAAATGTA TTAAAAATC CAATTGCCTC ACCTGATATT ATCGGTATCA	2340
	CAGGTGGTGC TAGCTTAAGT GCTGTTGTCT TTATTGCATT TTTCAGCCAT TTAACAATAC	2400
25	ATTTACTTCC ACTATTTGCA GTATTAGGTG GCGCAGTTGC AATGATGATA CTATTAGTGT	2460
	TTCAAACGAA AGGACAAATA CGCCCGACAA CACTCATAAT CATCGGTATT TCGATGCAAA	2520
	CGTTGTTTAT TGCCTTGTC CAAGGATTAC TCATTACAAC GAAGCAATTA TCTGCTGCCA	2580
30	AAGCTTATAC ATGGCTAGTC GGAAGTCTTT ACGGTGCTAC GTTTAAAGAT ACAATCATTT	2640
	TGGGTATGGT TATTTTAGCT GTTGTGCGGT TGTATTCTT TGTATACCA AAAATGAAAA	2700
	TATCTATACT TGATGACCCT GTAGCGATTG GCTTAGGCTT ACATGTACAA CGTATGAAAC	2760
35	TAATCCAATT AATCACTTCT ACTATACTCG TATCTATGGC AATCAGTTTA GTAGGTAACA	2820
	TTGGSTTTGT CGGTTTAATC GCACCACATA TCGCGAAAAC AATCGTTCGC GGAAGTTATG	2880
	CTAAAAAGTT ACTAATGTCA GCAATGATTG GTGCCATATC AATTGTTATT GCAGACTTAA	2940
40	TTGGGCGTAC CTTATTCTTG CCTAAAGAAG TGCCAGCAGG TGTATTTATT GCTGCTTTTG	3000
	GTGCCCCATT CTTATATAC TTATTATTAA CCGTGAAAAA GTTATAACGA TATTATTAA	3060
45	ACAAATGAC CTCACAACGA AGTTAGCTAA ATGATTCACT TAACTAACCG TTGCGAGGTT	3120
	TTTTTATACA TATAGTTGTT GTTATTGTTA ACAAGCGTCG ACTTTCTTAA TTACATATTA	3180
	ATACTTTATA TACAAATAAC ACCGACTCAT ATTCTATAAT ATCAATCAAT ATTCTTCGAT	3240
50	TTTTCAAATA TCGATAACTA TTTCTTATTT AAATATAGTG TTTGATAATG TCATTTATTC	3300
	AAAAACACAA ATTTTAATAA AAATATCATA TTATTTTTAA TTGTAAATTA TGGATTATTT	3360

ATATAATAAC CATGACAAAA TTAGAGATTT TATAATCATT GAAGCATATA TGCTCGTTT 3480
 TAAGAAAAAA GTCAAGCCTG AAGTCGATAT GACTATAAAA GAATTTATAT TACTGACTTA 3540
 5 TTTATTTTCAT CAGCAAGAAA ACACACTTCC ATTTAAGAAG ATTGTTTCAG ATTTATGTTA 3600
 TAAACAATCG GATTTAGTAC AGCATATAAA AGTACTTGTG AAACATTCAT ATATTAGTAA 3660
 AGTTCGAAGT AAAATTGATG AGCGTAATAC TTACATTTCa ATATCTGAAG AACAACTAGA 3720
 10 nAAAATTGCA GAACnTGTTA CATTGTTTGA TCAAATCATT AAACAATTTA ACCTT 3775

(2) INFORMATION FOR SEQ ID NO: 239:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1361 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

AGAAAAATTA GCCTACCTAT GCAAGTTGTT ATTGCATTAG TGTTAGGTGT TGTCGTAGGA 60
 25 CTTTTATTAT ATGGACAAGA AAACGTAGCA AATTACATTA AACCATTGGG TGACGTATTT 120
 TTAAATTTAA TTAAATGAT CGTTATACCA GTTGTATTTT GCTCACTAGC GCTTCTATT 180
 TCGAACGTTG GGAATCGAA AACTGTAGGG CGTTATGGCT GGaAAACAAT TTTATACTTT 240
 30 GAAATTATTA CAACAATCGC AATAGGTTTA GGGATTATCT TCGGTAACCT ATTTAAACCA 300
 GGTGCTGGAT TAGACCCAAC AAAATTACCT AAAGGTGATA TTTCTAAATA TCAATCAACT 360
 GCACATGCAG CAGAACAATC TACATATGGA AATCATTTTA TTGATACCAT TGTACATATT 420
 35 ATTCCGACAA ACTTTTTTGA AGCTTTAAAT AAGGGTGAAT TATTACCTAT TATCTTCTTC 480
 GCAGTATTCT TTGGATTAGG ATTAGCTGCT GTAGGTAAAA AAGCAGAACC AGTTAAAGAA 540
 TTTTAAAGCG GATCGCTTGA AGCTGTGTTT TGGATGATTA ATAAAATTTT AAAATTAGCA 600
 40 CCACTTGGAG TGTTTGCAAT CATTTGTAAT ACAATTATTA CATTTGGTGC ATCCGCATTA 660
 TTACCACTAT TAAAATTAGT ATTAGTTGTT GTCTTTGCAA TGGTGTTCCT TGTATTGCTT 720
 ATACTAGGAC TAGTTGCATG GATGTGTGGT ATTAATATCA TGAATATTAT TAGAATCTTG 780
 45 AAAAGTGAAT TGCTTTTAGC ATTTTCTACA TCAAGTTCGG AAGCTGTACT TCCTGTAATG 840
 ATGAAGAAAA TGGAAACTT CGGTTCTCCA AAAGAAATTA CTTCTTTTGT TATACCAATT 900
 50 GGTATACGT TTAACCTAGA TGGATCAGCA CTTTATCAAT CTATTGCAGC ATTATTCGTT 960
 GCACAGATGT ATGGAATGCA CTTAACATTA TCAGAGCAAA TTGTGTTGAT GTTAACATTA 1020

55

ACATTAGGTG CCATGGGCTT ACCGGCACAA GGTtTAGCAT TAATTATTGG TGTGACCGT 1140
 ATCTTAGATA TGGTACGTAC ATGTGTAAAC GTTATTGGTA ATGCATTATC AACAAATCGTT 1200
 5 ATAGCTAAAT GGGAAAACGT ATATGACAAA GCAAAGGTC AAGAATATTT AAAATCAATT 1260
 TAAAAAATAC TATCTGACAT TTAArGnCCC TTACAACCTT TGGTTgTnAG GGCTnTTTTA 1320
 TGTCATGCGT CTTAAAGCCA GGCCGTATAn CGSTAAGCGT A 1361

10 (2) INFORMATION FOR SEQ ID NO: 240:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1489 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

TCAATATGTA AATCGGCGAT TGTCCCTATC tTCATACCCT TACCTTCTTT ACTCAAATTA 60
 AAATGTTATT TATTATGCCA TaGCTTATCT AATATATATA GTTAACTGCT TCATTTAGAT 120
 25 GATTATTTTA TATTTTTCGA TAAAACTTA TATCTTTTCA AAAAAATCGA TAAGTTTTAG 180
 TTATCATACC CTTACCTATC AAATGTTTTT TCTTATATTT AAAAAAATAA TTGCTTTATT 240
 AAATGGATTT CTTTAGTATT TATAATTAAG AAAACGCTTA CACACAACTT TTTTATTTGC 300
 30 TTTATCCTGA GGAGGAAAAT TATGGCAAGA AAATTGCATA GAGAGTTGAA TAACAGACAC 360
 ATCCAATTAA TAGCAATTGG GGGCGCAATT GGAAGTGGT TATTCCTAGG ATCAGGTCAA 420
 ACAATATCTT TAACTGGTCC ATCACTGTTA TTCACATACA TGATTATTGG GGTGTACTA 480
 35 TTCGCTTTTA TCGCGCATT AGGCGAATTG TTGTTGAGCA ATACAAGATT TAATTCATT 540
 GTTGATATTG CAAATGAATA TTAGGCCCT TTTGGTGGCT TTGTCATTGG CTGGACTTAC 600
 TGGTTATGTT GGATTGTATC AAGTATGTCA GACCTAACTG CGATGGGACA ATACTTTGCA 660
 40 TTTTGGTATC CACAAGTCCC AAATGGATT ACCGTGCTAT TTATTGTTTT AATCTTGATT 720
 AGCTTCAACT TATTAGGTGC CAGATTATTT GGTGAAGTGG AGTTTTGGTT CTCGATTATT 780
 AAAGTTGTCA CAATTATTGC GATGGTTATC GTTGGTCTTG TATTAATCTT TTTCTCATT 840
 45 AAAACACATT ATGGACATGC ATCATTCACT AACTTAATCA GTCACGGTGG CATGTTCCCT 900
 GGTGGAaCAT TTGGTTTCTT AATGTCATT CAAATTGCTG TATATTCACT CATTGGTATT 960
 50 GAACCTATAG GTGTaCTGC TGGTGAAaCG AAAGATCCTG AAAAAACCTT ACCGAAAGCA 1020
 ATTAATAATG TACCTATCCG TATTTTATTA TTCTATATCG GTGGTCTATT AGTAATTATG 1080

55

TTAATCGGCG TACCATTTC AGCAGGTGTC GTTAACTTTG TCGTGCTAAC TGCCGCGGCC 1200
 TCTGCTACAA ATAGTGGTAT CTATTCGAAT AGTCGTATCT TATTCGGACT GTCACAACAA 1260
 5 GGGTTAGGTC CTAAAGTTTT AAATAAAACG AATAGTCATG GCGTGCCTTA TTTATCAATG 1320
 TTAGTTTCAT CAATTGCATT ACTTATAGCA GCCTTGTTAA ACTACATTTT CCCTAATGCA 1380
 ATTCAACTAT TCATATACGT TACAACGTTA tCAACTGTGT TGTTTTTAGT TGTtTGGGCA 1440
 10 ATGATnATTG TCGCTTATCh AATGTATTTG GAAAAAGCAT CCTGAGGCA 1489

(2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 5000 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

TTTTCCATCA TrATcACCAT TTAAGACAAT AGCTGTATGA CCATATCCTC CACCAAATCT 60
 25 TCCACTAAAA ACCACTAAGT CCCCAGGTTC AGCTTTGAAA GTTGGTGTGT TGTGATAAAT 120
 TTTAGCTTCA CTATTAAAAT TATTGTCATA TGGTATATCT TTAGCTCCAT ATCCTTTTAA 180
 TCCAkGACCA TAAAGAkGAT TCCAATATAC ATTAACTAAA TCGAAACATT GCCATCCATA 240
 30 ACTACCGTCG AAATCCCATC CTCTATTCTC TAATGTTTTT AAATAATTTA ATGTACTGCT 300
 ATTACTATTA CTTTTATTAT TTGAAGACAC TGTTTTTGGT TTTGGTTCTA CTAATGGTGT 360
 CATTGGCACT TTTAATTTTT GACCAATAAA TATTAAATTA GGATTGCTA TATTATTGT 420
 35 ATTTTGAATA TTTGAAACTG TAGTTTTGTA TTTAATGCT ATAGCACTAA GTGTGTCTCC 480
 TTTTtTACA GTATAGATTT GTGTTTTTGG AGCTTCTCTA AACTGTAGT AACCAAAGTA 540
 ATTATTAGTA ACTTTATTCTG TTTTCTTCTG ATTAGAATTT TGAGCTTCCA AGTTTGCAAT 600
 40 TTTAATTTCT TTAGTAAGTT CATTGTTATT AATAACTAGA TTGTTACCTT GGCTTGAGTT 660
 TTTCCGAGTA TTTGAAATCT TTATATCTTG ATTAATTTCA TTTCCGTTTG AAATTGCTGA 720
 TTTGTTGTCT AACTTTAAAC TTGTGTCCGA TGTTTTAAAC GCACCTTCAT TTTTTATTTT 780
 45 GTCTTTTGTC GTATTTTTAT TAGCATTTAA CTCTGATTTT CGGAATACAT TTTGCTCATA 840
 CCCTCTTGTA AAATCTTTAG ATTTATCAAT TTCATCTGCA TATGCTTTGT TCGACATACC 900
 50 CAATGCCAAA AACATACCTA TTGAAATTGA CAAAATTCCA ATACTAACTT TTCTAATTGA 960
 ATAGCGTACT TTACTTTGTT GTTTATTCAT GATGAAACAC TCCTTACAAT AATATACAAA 1020

55

GTAGGAGGTA AAATAATTAA CTTGTCTTTC CAAAATATGA AAAGTGTACT AAAaTTCATC 1140
 GCACGACAAA TAGCCCATT TCCGATACTTT TATAAAGTAT GGAaTGGGCT ATAGCCATTT 1200
 5 ATATCATCTT TTAACCTTAT TTATTAACAG TTAATAATGA TTCATAAATA CCTGCTTCTT 1260
 TAGCAGCTTC AATTAATGTT GAACCAATTT CTGAAGGTGT TGCCGCTGTT TTCACACCAC 1320
 AACTATTTAA TGTTTTAATT TTCTCTTCAG CAGTACCTTT ACCACCTGAA ATGATTGCAC 1380
 10 CAGCATGTCC CATACGTTTT CCAGGAGGTG CTGTTTGTCC ACCGATAAAG CCTACAACCTG 1440
 GTTTTGTCAT ATTCGCTTTA ATCCATTCAG CTGCTTCTTC TTCAGCCGTA CCACCGATTT 1500
 CACCAATCAT AACAACTGCT TTCGTTTCGT CATCTTCATT GAATGCTTTT AAAACATCAA 1560
 15 TAAAGTTTGT TCCGTTGACT GGGTCTCCAC CAATACCAAC AGCTGTAGTT TGACCAATAC 1620
 CTTCTTCAGT CAATTGGTGC ACTGCTTCAT ATGTTAATGT ACCTGAACGA GATACTACAC 1680
 20 CAACATGACC TTTTTTGTGA ATATAGCCAG GCATAATACC AATTTTACAT TCATCTGCTG 1740
 TAATCACACC TGGACAGTTC GGACCAACTA AACGTGTTTT TCTACCTTGT AAGTAGCGTT 1800
 TAACTTTAAC CATGTCTAAT ACAGGAATAT GTTCAGTGAT ACAAATAACC ATATCTAAGT 1860
 25 CTGCATCAGC TGCTTCTAAA ATTGAGTCTG CAGCAAATGG TGCTGGAACG TAAATGACTG 1920
 AAACCGTTGC CCCAGTTTCA TTTTGTAGCTT CTTCAACAGT GTTGAAAACA GGAACGCCTT 1980
 CAACAACCTG ACCACCTTTA CCAGGCGTCA CACCTGCTAC TATTTTCGTA CCATAATCAA 2040
 30 GCATTTGTTT TGTATGAAA AGGGCAGTAG ACCCTGTAAT ACCTTGTTACC ATTACTTTAG 2100
 TATTCCTATC TATAAATACA CTCATCTTAG TGCTCCCATC CTTTCCTTAT GCTTCTTTGA 2160
 CTAGTTTAAC AATTTTTTGT GCACCTTCAG CCATTGTTGC TGCTGGTTCA ATTGCTAATC 2220
 35 CTGAGTCTTT TAAGATTTTT TTACCTAACT CAACATTTGT ACCTTCTAGG CGTACAACCTA 2280
 GTGGTAAAGT TAAATCTACT TCTTTTACAG CTTCAACGAT ACCTTCTGCG ATAACATCAC 2340
 ATTTTCATAAT GCCACCGAAA ATGTTTACAA AAATACCTTT AACATTTTCA TCACCTAAAA 2400
 40 TGATTTTAAA TGCTTCAGTT ACTTTTCTC TAGTAGCGCT TCCGCCTGCA TCTAAGAAAT 2460
 TGGCTGGGTT TCCACCGAAA TGATTAATCG TATCCATTGT TGCCATGGCT AAACCTGCAC 2520
 CATTAAACCAT ACATCCGATG TCACCATCTA ATGCAATGTA TGATAAATCA TGTTTAGACG 2580
 45 CTTCAATCTC TTTCCGATCT TCTTCTTCTA AATCACGTAA TTCTACAACA TCTTTATGTC 2640
 TGAATAATGC ATTATCATCA AAATTAATTT TAGCATCTAA TGCCAATACA TCACCATCAG 2700
 50 CTGTTGTAAC TAATGGGTTG ATTTCTACGA TTGAACAATC TTTTCAATG AATACATTAT 2760
 AAAGTGCTAA TAAGAATTTA GCAGCTTTGT TAACAGATTC TTTAGGAATA TTAATATTAA 2820

	AGATCTTTTC AGGAGTCTTC GCAGCAACTT CTTCAATCTC AGTGCCCCCT TCTTCAGACG	2940
	CCATCAATGT TACTTGGTCA GTCGCACGAT CAATAACGAA TCCAACGTAA TATTCTTTTT	3000
5	GAATAGCACA ACCTTCTTCG ATATATAAAC GCTTAATTTT TTTACCTTCT GGACCAGTTT	3060
	GATGTGTCAC CAAAGTTTTT CCTAATAATT CTTTTCGATA TGTTTCTACC TCAGATAAAG	3120
	ATTTAGCAAT TTTTACTCCG CCTGCTTTAC CTCTACCTCC AGCATGAATT TGTGCTTTTA	3180
10	CAACATAAAC ATCAGAAATT AATTCTTTTG CTTTCTCCAC CGCTTCTTCA GCAGTAAATG	3240
	CTACTCGTCC TTCTGGAAC GCAACGCCCA TTGAACGAAA TATTTCTTTA CCTTGATACT	3300
	CGTGGATATT CATCTTCCAT CCTCCTGTTA CTTAGGTTAA GTTCCCTTAC AATTATAAAA	3360
15	AATGTAAGCG CTATTGTAAA CTTAAATGCT ACTTTTTTAT CATTTAATG AATTTTACGA	3420
	TTTACAGTAA CGATTTTATA GGTCAAAGC TTTTCTATG CTCTTTCATA ATGCCAATAT	3480
20	CATCGATTGC TAGTAAATGT TGTGTGTAC CGTAACCCGC GTTTTTTTCA AAACCATATT	3540
	CAGGATAATC TTTAGATAAC TGTGTCATAT AATCATCAG AAAAACCTTT GCCATGATAC	3600
	TTGCAGCTGC AATGGACACA CTTCTTGCAT CACCCTTGAT TAAAGATACT TGAGGCAGTG	3660
25	CATTATCAAG CGTCATCGCG TCTATCAATA AATGCGTTGG TTGTACTGAT AATCCATCAA	3720
	TAGCTCGCTG CATGGCGATT TGAGTAGCTT TATAAATATT AAATTCATCT ATTTCTTTCA	3780
	GtGTCGCGAT CCCATATGCA AAAGCAGTAA CTTCATTTTT TAGTGCTTCA TTTAATTCTA	3840
30	GACGTTTCGT AACAGGTACT TTTTCGAGT CATCAAGGCC CAAATAATTG TGATTTGAAT	3900
	TTAAATTGT TGCGCATGCA ACGACTGGAC CTGCTAAAGG TCCTCTTCCA ACTTCATCAA	3960
	TCCCACAAAT AATAGCATTG GGATGCTCTT TTAATATTTT ATTTTCAAAG TAAGTCATTT	4020
35	CAACATACTT TTCTTTTAA GCTTGTTCTT TTTCTAACGC TTTTCTGCGC CTAGCTATGG	4080
	CATTTGAAC ACCTTTTTCG TCATCTAAAA AGCATTCATG ATTTTCTAAT TCTTCTATTG	4140
	TATTAACCGC ATTAATCAAC TGCCTAACTT CTTTAATTGT TAGCGTCATT TGCTAATTCC	4200
40	TCAGTCATAT CTTTAAAAAT ATCAAAACAA TAATTTCTTA TTTTAGCATT TCGAATATCA	4260
	TAAATAATCA GTTCAATGAC TGCTTCGTAA TCAATTTTAT TACCACGTCG AATTAAGCCC	4320
	ACGTTTTTTC CCTATCGCAT CAAACCACGC TATGATTTCT GCACTTTCAG GAACTTCAAT	4380
45	ATTATAATGT GACTTTAATC GCGCTAAATC ATTTTGAATT AAAAAGTTTA ATCCATAGAT	4440
	GGCAACTTCA TCTAAGTGCA CAATACTATC TTTTATCGCA CCAGTTAAAC TCAACTTCTT	4500
50	ACCGACTTCT TCATCTTCAA ATTTAGGCCA AAGTATCCCT GGTGTGTCTA ATAGTTGTAA	4560
	TGCATTACCA ACTTTAATCC ATTGTTGTTG TTTGGTCACA CCTGGTTTAT TACCAGTCTG	4620

55

AACGATCATT GCTCTTATCG CTCTAGGTTT AAGTCCTTTC GCTTTTTTCGC GTTCAAATTT 4740
 TTCAGCAGTC GCCTTAATTG CTGCAGCTTC CACTTTCTTT AAATTTTAC CGTGCTTAGC 4800
 5 ATCCACTGAT ACAGGATAgT AtCCTTTATC AATAAAAAAT tGTTCCCATTT TTGACATCTC 4860
 ATTTAAATTA GACATATCTT TTTTATTTAA TATAACAACA CGTGGTTTTT GGTTAATAAC 4920
 TTCATCTATC ATAGGGTTTC TTGAAGTATA TGAATTCTT GCATCTACTA GTTCAAACAC 4980
 10 TACATCTACT TTTTTTAATT 5000

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 1700 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

AACCCGnAGA CGAAACTnCA TAGTTGCCTG TTATGTTGAT TAGTGCCTTT GTAGCTATTT 60
 25 nAAATCAAAC ATTATTAAAT ACAGCGTTAc CTAGTATAAT GAGAGAATTA AATATCAATG 120
 AAAGTACATC GCAATGGCTA GTTACTGGGT TTATGCTTGT TAATGGCGTC ATGATACCTC 180
 TGACGGCATA TCTAATGGAT AGAATTAAAA CTAGACCTTT ATACTTAGCG GCGATGGGGA 240
 30 CATTTTTATT AGGTTCTATT GTTGCAGCCT TAGCTCCGAA TTTTGGAGTT TTAATGTTAG 300
 CTCGTGTAAT TCAAGCGATG GGTGCAGGCG TACTTATGCC CTTAATGCAA TTTACGTTAT 360
 TTACATTGTT CAGTAAAGAA CATCGAGGTY TtGCAATGGG ACTAGCAGGT TTAGTAATTC 420
 35 AATTTGCACC AGCAATAGGA CCTACAGTTA CAGGATTAAT TATTGATCAA GCGAGTTGGC 480
 GAGTtCCATT TATTATAATT GTAGGAATTG CTATACTTGC CTTTGTtTTC GGTTTGGTTT 540
 CAATCTCGAG TTACAATGAA GTGAAATATA CGAAATTAGA TAAGCGTTCA GTAATGTATT 600
 40 CAACTATTGG GTTCGGGTTA ATGCTATACG CATTTAGTAG TGCAGGAGAT TTAGGATTTA 660
 CAAGTCCAAT AGTAATAGGT GCGTTGATAT TAAGTATGGT TATTATCTAT TTATTTATAC 720
 GTAGACAATT TAATATTACT AATGCACTTT TAAATTTAAG GGTTTTTAAA AATAGAACAT 780
 45 TTGCATTATG TACGATTAGT TCAATGATTA TAATGATGTC AATGGTTGGA CCTGCGCTGC 840
 TTATACCGCT ATATGTTCAA AACAGTTTAT CTTTATCTGC CTTGTTATCA GGAATTGTTA 900
 50 TCATGCCTGG TGCAATAATA AATGGTATTA TGTCAGTTTT TACAGGTAAA TTTTATGATA 960
 AGTATGGTCC TAGACCATTG ATTTATACTG GTTTTACAAT TTTAACAATT ACTACAATTA 1020

55

GAATGTTTTT AGTTTCTTTA CTCATGATGC CGATAAATAC TACAGGAATT AATTCTTTGA 1140
 GAAATGAAGA AATCTCACAT GGCACGGCTA TTATGAACCT TGGTCGTGTA ATGGCTGGTT 1200
 5 CACTAGGCAC AGCTTTAATG GTTACATTAA TGAGTTTTTG TGCAAAAATA TTTTATCTA 1260
 CATCGCCATC GCATTTAACT GCAACTGAAA TTAAACAGCA ATCCATTGCT ATAGGGGTGG 1320
 ATATCTCATT TGCTTTTGTA GCTGTGCTTG TTATGGCAGC TTATGTGATA GCACTTTTTA 1380
 10 TAAGAGAACC TAAAGAAATA GAAAGTAATA GAAGGAAATT TTAAAATAAT TATAGTAGTT 1440
 GGTCTATTTA AAATAATAGG CTAAGTCTT TTTTATTTA ATAAAAAGTT TTATACTTTT 1500
 AGTGATAGAC TAAGCAAAAA TTGTTATTTG CTATGATGTA GATGTCTTAA AATGATTAAG 1560
 15 GGGGATTTGC TTTGTTAACG GTAGATCAAG TGAAAGAATT GGTAGGAGAA ATTAAGATC 1620
 CTATTATAGA TGTGCCTTTA AAAGAAACAG AAGGTATTGT TGAGTTTCT ATTAAGGGAG 1680
 20 AAnAAGAACA TGTGAGTGT 1700

(2) INFORMATION FOR SEQ ID NO: 243:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10146 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

TGCATCAACT TTCAAACAAT AAATCATCAC AATAACCACA CCTAATTCAA CACTTTTCAA 60
 ACATAAGTAT TGACACATTG AGCAAAATGA TTTTAAATTG TAACTAATAC AGTTACAATT 120
 35 ATGAGGTGAG AAACATTGAA TTTAGAATTT AACATTGCCG TGCATGTATT AGCTTTTTTA 180
 ACTAAGCATC aTTCAGAAAA ATTCAATAGT AGTTCATTAG CAGaaTTAaC TTGTTTAAAT 240
 CCTGTTCAAT TACGACGCGT GACGACTCAA CTTGTGCGATT TAAAAATGAT TGACACAATA 300
 40 CGAGGTAAAG ATGGCGGTTA TTTAGCAAAT GATCAAAGTG CTGATGTCTC TCTAGCAACA 360
 TTATaTAAAC ATTTTGTCTT AGAGAAAGAA CACCACACAC GTCTATTTAC TGGCGACGAA 420
 GGCAGTCACT GTCAAATTGC TCGTAATATT GCAACTACCA TGTACATTA TCAGCAAGAC 480
 45 GAACAGAATA TCATTATTAA TTTTATAAT GGAAAAACAA TCAAAGATGT CATTGAAGAC 540
 ATTCAAAAGG AGGATTTATG TCATGAAAAC ATATGATTTA ATTGTAATAG GATTGGGAA 600
 AGCTGGTAAA ACTTTAGCGA AATACGCTGC ATCAACAGGT CAACACGTCG CAGTTATCGA 660
 50 ACAATCTCCG AAAATGTATG GAGGCACTTG TATAACATA GGATGTATAC CTTCGAAGAC 720

	TGTTGTCAAT GCGCTAAACA ATAAAAATTA CCACTTATTA GCAGACGATA ACAACATTGA	840
	TGTA CTGGAT TTTAAAGCGC AGTTTAAATC TAATACTGAA GTGAATTTAT TAGATCAACA	900
5	TGACGATATC GTCGATAGTA TTACTGCACC TCATATCATT ATTAATACCG GTGCTACCTC	960
	TGTCATTCTT AACATTAAAG GCCTTGATCA AGCTAAACAC GTCTTCGATT CGACAGGTTT	1020
	ATTAAACATT AGCTATCAAC CTAAGCACCT CGTTATTGTA GGTGGCGGTT ATATCGCCTT	1080
10	AGAAATTGCT TCAATGTTTG CGAATTTAGG TAGTAAGSTA ACAGTATTAG AACGTGGCGA	1140
	ATCATTATG CCACGCGAAG ATCAAGATGT CGTTGCATAT GGTATTACTG ACTTAGAAAA	1200
	TAAAGGCATT GCATTGCATA CAAATGTTGA AACGACTGAA TTGTCATCTG ACAATCATCA	1260
15	TACAACAGTC CATAACCAACG TTGGTAACTT TGAGGCTGAT GCAGTACTTT TGGCTATCGG	1320
	GCGCAAACCG AATACGGATT TAGCTTTAGA AAATACTGAT ATCGAATTAG GCGACAGAGG	1380
20	CGAAATTAAA GTCAATGCTC ATCTTCAAAC AACTGTGCCG CATATTTATG CTGCAGGTGA	1440
	TGTTAAAGGC GGACTTCAAT TTACGTATAT ATCTTTAGAT GATTATCGAA TTATCAAATC	1500
	AGCGTTATAT GGTAAATCAGT CACGTACGAC TGACAATAGA GGCAGCGTGC CTTATACAGT	1560
25	ATTTATAGAT CCACCATTAT CACGTGTTGG ATTAAGTAGT AAAGAAGCTG CCGCTCAACA	1620
	TTATGATTAC ACTGAACATC AACTTTTAGT AAGTGCTATA CCTCGTCATA AAATTAACAA	1680
	TGATCCAAGA GGTATTATTA AAGTAGTCAT TAATAATGAA AATAATATGA TTTTAGGTGC	1740
30	TACATTATAT GGTAAGCAAT CTGAAGAATT AATTAATATA ATTAACTTG CGATTGATCA	1800
	AAACATTCCA TATACCGTAT TACGAGATAA TATTTATACG CATCCTACGA TGGCCGAATC	1860
	ATTTAATGAT TTATTTAATT TCTAGACAAA ACATAAAAAC CTGGTGGCAC GCATTGAATG	1920
35	ATGCTGCCAT CAGGCTTTAT TGTTGTGCTT TTCGCTTTTC TAATTTTCT TTAAGCTTTC	1980
	TATCTGTTC TTCTTTACGA CGTTTACGTT CTTCATGTG TTTTCTTAAA CGCTCTTCTT	2040
	CTTCAGGATC ACGTGGTTTC TTTAATTGTT GAGAACTTT TTCGATTAAT TCTTCTTCAG	2100
40	TAAGCGCagc CAGTGGGCGG TTATTAACAA AAGTGAATGT TTTTCGGCGT CCAGGTCCAC	2160
	AATAAGATTG ACAACCTATC ACGATTTTACG CATCGGGATC TAATTTTCC AACTTCTTTT	2220
45	GTAACGTTCT TATATTGACT GCCTGACATT CATCACAAAT AAGGAATGTA TTTTTCATAT	2280
	TGCTACCCAC CTTTCTTTAT CATATCTATA TCGTCGATTT CATTAATTTT TCGTTAACT	2340
	CTATCTATTT TACTCTTTTT AATATTTTTT TCAAGATACG TAACACGGCT GaCAATAAAA	2400
50	AATGGAGCAT TTATCTTCTA ATTAAATTAG ATGaTTGCTC CCCTATCAAA TCATTATTG	2460
	CCCATGATAA ATATTAAATT TTAATGGTTT AATACCATGT TTTGTCCATT CATCATAAAT	2520

	TGATGTTTTA GATGCGCCGT GATACTTTTC AGCAATATCA CACAAATATT TTAGCTTTTC	2640
	AGTTTCTATA TCAACTGTAG CTTCTTTATC CATACGTTGA ATAATTGTAC GATTCTGACG	2700
5	CACCATCTTT TGCACACCTT TAATGTTATT TGTTTTAAAA GCATGAATAA GTTTTTCAAC	2760
	ACAACGATGT GAATCTTCTA AGAAGTCACC GTAAAATGAA GGATCTGATT TCAAACGTTT	2820
	CACTTCGCTA ACAAAGTGTG GTGATGACGC CGGTGAGCCA GTCCAACCGA TAAGTACTTC	2880
10	CATATTTTCA GGTGCTTGTA ATGGTTCGAT GTGCAATCCA GGCCAGTTTT TGATTAAAC	2940
	TTCTTCAACC GTAGTATCTT CAATTTGATG CTTAACCCTAT TCATGATCAA AAGTACTATA	3000
	CGCTAGCCAT CCACTATATA CACTCACAGC AATATCTCCG CATGAACCTA AACTTTGTAA	3060
15	CTTCATATTT GCAATCACTG CTAGTTTATA AATGTATAAA TTAGATAACT TCATATCATA	3120
	AAATTCATTT AATACTTTTA TAACTGACAC AAGTACTGCT GCACTTGAAC CTAATCCATA	3180
20	TTTATGACCA TTTGAATCAT CTAAATTACT ATCAATAGTC AGATGAAAAT GCTTCATCGC	3240
	TATATCGCAA CTTTTCGCGT ATTGTTCAAA TATTTCAATA GCTGTGACCA CATAATTTAA	3300
	TTGTTTTGCT GcATGTGGAT CTGAAATGAC AATACTATCT TCATCTCTAC TAAATGTAAC	3360
25	TGGGTTATGA TGTAAATGCTT TTGAATGAAT GGTACCTTTA TATTGGTCTG CTTCTTCAAT	3420
	AGTAGCAGTT AAAAAACGAT CTAACGCAAT AAGTACAGAT TTATATCCTG GTTCTGTTAC	3480
	AGCATATTCT CCAGCAATAT AAAGTTTTCC GGGTGCTTTG ACCTGAATCA TTTTATCTCT	3540
30	TCCTTACTCA ATTATTTCAA TTCTGTGGC AATAATGTCA CTATCAATAA TTTGGTTATT	3600
	ATCAAACGTG GTTAATAATT TATCTATAAT CTGTTGCTTG TTTTCTTTT CTACAAGTAT	3660
	TTTCACATTA GGTCCCGCAT CCATTGTAAA ATAACACGGA TACCCCGCTT CTCGGCATTC	3720
35	GTGAACAAGC GCCATGACAT CATAACTTTT TTGCACAAGA TATGTGAACG GCGGTGTTGA	3780
	TCCTAGATTG GTGGCATGCA TACGCAAACC ATTTTCTTCA ATTACTTCAC CAAGGCGTTT	3840
	AAAATCTTTG TCTTGAATCG CTGCTTTTGC TTCAGCTAAA TCTTCATCAA TATGATCTAA	3900
40	CCAATATTGA TAAAACCTTG ATGTGTTTCG TGTCAATGAC ATACCATATC GACTAGGTAC	3960
	CTTTTITAGAA TGTGATTAA TCACAACAAA TATCATGGCA AGGTCACTTT CAAAATGATT	4020
45	CGATTCAAGT GGAACGGCAT ATGACGTCTC ATCACTATAC CCTTTTTCCC ATTCTGCAAA	4080
	TCCACCATAA ATACTACGCG ACGCAGAACC CGAACCAATT CGCGCCAATC TCGATAAATC	4140
	CTTATCTGAC AGCTGCATGT CTAGCGCTTG ATTACAAGCT GCTGCTAAAG CTGCATATGC	4200
50	GCTTGCCGAT GAAGCCAACC CTGCTGCTGT TGGTACAAAA TTGTCGCTTT CAATTTCTGC	4260
	ATACCAATCG ATGCCAGCTC TATTTCTGAC AATATCCATA TATTTTGAAA TTTTCTCTAA	4320

AAAAGTGACT TTCGTTTCAG TGTAAAATTT TTCTAATGTA ACAGATATGC TATTATTCAT 4440
 TGGAAATGATT AGTGCTTCAT CTTTTTTACC CCAATATTTT ATAAGTGCAA TATTCGTATG 4500
 5 TGCACGTGCT TTGCCACTTT TAATCAACGC ATTAACCTCC TAAATTCTCA ATCCAAGTAT 4560
 GTGCTGCACC AGCTTTTTCT ACAGCTTTTA CAATATTTTT CGCTGTTGGT AAATCTTTGG 4620
 CAAGCAATAA CATACTTCCA CCACGACCAG CGCCAGTAAG TTTTCCAGCA ATCGCACCAT 4680
 10 TTTCTTTACC AATTTTCATT AATTGTTCTA TTTTATCATG ACTAACTGTC AACGCCTTTA 4740
 AATCCGCATG ACATTCATTA AAAATATCCG CTAAGGCTTC AAAGTTATGA TGTTCAATCA 4800
 CATCACTCGC ACGTAAACT AACTTACCGA TATGTTTTAC ATGTGACATG TACTGAGGGT 4860
 15 CCTCACAAAG TTTATGAACA TCTTCTACTG CTTGTCTTGT TGAACCTTTC ACACCAGTAT 4920
 CTATAACAAC CATATAGCCG TCTAACTTA ACGTTTTCAA CGTTTCAGCA TGACCTTTTT 4980
 20 GGAACCAAAC TGGTTTGCCT GATACAATCG TTTGCGTATC AATACCACTT GGTTTACCAT 5040
 GTGCAATTG CTCTGCCCAA TTAGCCTTTT CAATGAGTTC TTCTTTCGTT AATGATTTCC 5100
 CTAAAAATC ATAACCTGCA CGAACAAAG CAACCGCGAC AGCTGCACTC GATCCTAATC 5160
 25 CACGTGATGG TGGTAAATC GTTTGGATCG TTACTGCTAG CGGCTCTGTA ATATTATTTA 5220
 ATTCTACAAA ACGGTTCAAC AAAGACTTAA GATGGTCAGG CGCATCATAT AACATACCAT 5280
 CGTAAACATC GCTTTTAATA GACGAATAGT TCCCGCTCTC TAAGGCTTCT ATTAAACTT 5340
 30 TGATTTTACC TGC GTTAAAC GGTACTGCAA TAGCAGGCTC TCCAAATGTA ACAGCATGTT 5400
 CTCCTATTAA AATAATCTTA CCTGTCGATT CCCCATATCC TTTTCTTGTC ATGTCAATAT 5460
 CACCTTTTAT ATTTATCCTA TACTTGATT CATTATTTTA TTTATTAGTA AAAGACATCA 5520
 35 TATTCTAAGT TGCAATngCAT TCGCGTTAAA TTTCATTGCA GTCTTTATCT CACATTATTC 5580
 ATATTATGTA TAATCTTTAT TTTGAATTTA TATTTGACTT AACTTGATTA GTATAAACT 5640
 AACTTTCGTT TACTTCAAAG TTTAAATCTT ATCGAGTGAT ATTTTCAGAT CTTTATCTTT 5700
 40 TTATAAAATA GCCCTACAAT TTATAATTTT CCACCCTAAC TATAATACTA CAAATAATAA 5760
 TTGGAATATA TAGATTTACT ACTAAAGTAT TAGAACATTT CAATAGAAGG TCGTTTCTTT 5820
 CATAGTCATA CGCATTATAT ATACCCTATT CTCAATCTAT TTAATACGTA AAACATGAAA 5880
 45 TTTTCTIATT AAATTTATTA TTTCCATCAT ATCATTACTT TTAATTTAAT GATGTTCAAT 5940
 TTAAATATTA GGTCAATAAC ATATTTATGC TTTTATGGA TACTTTCAA AATAACAGCC 6000
 50 CCAAACGATA ACTTGAAAGG GGCTGTAAA TATTTAACTA TTGCATTTGA TCTATCATTT 6060
 TCTTGTCTTCT TTCAATCATT TTATCAAAAT ACGTATCGTA TCTTTGCCAT TCTTCTTGAG 6120

55

	TTACATCTTG AACAGTAATC GTTTTGTTAA GCAATGTCTC TAATGAGGCC ATACAAGATG	6240
	GTTCAATTTC AGGATATTTA AATTTAGTCA CTTACACCTTT TAAAGCATGT TCATAAAATG	6300
5	TTTGCATCAT CAATGCACGT TCTGAACCAG AGCCTTCAAC ACAAAGATAA ATTTGTACAG	6360
	CAATACCGCC TCTAACTCTT CGTTGCGATA TGCCTGCAAA TTTCTTACCA TCGATACTTA	6420
	AGTCAAATTT TCCTGGGCAA TAAGAATGTT CAATTTCCAT CGTATCAATA TCAACATTCT	6480
10	CATTTTCGAA CATTTTGCTA ATTAAGAGGT ACATCACAGT AAACGCTTCA TCAATCGTTG	6540
	TTTCTGTTTG TCCTTTGAAC ATCAGCGATA TATTTAATAC ACCTTGATCT AGAACGACAC	6600
	CTAAGCCACC AGAATTTCTA ACAATGGCAT TATAACCAAT CTCATTGCTT AAATAATCAA	6660
15	TGCCATCTTT TAAAAACGGC AATCTTGAAT CATGAATACC AAGAATAACA GTATGTTGAT	6720
	GAATCCAAGT ACGCACAACA TTATCTGATA TATCTTTGCC CACACTTTCC CAAAATGTAT	6780
	CATCGAATGC GAAAGATTGC ATAGGTTCTA ATCCAGAAGA ATGATCGATA TATCGCCAGT	6840
20	TGACGCCATT AAAATATTTA CTCGCTAAAT CCATCGTTAT TGTAAGGCTT GCGCTGCTGT	6900
	AATAATTGAA AGATTGTATA CATCTTCAAT TGAGCAGCCA CGTGATAAGT CATTTACTGG	6960
25	AGAAITTTAAA CCTTGTAATA CTGGACCAAC TGCATCATAT CCACCTAAAC GTTGTGCAAT	7020
	TTTGTAACCA ATATTACCAG CTCTCTAACT TGGGAATACA AAGACATTTG CATCACCTTG	7080
	TAATTTAGCA CCTGGCGCTT TTTTCTCAGC AACACCTGGT ACAATCGCAG CATCAAATTG	7140
30	GAATTCGCCA TCAATGATTG CTCTCTAATTT TtCTTCTTCA GCTTTTTtGTT GTGCTAATTT	7200
	GACAGCTTCT TGAACTTTTG TCACGTCGTC TGATTTAGCA GACCCTTTTG TTGAAAAGCT	7260
	TAACATTGCA ACTTTTGGAT CCATGCCAAA GCTTAATGCT GATTTTGCAC TTTCTACTGC	7320
35	AATTTCTGCA AGTCCTTG TG AATCAAGTTC TGGATTGATT GCACAATCAC CAAAGATGTA	7380
	TTGTTCAATCA CCTTTAATCA TAAAGAAGAT ACCTGATGTT CTTGATACAC CTGGTTTCGT	7440
	TTTGATGATT TGTAAAGCTG GACGCACAGT GTCGCCTGTT GAATGTGCTG CACCACTAAC	7500
40	TAAACCATCT GCTTTACCAG CATAACAAG CATTGTACCG AAGTAGTTCA CATTGTTTAA	7560
	TAATCTTGT GCTTGTCTT CAGTCGCTTT ACCTTTACGT CGTTCAACAA ATGATTGAAC	7620
	TAATTCAGCT TTCAATTCAC TTGTCGCAGG ATTAATTAAT TCAATATTAG AAATATCAAG	7680
45	ATCAAGTTTT TGCCTAAAG ATTGAACCTT AGTCTCATCA CCTAACACGA TTGGTGTAAC	7740
	ATAATCTGTT GCTTGTAATT GTGTTGCAGC TGTTAGAACA CGTTCGTCCT CTCCTTCAGG	7800
50	TAATACGATT TTAACGTTTT TACCAGAAAG TTTGTCTTTT AATACATTTA ATAAATCAGC	7860
	CATAATGTCC TCCTGTAATA TAAATCTTAT TAATCATTCA CGGTATAATT ATACGCCATT	7920

55

	TATGATAAAA	TTTATAAAGA	ACTGATGATT	TTTGAAAAGG	AGCGATAAAC	ATGAGTCAAG	8040
	CAGCCGAAAC	ATTAGATGGT	TGGTATAGTC	TACATTTATT	TTATGCAGTT	GATTGGGCAT	8100
5	CATTACGTAT	AGTTCCAAAG	GACGAACGCG	ATGCACCTGT	CACTGAATTT	CAATCATTTT	8160
	TAGAAAATAC	AGCAACTGTA	AGATCATCAA	AATCTGGTGA	TCAAGCTATT	TATAATATAA	8220
	CTGGTCAAAA	AGCAGATTTG	TTATTATGGT	TCTTACGTCC	TGAAATGAAG	TCTTTAAATC	8280
10	ATATTGAAAA	TGAATTTAAC	AAATTGCGCA	TTGCTGACTT	CCTAATCCCT	ACATATTTCAT	8340
	ATGTATCAGT	CATTGAATTG	AGCAATTATT	TAGCTGSTAA	ATCTGATGAA	GATCCTTATG	8400
	AGAACCCTCA	TATCAAAGCA	AGATTATACC	CAGAATTACC	ACATTCTGAT	TATATTTGTT	8460
15	TCTATCCAAT	GAACAAACGT	CGTAATGAAA	CTTATAACTG	GTACATGTTA	ACTATGGAAG	8520
	AACGCCAAAA	ATTAATGTAT	GACCATGGTA	TGATTGGTAG	AAAATATGCT	GGCAAAATCA	8580
20	AACAATTTAT	TACTGGTTCT	GTAGGGTTTG	ATGATTTCGA	ATGGGGCGTA	ACATTGTTCT	8640
	CAGATGACGT	ATTACAATTC	AAAAAAATTG	TATACGAAAT	GCGCTTTGAT	GAAACAACAG	8700
	CACGATACGG	TGAATTCGGT	AGTTTCTTTG	TAGGACATAT	TATTAACACA	AACGAATTCG	8760
25	ATCAATTCTT	TGCGATTTCT	TAATACATTG	GTACGTTTAT	AAATTAATAA	AAAAATTCCA	8820
	AGCTTATCGG	TTTAAGCTTG	GAATTTTTTCG	TTTATCTTCA	GTATATTCCC	GTATACATAA	8880
	GACGTGATTT	GGTAAATAGT	TGAAATCTGT	ATGTTTAAAC	TTATATATAT	GTGCTAATGT	8940
30	ATTATCAATA	ACAAAGtACA	CTTTGCTCAT	AGCAAgTsac	CCGagTAGTC	TTCCTTGGGA	9000
	GAACTTTAAAC	TACTATCACT	ACATATAAAC	GTTAACCTCA	ATAGAAATTA	TACAGTCGCT	9060
	ACTCTATACA	ATTTTTGTAA	TGGTTAACTA	ATATTATTTT	AACCTATTTG	AAATATTTGA	9120
35	AACATATTTT	TGTCGAATTT	TTTTCAATAA	TTTTTCCTTT	TTTACTTCA	AGAGAATTTT	9180
	AACTACTAAA	AATTCCGATG	ATTATTATTA	CAATAGTATC	AAATATTAGT	TTTTTAAAT	9240
	CAATAACAAC	TTATCAAAAA	GTCATGTGG	TTATTTTATA	GTGTATAAAC	TATAATGAGT	9300
40	ATTAAATTCT	TATAAACAAT	GGTGATGAAA	TGGACATAAA	TTCAGAAGAA	TACAAACAAG	9360
	AGGTACTTAT	CAAAGACGTT	GTCTGCTTG	CTGCTCGCAT	ACTATTAGAA	TCTGGTGCAG	9420
45	AAGGTACGCG	TGTAGAAGAT	ACCATGACAC	GTATTGCAAA	AAAACCTGGT	TACAGTGAAA	9480
	GTAACAGCTT	TGTTACAAAC	ACTGTCATCC	AGTTTACGTT	ACATTGGGAA	TCGTTTCCTA	9540
	GAATATTTAG	AATTACCTCT	CGAGATACAA	ACTTAATAAA	AATTTCTCAA	GCTAATAAAA	9600
50	TTTCGCGTCA	AATTACAAAC	AATGAAATTT	CTTTAGCCGA	AGCAAAAACG	CAACTTGAAA	9660
	AAATATATGT	TGCTAAGCGT	GACAGCAGTC	TTCCCTTTAA	AGGTTTTGCT	GCAGCAATGA	9720

55

TAGCAGGTAG TCTAGGATAC CTAGTCACTG AGATTTTAGA TCGTAAGTwa CACGCACAGT 9840
 TTATCCCAGA ATTCaTTGGT TCaTTAGTTA tTGGGATTAT CGCCGTTATT GGACATACAC 9900
 5 TTATTCCAAC AGGTGACTTG GCAACTATTA TCATTGCGGC AGTCATGCCT ATTGTTCCCTG 9960
 GTGTATTAAT AACAAACGCA ATACAAGATT TATTTGGTGG ACACATGTTG ATGTTCCACAA 10020
 CGAAATCATT AGAAGCATTG GTTTnGCGTT TGGCATCGGT GCTGGCGTTG GTAGCGTATT 10080
 10 AATTTTAGTA TAGGAGTATC AGACTATGTT TTGGATCTTA AACTTTATCT TTAGCTTTTT 10140
 AGCTTC 10146

(2) INFORMATION FOR SEQ ID NO: 244:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2022 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

25 ATTTAATTGG TTGGTGGCGT ATTCTCnTTT CCAAGATTGG CCACGATTGG ATGTTCAGTC 60
 GACAAATATG TAACCAGTCC CGATAAATTA CGTTCCTCAA TCATTGATAG TATGTCTGGT 120
 GTTTGTAACG ATTTATCGAG TAAGGCATCA AGCAATTGAT AATGTCCCAA TACAACAAAA 180
 30 TGCACGTTGT CTCTCAACTG CTGTTGAATA AACTGAATAA AGAGCTTTAA GCTCTGTTGC 240
 ACATTCGCTA ATGATGGTGC ATAGTTTTC AAACCAACTT GTACAGCCGC TTCATTATTC 300
 CGAATGATTA AACCTGTGTA TGCCACTTTT GTTGCTGCAG TTGGATACAT TGAGTAATAA 360
 35 CGCAATAATT GATCTGTAAA ATCATTTCGA AGTGCATAAA TTTGATGCTC ATGTTGCCAA 420
 AAATtTCGCT CACCCATCTG CTgCAAATCC TCATGGTTCA ATTGTTTCCA GTCCAACTTT 480
 TCAACCACAC TAAATCAAC TAACTCATAA TCCGCTTTAT TAAAATATTT TAAAATGCT 540
 40 GTTCCGATT CTTTAAACGC AATTAATTGT TCTGnATTAT TCACTCGACC ACCCTTTACT 600
 TTCAATACTG TATTTAAAAT CACTTGGTAT TTTCGTTGTT TGCTTTACTT CTCTACCACG 660
 CTAAAGTGTA ATATGATTAA TAACTTATCA TTTTtagCAA TACATTACAA CCTTTTTTCAG 720
 45 AAAATTCGGT GTATTGATTT TAAAATTTTT TAAAATAnAA AAGGCAAGAC ATTTGTGCCT 780
 ATAAAAATGC TTAACCAAGA TTTTATATT GaAGTTGTAC TTCTGCACA TATTGTCCTT 840
 50 GCCTTATTAT GTAAAGTTAT TTTCTTTCTA TCTTTTTIATT AAATTTAACT ATTCTTCATA 900
 ATCCCCGATTC CCTTAAAGT AACGTCTATC TTGTTTACTA TATACATTTT CAGGATTAAA 960

TTGGTAAACG TTCGTTGCTG ATATATCTGT AAAATTGTTT GGACCGACAC CTGCAATAAA 1080
 CTTAAACTCT GCTTCATCTA CCAAATAATC ATACGCTTGT GTATGTCTAT CCTGTGCGCC 1140
 5 ATGTGGAAAT ACAAACATAT CTGTTTTACC TACAATTGGT TCAACTTCAT CTTTCCATCT 1200
 TTTAGTATCA CGTTTAATAC CTTCTAAAGA TGTTTTTCA AAATTAATGT GACCATATGA 1260
 ATGACTCGCA AATGACCATC CATCCCGTTT CATTGCGCGA ACAACTTCCT CAGCTGCCTT 1320
 10 TTTATTCTTT GTATAATCTT TACTCGTTAA TTCATTCTGT CGATAACCTA ATACGCCCTC 1380
 ATAACCGGTT AAAGCAACAA CACCTTTTTT ACCATTAAA GAAAAATCTG GATGCTCTTT 1440
 TACAAATTTA TTTAAAATG GCACGATATC ATTGTCATCA GAATAAGTAG CATGGCCTTT 1500
 15 TTTGTCTGTA GTTTCAGAAA CAACATGTTT ATTTTATCG AGTACTAAAC GGTGAGCATA 1560
 ACCATGGTGT CTCATGTAAC TATAGTAATT CATATCATCA ATTGAGATGA TTAGTGGCTT 1620
 20 TTTACCTTTC GGCAATTTTA TTTTTTTGGC TTTTACATGA TGAGATGATA AGTCGTATAC 1680
 ATCATGTGGA TTAACGATGA TGTAATTATT TTTATATAAT TCGTTCAATG ATTTTTTAAA 1740
 TTCACCTTACA GTAATCATCC AATCATTGTT GCCCTTAGCT TGGTGTGTAT CTCCTGTAAA 1800
 25 CGCAACTTTT GGGTCTGTAA TTAATGGGTG ATAAACACA TGATAAACTT GGCCGTGATA 1860
 TGTTTCCCAA TGTTTCATCCA TTTTcGATTT aTGCTTTGCA TACTCATTTG GATTAACAGA 1920
 TTTATTkTGA GCTTTCTCAT TTTGCTTGGA ACAGCTATAT mACAATGCAA CTGATAATAA 1980
 30 CAGAAAAAAT AGCAATAAAT ATTTTTTATG CATTAAACAT TC 2022

(2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 1340 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

ATAGAGTGAT AAAGGATGGT TGTCAATGA TAAATGCAGT AGTAATAGCA GTAATTTTAA 60
 45 TGATTGTGCT ATGTTTATGT CGATTAAACG TnTTATTAAG CTTATTTATC AGTGCGCTAG 120
 TTGGTGGCTT AATTTcAGGC ATGAGCATTG AAAAAAGTTAT AAATGTATTT GGGAAAAATA 180
 TAGTCGATGG TGCTGAGGTA GCATTAAGCT ATGCTTTATT AGGTGGATTT GCAGCATTAA 240
 50 TTTcATACAG TGGTATCACA GACTATTTAG TAGGAAAAAT TATAAATGCA ATTCACGCTG 300
 AAAATAGTCG ATGGTCAAGA GTTAAAGTCA AAGTGACAAT AATCATTGCA TTATTAGCTA 360

55

CACCATTTGTT AAGTCTGTTT AATGACTTAA AAATAGATAG ACGTTTAAATC GGTTTGATTA 480
 TCGGTTTTTGG TTTATGTTTC CCGTATGTGT TATTACCATA TGGATTCCGT CAAATTTTCC 540
 5 AGCAAATTAT TCAAAGTGGC TTTGCAAAGG CAAATCACCC AATTGAGTTT AATATGATTT 600
 GGAAAGCAAT GCTTATTCCT TCAATGGGGT ATATTGTTGG CTTACTTATC GGTTTATATG 660
 TATATCGTAA ACCACGTGAA TATGAAACAC GTAAAATTTT AGATAGTGAC AATGTTACAG 720
 10 AGTTAAAACC ATATATCTTA ATAGTAACAA TTGTAGCAAT ACTAGCTACA TTTTGTAGTAC 780
 AAACATTTAC AGATTCAATG ATTTTGGTG CACTGGCAGG GGTACTCGTA TTCTTTATTT 840
 CACGTGCATA TAATTGGTAT GAATTAGATG CTAAGTTTGT TGAAGGTATT AAAATTATGG 900
 15 CTTATATTGG TGTAGTTATT TTAACAGCAA ATGGATTTCG TGGTGTAAAT AATGCTACTG 960
 GTGATATAGA TGAATTAGTT AAAACTTTAA CAAGTATTAC TGGTGATAAT AAATTATTTA 1020
 GCATTATCAT GATGTATGTG ATAGGTTTAA TTGTCACTTT AGGTATTGGA TCATCATTTG 1080
 20 CAACAATTC TATTATCGCA TCATTATTCA TTCCTTTTGG AGCGTCAATT GGACTAGATA 1140
 CAATGGCATT AATCGCATTG ATTGGAACAG CGAGTGCATT AGGTGACTCA GGTTCGCCTG 1200
 25 CAAGTGATTC AACATTAGGA CCAACTGCGG GATTAAATGT TGATGGCCAA CATGATCATA 1260
 TACGTGATAC ATGTGTACCA AACTTCTTGT TTTATAATAT TCCTTTAAAT GATTTTCGGT 1320
 ACTATTGCTG CTATGGTACT 1340

30 (2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3365 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

40 CAAAATCTAA GAGAATAAaA TTTGTTAATT TnAAATAGCA AGCAATTCAA AGTTATATGT 60
 GTAATAGATA AAATAGATAT CCCTATAGTG ATGCGTTACT AGCTAAACAT AATAACACAT 120
 TAGAAGATAA TGAAGTTAAG GAGTTACTGG ATTGTTTCGA CTATGTAATT AAGTATAAAA 180
 45 ATATCCAACG ACAAACGTA ATTATAAAAT GGTAAGCT ATGGTACAGT TTCAAATTGC 240
 TAATGACATG CGTATCGGTG AGCTACTTGC AATAAAGAGA GTAAATATAA ACTATGAAGA 300
 TAAAACGCTA GATATCGACG GTAAAGTTAA TTGGATAACT GAAAAAGAC GGGAGCATTC 360
 50 GGAGTAAAGG AGACAACTGA AAGAAGTAAT AGCTATAAGG CCACAGGGCT CACTACCCAA 420

55

	TATTCACAAA TACGTCTGGT AGCCCTATCG ACTCGAACAA AATTAGCCAC ATTATTAAAG	540
	GGGGGCGCTG ATATTAGTTC TATTAAGAAA CCTATAACGA CGCATACATT ACATCATTCTG	600
5	CATATATCTA CACTTGCTCA ATTAGGAATT AACTTAAAAG CAATGCAAGA GCATGTAGGT	660
	CATTGAGATT ATAAAArAAA TCTAGAGATA TACACACATG TTACTAATCA GATGGCGAAA	720
	GATATGATGA ATAAATTTGA ACGATTGGGG AGTTAAAATT GGAAAAAGAT GaTACACTAG	780
10	CAGAAATTAA GCCTATGCTC AATTTTGTATG AGCAAAATAGC AAAATTAAAA CAGATGAATA	840
	TATTTTTTAA TATTATTGAC ACCGAAAAAG CAAATGAAAT TCTTAGAAAA AATAATTACT	900
	TCTTCAAACt wGcTTATTTT CGaAAAAATT TCGrAAAAAA GaATGGCGGC TATTTTCATAG	960
15	AATTTGCTTA TTTATCAGAT TTAGCAACTA TAGATATGaA ATTAAGATAC ACAATGTTGC	1020
	ATTTAACTTT AGATATTGAA CATAGTTTAA AGTATCTAGT CTTAAAACTA ATAACAGAAA	1080
20	ATAACCAAGA AGATGGTTAT AAAATAATAG ATGAGTTCTT ATGTATTGaT AAATCATATA	1140
	GCAATTCAAA TTTTGACACA AATTCAAGAA CACCAGAAGA AGTTATGGAA ACCAAAATCA	1200
	AAAATAAAAA CGAAATATTC AAGCATATGA ATAAACGAGG ACAACTACCC GAGAAGTTGa	1260
25	ATAAATACTA TcmAAATCCA CCCGChnGGk TTTGCaTTGr ATTcATGCAA CTAGGTCAAT	1320
	TCGTTTCGTT TCTCAACTTC TATTACAAGA AGTACAATGA CGAAGAATTG AGAGTTGCTA	1380
	ATATTTTAAAT GCCTTTTAGTT AAAAATATAA GAAaCAATC AGCTCATAAC CAACCCATCA	1440
30	TAGCAAATCT AAATTATGAC AGTAGATTAC CTCAATATTT ATTTGAAAAA GGAATAATA	1500
	TAGGCATATC TAGAAACATG TTCGGAATAA AAAATTTcAT AGATACTkTC ksTACGCTAG	1560
	AATTACATAA TCAAGTTTGT AGTAATGCAA TTATCCAAGC AAGATATCAC GATTGGACC	1620
35	AACTTCAAAA GCGATATAAA AGrAACGrAA GCTATTATAA TAATGCATTA GCTATCAAAA	1680
	GATTfTTTTAT AGCTTTTAGAT AAAATTATTG ACTTCAACAG ACCAAAAGTA TAACTATCT	1740
	AGTGAGGAAA GAGACTTATA GGTCTCGCGA GTTATTTTAA TTCGTATGCA AGAAAAAGAA	1800
40	GAGCTATGCA TTTTATTTAA AATGCGTAGT TCTTcTTTTA TGCATCTAAA TTCATATTAT	1860
	TTTTGCAATA TAAACATATC TTTGTGCAAA TTCCGAACAC AAAACATTCA CATCATCCTT	1920
	TTTTGCCCTT TTTCTATACC CCAAAACACA AAAAGCCCCG TAAGCCTATG CcTACGGGgT	1980
45	TTGACAATAA ATTATATATT ATTGTTCTTC TTTAACATAT GGTAATAATG CCATATGACG	2040
	AGAACGTTTG ATAGCTGTAG TCAaCATACG TTGATATTTA GCTGAAGTAC CAGTTACACG	2100
50	ACGTGGTAAA ATTTTACCGC GTTCTGAGAT AAAACGTTTT AATAATTCAG TGTCTTTGTA	2160
	GTCGATATGT GTAATACCAT TTGCTGTGAA ATAGCATACT TTTTACGAC GACGTCCGCC	2220

55

CGTTAATTTT TATTAGAATG GTAAGTCATC ATCACTTATA TCAATCGGTC CGTTTGCATT 2340
 TGCAAATGGA TTATCAGATT GTTTCGTGTT TGATGAATTA TTGTACGAAT TGTTTTGTCC 2400
 5 TGATTGTTGA CCACCGAATC CTTGACCGTA ATCTTGAAT TCATTTTGTT GACGTTGGCC 2460
 ACCATTTTGT TCGCATTTT TAGGTTCAAG GAATTGAACG CTATCACACA CAACTTCAGT 2520
 AACAAACACA CGACGACCTT CTTGATTTTC ATAATTACGG GATTGTAAGC GACCATCTAC 2580
 10 ACCAGCTAAA CTACCTTTAG ATAAATAGTT ATTTACATTA TCTGCTTGTC TTCTAAAAAC 2640
 AACACAGTTA ATAAAActG ctTCGCGCTC CCCTTGAGCA TTCGTGAACG TACGATTTAC 2700
 TGCAAGAGTG AATGtCGCTA CACTCACACC TGAGGGAGTG GTTCTGTATT CCGGATCTTT 2760
 15 CGTTAAACGA CCTACTAATA CAACTCTATT TAGCATTTAA ACGCCCCCTC TAATTATTAC 2820
 TTGTCTTCGT CTTACGAAT AACCATGTAA CGAATGATAT CGTCACTGAT TTTAGCTAGA 2880
 CGTTGGAATT CGTCAGTAGC TTTGTTGTTA TCAGATTTAA CACGTACGAT GTTGTAGAAG 2940
 20 CCATCTTTGA AATCATTGAT TTCATAAGCT AGGCGACGTT TACCCAGTC TTTTGCTTCT 3000
 AAAACTTCTG CACCTTCAGT AGCTAAGATA CCGTTGAAAC GTTCAACTAA CGCTTTTTTA 3060
 25 GCATCTTCCT CAATGTTTGG GCGTACGATG TACATAACTT CATATGTTCT CATTTTATAT 3120
 TTGCACCTCC TTGTGGTCTA TACGGCTTAT CAATCTTAAA ACAGATAAGC AAGGaATAAT 3180
 TTTCACTACT CACAATAAAG AATTaTATCA TGCGCCATTA CTTTTTACAA TaATAATTca 3240
 30 AACTACTCTT CATATCATTT TTGATATtAA TTCATTTGaA ACTTTChATG ATATTTThAA 3300
 AAATACACTT CACAAAAGCG AACATATGTn CTATAAnAGT TGTGAGGTGG TAAGGAATGA 3360
 ATTTA 3365

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1032 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

GTTAAAGTA ATTGGTGGTA TTGATGATGA TTTTACAGCC AATGTTATGC ATCCAAATCA 60
 ATATCGAATT CGATATTCGT CTCAGAAACA GGACCTTAAT GAAGATATGA CAGTTTTTGA 120
 50 TGCAGTATTA AGTTCTGATA CAACAACCTT ACGCATCATC AAGCAATATG AGCAGGCAGT 180
 ACAAGCTTAT GCGGATGACC AAAGTGATAA ATTGTTCAAG CGAATGATGG ATGCGCAAGA 240

ACTAGGTATA CATGATACTA CTAAATACAT TAAAGAATTA TCCGGCGGAC AACAAAAACG 360
 TGTGTACTT GCTAAAACAT TAATAGAACA ACCAGATTTA TTGTTATTAG ATGAACCTAC 420
 5 GAACCATTTA GACTTCGAAT CAATCAGCTG GTTGATCAAT TATGTGAAGC AATATCCTCA 480
 TACTGTTTTA TTCGTAACCC ATGATCGATA TTTTTTAAAT GAAGTTTCCA CTAGAATTAT 540
 TGAACTAAAC AGAGGTAAGT TAGCGTCATA TCCTGGTAAC TATGAATCTT ATATTGAAAT 600
 10 GCGCGCTGAA AGAGAAGTAA CACTTCAAAA GCAACAACAA AAGCAACGAG CTTTATATAA 660
 GGAAGAACTT GCTTGGATGA GGGCTGGgAG CTaaggCTCG TACTACAAAG CAACAAGCTA 720
 GAATTAATCG ATTTAATGAC CTAGaAATG AAGTTaACCA GCAATATAAA GACGATAAAG 780
 15 GTGAATTGAA TCTTGCTTAT TCaAGATTAG GTAAGCAAGT GTTCGAATTA GAAGACTTAT 840
 CAAAGGCTAT TAATGATAAA GTATTATTG AACATCTGAC GGAAATTATT CAAAaGGTG 900
 20 AGCGTATTGG TGTTGTTGGG CCAATGGAG CTGGTAAAC AACACTCTTA AATATTTTGA 960
 GTGGAGAAGA CCAACAATTC GAAGGTAAAT TGAAGACTGG GCAGACGGTT AAAGTAGCTT 1020
 ATTTTAAGCA AA 1032

25 (2) INFORMATION FOR SEQ ID NO: 248:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 852 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

35 TGTGATTAC GAAGCTTATT TTCGTACACC TTCAACAACT GATTACAACG GCGTTTATCA 60
 AGGTATTAT ATTGATTTTG AAGCAAAGGA AACTAAAAAC AAGACGTCCT TTCCTTTAAA 120
 TAATATTCAT GACCATCAAG TCGAACATAT GAAAAATGCA TATCAACAAA AAGGTATTGT 180
 40 GTTTTTAATG ATTCGTTTTA AAACGCTAGA TGAAGTTTAT CTTTACCCT ATTCAAAATT 240
 CGAAGTATTT TGGAAGAGAT ATAAAGATAA TATTAAAAAG TCTATAACAG TTGATGAAAT 300
 ACGAAAAAAT GGTTACCATA TTCCTTATCA GTATCAACCA AGATTAGACT ATCTAAAAGC 360
 45 AGTTGATAAG TTGATATTAG ATGAAAGTGA GGACCGGTA TGACGGA AAA CAAAGGATCT 420
 TCTCAGCCTA AGAAAAACGG TAATAATGGT GGGAAATCCA ACTCAAAAAA GAATAGAAAT 480
 50 GTGAAGAGAA CGATTATTAA GATTATTGGC TTCATGATTA TTGCATTTTT CGTGTCTTCTT 540
 TTACTAGGTA TCTTATTGTT TGCTTATTAT GCTTGGAAAG CACCTGCTTT TACCGAAGCT 600

55

TTAGATAATG GCCAAGACA TGAGCATGTA AATTTAAAG ACGTGCCGAA ATCAATGAAA 720
 GACGCAGTAC TTGCAACTGA AGACAATCGT TTCTACGAAC ATGGCGCACT TGATTATAAA 780
 5 CGTTTATTTCG GTGCAATTGG TAAGAACTTG ACTGGTGGAT TTGGkTCtGA AGGtGCCTCA 840
 ACATTAACAC AA 852

(2) INFORMATION FOR SEQ ID NO: 249:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5804 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

CACTTTTTTC ATTAAAAATC TCATATTTAT ACACTGAACG TAATCTCGAA TATTTTTCAA 60
 CCCAAGTTTT AACTTTAACT TTTTCTGGAT AAAAAATAGA CTTTTTATAA TTGACATTGA 120
 GGTCACTCAC AGGTGAAATG ATTCCTTGTT TTTCCATATC AGCATAACTA AAACCTAACT 180
 25 TCGATATATA ATCCAACCGC GCAACTTCAA ACCAAGTTGC ATAATTCCCG TGATAAATTA 240
 CACCCATCTT ATCAGTTTCA GCATAACGCG CTTCTATTTT TGTAATACTA TATATCATTT 300
 TAAGCCTTCT TTCAGTTTAA CTTTATATCT CATTCTAACA TAAAATACAA GAAGAGGCCG 360
 30 GCCAAGAACA CAAAGGkTTT GAACCGACCT ATTATATCAT AAaGTTTATA GAAGTATTTT 420
 TGAGCACTAT CAAAGTGCCT CAAATACCGA TTAAAATTTT ACTGTGATAT CTATTTTTTA 480
 TTGCGCTAAT TTATTTCTTA AAACCATTG TAAAATTCCA CCGTGACGAT AGTAATCCAT 540
 35 TTCAACAAGT GAGTCAAAAC GAACCATAGC GTCAAATTCT ACCAAATCAC CATCTTGCTT 600
 CTTAGCAGTA ACTTTGACGT AGtCATGTGG TTGAACATT TCATCAATAT TAACAGtAAT 660
 TTCTTCTGTA CCATCTAGAC CAAGAGAATC AGCTGATTCA CCTTTTTTAA ACTCTAATGG 720
 40 TAATACACCC ATCATAACTA AATTTGAACG ATGGATACGT TCATAACTTT GTGCAATAAC 780
 TGTTTTAACA CCTAATAAGT TTGTACCTTT TGCTGCCAG TCACGAGATG AACCCTATACC 840
 45 ATAATCGTTA CCAGCTAATA CAACTAAACC TGTACCATCT TCTTTATATT TCATTGCAGC 900
 ATCAAAGATA GGCATTACTT CATTTGTTGG CCAATAAGTT GTAAAACCAC CTTCAGTACC 960
 TGGCGCTAAT TGGTTTTTAA TACGTATATT AGCAAACGTA CCTCGAACCA TTACTTCGTG 1020
 50 ATTACCACGT CTGAACCAT ATGAATTAAA TTCACGAATA GGCATTGAT GATCTTGTA 1080
 ATATTTACCA GCTGGCGTAT CTTTACCAAT TGCACCTGCT GGAGAGATGT GGTCACTTGT 1140

	TTCTTTAGAT AATCCTTGGA AGAATGATGG ATTTTGAATG TATGTTGAAT TAGGATCAAA	1260
	GTCATATAGA GGTGATCAG TTACATCAAT CTCATTCCAT AATTCGTTGT TATTGTATAC	1320
5	GTTATTATAT TCTTCAATAA ATAATTCAGG TGTTACAACA CTATCAACGG TATCTGAAAC	1380
	TTCTTTAATT GATGGCCAAA TATCTTTCAA ATATACATCT TCACCGTCAT TACCTTTACC	1440
10	AATAGGTTCA TTTTGTAAT CAATATCAAC CGTTCAGCT AATGCATAAG CAACAATAA	1500
	CTGTGGTGAA GCTAGGTAAT TGGCTTTAAC AAGAGGATGG ATACGACCTT CAAAGTTACG	1560
	GTTACCAGAT AATACAGATG TCACTAATAG GTCCTCATCA GCAATCGCTT TTTCAATTTT	1620
15	TGGTAATAAA GGACCTGAAT TACCGATACA AGTTGTACAT CCATAACCAA CCAAGTTGAA	1680
	GCCTAAATCA TCTAAATAAG GTTGTAAGCC AGCATCTCTT AAATATCCGG TAACAACCTT	1740
	TGATCCTGGT GCTAGAGAAG TTTTAACGTA TTCAGGAAGT TTCAAGCCTT TTTCAACTGC	1800
20	TTTTTTAGCA ACTAAACCTG CACCTAACAT TACATAAGGG TTAGATGTAT TTGTACATGA	1860
	TGTAATTGCT GCTATTGCAA TATCACCTGT TTTCAATTGA GCTTTTGATC CATCTTTAAA	1920
	GTTAATTTCA GCTTTCTTAT CAAATTCAGT TTTATCTAAA CCGTGTCTT GGTGCTGCTG	1980
25	TGGAGCTGTT ACAGAATTTT CAAATGATGA TTTATATCA CTTAAGAAA TTAAATCTTG	2040
	AGGACGTTTT GGTCTGAAA GCGATGCTTC AACTGTTGAT AAATCCAATT CGATAACATC	2100
	TGTATAATTA GGATCTTCTT TCTCAACATC AAAGAACATA TGGTTTTGTT TCAAATATTC	2160
30	TTTTACTAGC GCGATATGTT CGTCTGATCT ACCAGTTAAC TTCATATATT TAAGAGATTC	2220
	ATCATCAACT GGGAGAATC CGCAAGTTGC TCCATACTCT GGTGCCATGT TTGCAATTGT	2280
	AGCACGGTCT GCTAGTGGTA AATGTTGTAC ACCTGGACCA AAGAAGTCCA CAAATTTACC	2340
35	AACAACACCT TTTTACGTA GCTCTTGAGT TACTCTTAAC GCTAAATCAG TTGCTGTTGC	2400
	GCCTTGTTGT AATGAATTTA CTAGTCGTAC ACCAATAACC TCTGGAATTG GGAAATAAGA	2460
40	AGGTTGTCCA AGCATTCCAG CTTAGCTTC AATACCACCA ACACCCATC CTAGTACGCC	2520
	AATACCATTT ATCATTGTTG TATGTGAATC AGTACCAACT AATGTATCTG GAAATGCAGT	2580
	TTTTTCACCA TCTACATCAC GAACATGTAC AACACTTGCT AAATATTCTA AGTTAACTTG	2640
45	GTGAACTATT CCAGTTGCAG GAGGAACTGC ATTGTAATTA TCAAATGCTT TCGTTGCCCA	2700
	ATTTAAAAAC TGATAACGTT CATAGTTACG TTCAAATCTT AATTTTCATAT TACGTTCAAG	2760
	AGCTTCTGGA TTGTCATAGC TATCCACTTG AACTGAGTGG TCAATAACTA AATCCACCGG	2820
50	TACTTCTGGA TTAATTTTAG TAATATCTCC CCCAACGTCA TCCATTGCTT TACGTAAAGA	2880
	AGCTAAATCA ACTACGGCTG GTACACCTGT GAAATCTTGT AAAATAACAC GAGAAGGTTT	2940

55

	GTCTGTAATT	ACAAAATCAT	CTTCTTGACG	AAGTAAAGAT	TCTAACAAAA	CACGAATTGA	3060
	ATAAGGTAAA	TTGGAAACTT	TAGTAATACC	TTGCTCTTCT	ACAGCTTTTA	AATCATAGTA	3120
5	AGTATAACTT	TGGCCATTCA	AGTCAAAATG	TTTTTTTGAT	TGCTCTTTAA	AATTTGCAGC	3180
	CATTTAATGA	TCCCCCTTGA	TACATTTTTA	TATTTATATG	CCTTGATTAA	ATTGTATTAT	3240
	TATATTTATT	GATAAACAAAC	TCATCATGCT	TAGAAAACGC	TTAATTTAGG	TTTTGACTTT	3300
10	TTAATCAGAG	TATATAAGCA	AAACTTATCA	TACAGGTAAG	GTGTAATAAG	TATTTTTTAT	3360
	TAATTGAGAA	TAATTATCAA	TTTCGCGAAT	GATTCAATTC	AATTTTTTAA	CGTATTATTT	3420
	CATTGAGCAG	AAAGAAAATT	ATGGCACCAA	ACTTTAATAT	TTTTTTCAAT	GTCATTCTTT	3480
15	TGATGGGAGT	GGGACAGAAA	TGATATTTTC	GCAAAATTTA	TTTCGTCGTC	CCACCCCAAC	3540
	TTGCATTGTC	TGTAGAAATT	GGGAATCCAA	TTTCTCTTTG	TTGGGGCCCA	TCCCCAACTT	3600
	GCACATTATT	GTAAGCTGAC	TTTTCGTCAG	CTTCTGTGTT	GGGGCCCTCA	CCCCAACTCG	3660
20	CATTGCCTGT	AGAATTTCTT	TTCGAAATTC	TCTGTGTTGG	GGCCCTGAC	TAGAATTGAA	3720
	AAAAGCTTGT	TACAAGCGCA	TTTTCGTTCA	GTCAACTACT	GCCAATATAA	CTTCGTAGAG	3780
25	CATAGAATAT	TGATTTATGT	CCCAGCCTGA	GTTAATTTTC	TATAAAAGTA	TATTTAATTT	3840
	GCGTTTATAC	CGTCAAACCT	CACTTTAGCT	TTGTCAAACC	CCTTTCTATT	AAGTTTTCAG	3900
	AAATAAACCT	ATCTTAAAT	ATAAAAAAAT	CGAGAATTCG	TAGTTTAATA	ACGAAATTCT	3960
30	CGTCTTATC	CTTTTGAATA	TACTCAATTT	TCCACAAAAA	CAAACAAGTA	GTATATCTGT	4020
	TCTAGCTACT	AGAATGACAT	ACTACTTGTT	ATTAAATAC	TTAACTAAAC	TTTATTAGTT	4080
	ATCTTTTTTC	TCTATATTTT	TACGTGACTG	ACGCTTTTCA	AGAATGTCAG	ATTCATAATC	4140
35	TTCTTGTTGA	CTCTTGATAT	ATTCTTGTA	GCGATGTTTA	TTCGGAGTCA	ATGTTAAACC	4200
	TAGGAAATTTA	CGTTCCTGGT	TCGCATCCTT	GTAAGAACTT	ACCATCATGA	GTATGACGAC	4260
	AAAGGAGAAT	GGGAATGCAC	TTATAATTGC	AGCACTTTGA	ATCGCAATTA	AAGCTTCAGC	4320
40	GCCGTTACCG	CCACCAGCTA	ATAAAAGTAC	AAATGCTATT	AAGGCCTGTG	AAATTCCCCA	4380
	AACAACTTTT	ACCATACTAG	ATGGATTTAA	TGAACCAAAT	GTTGTTTGCA	TTCTTAATAC	4440
	AAATGTTGCT	GAGTCAGCAG	ATGTAATAAA	GAATGATGCA	ATTAATAATA	ATGCAATCAA	4500
45	CGATAAAACA	ATGCCAAATG	GCACATGATT	AAACTCTCCA	AATAGCTGTG	TTTCAGGAGT	4560
	CATATCAAAA	ATTTCTTTGT	GTTTCTTACC	TGTCTCGATG	CCTAATACAC	CAAAGACACT	4620
50	AAACCAAACA	AACTAACAA	TTGCTGGAAC	TAGCAAGACA	CCAGAAATGA	ACTCTCTAAT	4680
	TGAACGTCCT	TTTGAAACTC	GTGCAATAAA	CACTCCAACG	AATGGACTCC	AACTTAACCA	4740

55

TGCTGTATCA AAACCTATTAA ACAAGAATGT GTTTAGTAAA CTACCCGTAG AGCTAGTTAA 4860
 CATATTTAAA ATAAGAACAG TTGGTCCAAC AATTAAAGCA GCTACCATTA AAATAGTACC 4920
 5 TAAACCAATG TTCAAGTTAC TTAAGTATTG AATACCTTTA CTTAATCCAG ACCATGCACT 4980
 TGCTATAAAT AAGATAGTAA CAACAATGAT GATAATCGCT TGTACAAACG TATTGTTTGG 5040
 AACATTGAAC AAGTAATGTA AACCACCATT AATTTGTAGA GCACCCATAC CTAACGAAAC 5100
 10 GGCTACCCCA ACGATTGTCG CAAATACAGA TAAAACGTCA ATAAAAATCC CAATAGGACC 5160
 TTCTACTTTA TCACCTAAAA GAGGACGTAA AGTTCTAGAT AATAAACCTG GTTCACCTTT 5220
 ACGGAATTGC GAATATGCCA ACGTAACGCA ACAACACCAT AAACAGCCCA AGCATGGAAT 5280
 15 CCCCATGGA AAAATGTTGA ACGTAGAGCT TCAGTATAAG CTTCAGTAGT TTTGGGATCT 5340
 GCTGTAGGTG GCGTAGCAAA GTGCGCCATC GGTTCAGCTG CACCATAAAA CACCAAACCT 5400
 ATCCCCATAC CAGCACTAAA CAACATAGCA AACCATGAAA TTGTATTAAA CTCAGGTTTG 5460
 20 TCATTTGGTT TACCTAGTTT AAGTTTTCCA ATAGGACTAA AAATAAGGAA TATACAGAAG 5520
 AACACGATAA TCGTAGTAAG AATAAGATAA TACCAACCTA ACTTTTCTGT AATCCACATT 5580
 25 TTAATATTAT TGGTAACATA GTTGAATTGT TCAGGTAAAA ATGCACCAAG TAATACGACT 5640
 ATAGCAACAA CAATTGCACT ATAGATGAAG ACTGGTGAAT ACTTCTTTCC ATTTGGATTG 5700
 TCTGGTGAAG AAGAATTCAT AATTAATTAC TCCCTTCAAT TCTATATTTA ATTTTATGTA 5760
 30 GTAGAATAAA AATATTATCT AAACATTTTA TTCAATAACT CACG 5804

(2) INFORMATION FOR SEQ ID NO: 250:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

ACCCGCGAAT ATGGTCCATC CTATCGATTT ATTTTAACT GGTTCGACAA TATTTAATTT 60
 TTCATAATCA TTCTTAGTGA TTTTGACATA TGTTTTCGGT ATGAGCCAGT TAATAAATGG 120
 45 AAAGAAGAAG ACAATCCAAT TACTTGCCAA ATCaATCATT AAATATTCAC TATCGTATTT 180
 GATTATTCGA TATTTAGGGT TTTTATTAAT AACTTTAGAT TCGCAAAGCA ATGTCTCCAC 240
 50 ATCCCTTTAA TTTTATGTGT AATACATTTT TCGATACTTC AAAAGACATT CAAATACTAT 300
 CAAGTTACTG TCATCAAAGG TTTTATTAAC TGATATTtTC ATATTTTTAA TCTGAATTTA 360

(2) INFORMATION FOR SEQ ID NO: 251:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 964 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

```

CCAGGGTGCG GAAAGCTTTA AAATTTGGAC TAATAAAGAT GCTGATATTA ATTCTATGAA      60
AACAGCAGTT TTACAACAAT TAAAGGAGA ATAACATATG CTTACTGGCA AACAAAAAG      120
ATACTTAAGA AGTTTAGCAC ACAATATTGA TCCGATTTTT CAAATTGGAA AAGGCGGTAT      180
CAACGAAAAT ATGATTAAAC AAATAGATGA TACGTTAGAA AACAGAGAAT TGATTAAAGT      240
ACATGTACTA CAAAATAACT TTGATGATAA AAAAGAATTA GCTGAAACAT TAAGCGAAGC      300
TACTCATAGT GAATTAGTGC AAGTGATTGG ATCTATGATA GTGATTTATA GAGAATCTAA      360
AGATAATAAA GAAATTGAAT TGCCATAATA ATGAAAAAGA TATyACTTTA CGGCGGTCAG      420
TTTAACCCTA TCCATACTGC ACATATGATA GTAGCTAGCG AAGTATTTCA TGAATTACAG      480
CCAGATGAAT TTTATTTTTT ACCTAGTTTT ATGTCTCCAT TGAAAAAGCA CCATGATTTT      540
ATAGACGTTT AGCAGAGATT AACATGATA CAGATGATTA TCGACGAGCT TGGTTTTGGA      600
GATATTTGTG ACGATGAAAT TAAACGTGGT GGTCAAAGTT ATACCTATGA CACGATCAAG      660
GCATTCAAGG AGCAACACAA AGACAGTGAG TTGTACTTTG TTATTGGGAC GGATCAGTAT      720
AACCAACTAG AGAAATGGTA TCAAATTGAA TACTTAAAG AAATGGTTAC TTTTGTAGTT      780
GTAAATCGAG AAAAAAATAG TCAAATGTT GAAATGCTA TGATTGCAAT TCAGATACCT      840
AGGGTAGATA TAAGTTCGAC AATGATTCGA CAAAGAGTTA GTGAAGGGAA ATCTATCCAA      900
GTTCTTGTTT CTAAATCCGT TGAAACTAT ATTAAGGGGG AAGGATTATA TGAACATTGA      960
AAAA
  
```

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1193 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:

TTGTCATGAA ATAAATGGGA TGAATATCAC GACTAGAAGT AATGTTACGA ACAGGAGCGT 120
 ATAACTAGA GACGCTAAAT TCGACATAGT ATGThGCTCA ATTATGGCTG ATGATGAATT 180
 5 TAAAGTATGT GCGTTGGAAC TGTGGGATTT TTGTTTCATAA TGTTTTGCAT ATTGCGCCAT 240
 GATGAATAGT GTAAAAATAA ATAAACAAT AAGAGATATA ATGCCCATAA TCAAAAGTAT 300
 TTGTTTAGAG CCTTTCATTA TTTCACATCC TTTCTAAAAT ATATTGTAA CTAAATTTAA 360
 10 AATAGTTATT TTGTAAATTC TAAACCTTTT TCATCGCGAA AACAATTAAA TAGGTCGCGG 420
 TATTAATTAT TATATTATTA CCGCTTAATA TGAAAAATAC ATGAAaATTA ATTTTCTAAT 480
 ATACTTTTGA AAAATTATTA CAAATTAGCC CCTTCAAAC GCGAAAACAT AAGGATTCTA 540
 15 GTTCAAAG GCGTGATAAG CATAAAATGA AATGTAATAT TTCGATGTAT AAAATTTTTA 600
 ATTAGCTAAA AATCATCGCA TTAATTTTTT GAGCTACATC ATCAAAATTC GGACATTTTA 660
 20 ACGACACATA TAATTTAATT TTAGGTTTCAG TACCAGAAGG ACGTAAAGCG ATAAATCCTT 720
 CGTCAAATAA GACACGAATA ACATTTGATT TAGGAGAATT AATCTGCGAC GTTGTATCTT 780
 TATCCAAATG ATAAACCTCG CTAGTTAAAT AATCTTCAAT TGCTTTCACT TTGAGTCCTT 840
 25 GAATCTCTTG CCGTGGATTT GAACGGAATT TGGTCATTAT TGCATTAATT TTCTTTTTCC 900
 CTCAAATCC TTCTAGCGTA TGCGgAATAA TGTATCTCA TGTCTACCAA CAGTTTGATA 960
 AATCTGTTCT AATTCATCTT TCAATGTTTT GCCATATAAT TTTAACTCAG AAGCGTATTT 1020
 30 TATAATGAGT GGCACAATTT GTACGGCATC TTTATCACGT ACAAAGGCT CTGATAGAAA 1080
 ACCGTAATC TCTTCAAATG CGAAAATCAT ATTTGATGAT CATCCAGTTG TCTTATTTCC 1140
 TGAGCAATAA ATTTAAGCCC GTCAGCACCT CTTTGGTATT CAACATTATT ATA 1193

(2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

TATCAGGATG ATTATGnCGG GGTTTTATTA AGTCTGAGCG TATATnCCTT TTATTCTCCT 60
 AAATGGTTTG GAATTCAAGG AGAAGACGAT GAAATGGTAT CCAAAAACCA AnGAATATAT 120
 50 nGCATTAAGT CTGAGGATGA TAGTGGGTG GCAATTCGTT CATTAATTTT GCATAAAGAT 180
 GAACCTATGT ATTTAAAAAA ACGTACATGT GTACCTACTT TGTTAATTAA TGGGGAACAT 240

AAAAAAATCT TCGAACATTG AGGACATGCA CCGCATATTG AAGAACCAGA AGCATTATTG 360
 AATTATTATT TAAAATTTTT AAAAAGCGTA TCATAATATG TGATATATAA ACCTAGGGCA 420
 5 TAAAGTCCTT AGGCAATGTG AAAAAGCTGA TTAATATTCA TTATTTGATA GAAATCAGCT 480
 TTTTTTGAAA TGTATTGAT ATATACTGCT CGTTATGCGG CTATCTTCCT TATATTAAGT 540
 GCCATTAGTG CAAAACCTCT TAACAATTAG GTAAAAAGAG CATAAAAAAA GGAAGTTTAA 600
 10 TAGAATGTAT CATCTATCAA ACTTCACCAA ATTGCGCTAA ACAAATTAT AGTTCAATTT 660
 CGTTGTTTGC TTCAGTGATT CGTTTATTTA CTCGACTCAA TAATGATTG ATTTTTTTAC 720
 GTTGTGTGTC ATTAACAAGA ATTAATACAG TTCTTTCATC ATGCTCATT CGTTTTTTAT 780
 15 CGAAGTAATC TTCTTGAGAT AAAATTTTAA CTGCTTTAAC AACTTGTGGT TGTTTGTAGT 840
 TTAAATGATT AATAATATCT TTAAGATAGT ATTCTTCTC TTTGTTTTCG CTGATGTATG 900
 20 TCAATACAGC GAATTCCTCA AAGCTAATTG AATAATCCTT TTTAATTAAA CTTTTTAATT 960
 TGTCAGCATA AGTGACCATT GATAACAAC CAAAGCAATC ATTGATTTTT GTAATTGCCA 1020
 TGTTTAAAC CTCCCTATTT GATGCATCTT GCTCGATACA TTTGCCCCGA TAATATATG 1080
 25 TATCTAATCT TTATGAT 1098

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 2881 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

CCAGGTAAAA TTGTGCAATC ATTTGACGCA TTAATGGACG CCTTGGACAA TGAAGATTAT 60
 GAAGGAGAAA AAGTCATTCC ATTCTTAGAT AAACATTTTA AATATCAAGA TGGCCGATCA 120
 40 AGTGAGCGTT TAGTCAGAAA TTTATTTGGT AGCTAAGTTT ATATAGTAGT CAAAGTGGGA 180
 GAGGTATAAT GATGAAATTT TCAGTAATAG TTCCAACAT CAATTCAGAA AAGTATATAA 240
 45 CAGAACTACT TAATAGCCTT GCGAAACAAG ATTTTCCGAA AACTGAATTT GAAGTGTTG 300
 TAGTTGATGA CTGTTCAACA GATCAAACGT TACAAATAGT TGAAAAGTAT CGCAATAAAT 360
 TGAACCTGAA AGTAAGTCAA CTCGAAACAA ATTCTGGTGG TCCAGGTAAA CCTAGAAATG 420
 50 KGGCGTTAAA ACAAGCAGAA GGTGAATTTG TATTATTTGT GGAATCCGAT GACTATATAA 480
 ACAAAGAGAC TTAAAGGAT GCAGCAGCAT TTATTGATGA ACATCACTCA GATGCTTAT 540

55

	CACCTGAAGT TACTTTGTGA AATTCAAGAA TTATCTATAC TTAAAGCCCG ACTAAAATCT	660
	ATAGAACAGC ATTACTAAAA GATAATGACA TTTATTTTCC AGAAGAATTA AAGAGTGCAG	720
5	AAGATCAATT ATTTACAATG AAAGCATATT TAAATGCAAA TCGAATCAGT GTGTAAAGTG	780
	ATAAAGCGTA TTATTATGCT ACAAAGCGTG AAGGTGAACA TATGAGTAGT GCGTATGTTT	840
	CACCTGAAGA CTTTTATGAA GTCATGAGAT TGATTGCTGT AGAAATATTA AATGCAGATT	900
10	TAGAAGAAGC CCATAAAAAAT CAAATCTTAG CAGAATTTTT AAATCGTCAT TTTAGTTTTT	960
	CTCGTACGAA TGGCTTCTCA CTTAAAGTTA AACTAGAAGA TCAACCACAA TGGATTAATG	1020
	CTCTAGGAGA CTTTATACAA GCAGTTCCAG AACGTGTAGA TGCATTGGTG ATGAGTAAAT	1080
15	TACGACCATT GTTGCACTAC GCGAGAGCGA AAGATATAGA CAACTATAGA ACTGTGGAAG	1140
	AAAGTTACCG TCAAGGTCAA TACTACCGTT TTGATATTGT AGATGGTAAA TTAAACATTG	1200
20	AATTCAATGA AGGCGAACCA TACTTTAAAG GCATTGATAT CGCTAAGCCA AAAGTGAAAA	1260
	TGACAGCATT TAAATTTGAT AATCATAAAA TTGTTACAGA GCTAACGTTA AATGAATTTA	1320
	TGATTGGCGA AGGACATTAT GATGTCAGAC TTAAATTACA TTCACGAAAC AAGAAGCACA	1380
25	CAATGTATGT ACCTTTAAGT GTCAATGCGA ATAAACAATA TCGTTTTAAC ATTATGTTAG	1440
	AAGATATTAA AGCGTATTTA CCTAAAGAAA AAATTTGGGA TGTTTTCTTA GAAGTCCAAA	1500
	TAGGTACGGA AGTATTTGAA GTGCGTGTG GTAATCAACG TAATAAATAT GCATATACTG	1560
30	CAGAAACAAG TGCATTAATT CATTGAATA ATGATTTTTA TAGATTAACA CCGTATTTC	1620
	CAAAAGACTT TAATAACATT TCGTTATACT TTACAGCTAT TACATTAACG GATTCAATCT	1680
	CATTGAAGTT AAAAGGTAAA AACAAAATCA TTTTAACTGG TCTGGATCGT GGTATGTAT	1740
35	TTGAAGAAGG TATGGCTAGT GTCGTACTAA AAGACGACAT GGTGATGGGA ATGTTAAGCC	1800
	AAACATCAGA AACGAAGTG CAAATCTTA CTTAGCAAAG ATATTAAAA GCGAGACTTC	1860
	AAAAATATTG TTAAGTTAAA CACTGCACAT ATCACTTATC CACTAAATAA ATAATAAATG	1920
40	CCCTCAAATC ATTGTGAGCC AACATGATTT GAGGGCTTTA TTTTGCTGTT TATGACATGA	1980
	TTATGACATT TCCCTGATTT TCATTTTCAT ATACATTAAA TTGTATACAC TGGAAATGAG	2040
45	GAGGTTATCT ATAATGATAA ATAAAAATGA CATAGTAGCA GATGTAGTAA CTGATTATCC	2100
	GAAAGCAGCG GATATTTTTA GAAGTGTGGG AATAGATTTT TGTTGTGGCG GACAAGTAAG	2160
	TATAGAAGCA GCAGCCTTAG AAAAGAAAAA TGTAGATTTG AACGAATTAT TACAGCGTCT	2220
50	CAACGACGTT GAACAAACGA ATACACCAGG TTCGTTAAAT CCTAAATTTT TAAATGTTTC	2280
	ATCACTTATT CAATATATTC AATCAGCATA TCATGAACCT CTAAGAGAAG AATTAAAAA	2340

55

TGAGTTAAAA GAAACATACG ATACATTTAA AAATGGCATG TTAGAGCATA TGCAAAAAGA 2460
 AGACGATGTC GATTTTCCAA AACTCATTAA ATATGAGCAA GGTGAGGTAG TAGACGATAT 2520
 5 TAATACTGTG ATAGATGATT TAGTTTCAGA CCACATTGCA ACGGGAGAAT TGTTAGTAAA 2580
 AATGAGCGAA TTAACATCTA GTTATGAACC TCCGATAGAA GCGTGTGGTA CTTGGCGACT 2640
 TGTTTATCAG AGATTAAAAG CACTTGAAGT GTTAACACAT GAACACGTAC ATTTAGAGAA 2700
 10 TCACGTATTA TTTAAAAAAG TATCATAAAT AACGCGATTAA GAAACTGTTG GCAAAAATAA 2760
 GTCCAGCAGT TTTTCGCTAT GTATAAAAGT CATAATAGTG ACATAAACAG CATTATTTGA 2820
 AAAGAAATAAT GGTCAACTTA GCATAAAAAT TGATATGAAT ATTTAATGGT ATAGATAATT 2880
 15 A 2881

(2) INFORMATION FOR SEQ ID NO: 255:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1056 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

ACCGTCGAAT ATCGCTTGTG ATTTACAATT TGTGTATTAA GATGCTCAAC TAATTTGGGT 60
 30 ACATATTCGG AATTTAGATT TGCAAGTACA ACAATTCCAT AATTTTGTTT TGGATTTAGT 120
 AAAATAAATG ATGAAAAGTT ATCTAGCGTT CCTGAATGAA AACTAAATG TTCATCATTAA 180
 TTGGTAAACC AGCCGGAAGC ATATGCATTG GCATTAGGTT CACCAATTGT TGAAGATAAA 240
 35 TTTTATGTG ATTGTTGAAC TAATGATTG TATTTATCAG GTGGATTAAG TTGGAATTTT 300
 ATCGAATGTT CCAAATCTTC AGTTGATGTC ATCATATATG CTGATGGTGT ATCCCAAAGG 360
 TTAAATTCAG GTTTAGAGAC GACAGGTGTC GAACCTTGTA ATTCATAGCC AATAGCATCA 420
 40 TGTTTTGATT TGTAATTGGT TTGTTTGAAT GATGTATGTG TCATATGCAA AGGCTTGAGC 480
 CATGAATTTG TAATATATTT TGTATAGGAT TGCTTCGTAA CGTTTGGAT AATTAAACCT 540
 AATAAATCAT AGTTCATATT TGAGTATTCA AATTCTTCTC CGGGCTTATG ATGTAATTCA 600
 45 TCACCCATAA TTGCATGGGT TACATCATT AAACGATTAT TTTTGCTTGT CACAGAATCT 660
 TCGCTTGTA TAACACTAGG TATACCACTT GTTTGAGCCA AAAGTTGCTT AATCGTAATA 720
 50 GTTTCATTTT GACCATTATA GTTCATTTTA AAATGAGGCA CATGTTTGA TACGGCATCA 780
 TTTAAGTTTA ATCGACCTTC TTGAGCTAAT TTTAAATTG CAAGACCTGT GAAAGCTTTC 840

55

TGATAACCAT AACCTTTATT TAAAAAACT TTGCCATTTT TTAATAyTAA AATTGATGCT 960
 CCAGGAATGT GTCCCTTTTG TAAATCATGC TCGATAATTG TATCTATTTG TTGTTGCGAA 1020
 5 TCATTGGTTA ACCGTGTCTT CGTATTGCTA TTTAAT 1056

(2) INFORMATION FOR SEQ ID NO: 256:

(i) SEQUENCE CHARACTERISTICS:
 10 (A) LENGTH: 1277 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

ATGCCaCACT TATTGGTGGC AGGATCGACG GGTAGTGGTA AATCTGTTTG TATAAATGGT 60
 20 ATTATTACAA GTATTTTATT AAATGCTAAG CCGCATGAAG TTAAACTTAT GTTAATCGAT 120
 CCGAAAATGG TTGAACTAAA TGTTTATAAC GGAATTCCAC ACTTATTAAT TCCGGTTGTT 180
 ACAATCCTC ATAAAGCTGC TCAAGCTTTA GAAAAAATTG TAGCTGAGAT GGAAAGACGT 240
 25 TATGATTTAT TCCAACATTC ATCAACTAGA AACATTAAAG GTTATAACGA ATTAATCCGT 300
 AAGCAAAATC AAGAATTAGA TGAGAAGCAA CCAGAATTAC CTTATATCGT TGTTATTGTA 360
 GATGAGCTTG CAGATTTAAT GATGGTAGCT GGTAAAGAAG TTGAAAATGC GATTCAACGT 420
 30 ATTACACAAA TGGCAGGTGC AGCAGGTATA CATTTAATTG TAGCGACACA AAGACCTTCT 480
 GTGGATGTAA TTACAGGTAT CATTA AAAAT AATATTCCAT CTAGAATAGC TTTTGCTGTG 540
 AGTTCTCAA A CAGATTCAAG AACTATTATT GGTACTGGCG GCGCAGAAAA GTKACTTGGT 600
 35 AAAGGTGACA TGTTATACGT TGGAAATGGT GACTCATCAC AAACACGTAT TCAAGGGGCG 660
 TTTTAAAGTG ACCAAGAGGT GCAAGATGTT GTAAATTATG TAGTAGAACA ACAACAGGCA 720
 AATTATGTAA AAGAAATGGA ACCAGATGCA CCAGTGGATA AATCGGAAAT GAAAAGTGAA 780
 40 GATGCTTTAT ATGaTGAAGC GTATTTGTTT GTTGTGaaC AACAAAAGGC aAGTACATCA 840
 TTGTTACAAC GCCAATTTaG AATTGGtTAT AATAGAGCAT CTAGGTTGAT GGATGATTTA 900
 GAACGCAATC AGGTAATCGG TCCACAAAAA GGAAGCAAGC CTAGACAAGT TTTAATAGAT 960
 45 CTTAATAATG ACGAGGTGTA AAAAAATGTC AGAAATGAAT GCGGTATATA ACGTTAAACA 1020
 ATaCATTTtA AATTTgATTA AGCAAAATAA ATTGGAATAT GGTGACCAAC TTCCAAGTAA 1080
 50 TTTATCAATT GCCAGAGAAT TAAATGTAAA AACCAGCAT GTTTATGAAG CAATTCAGCa 1140
 TTGATTACTG AACAAATCAT TAAAGATaT TTGAAGAGGG CACAAGTGTT AAGTCACTGC 1200

55

GrtTTGAATG CGGAACT

1277

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3557 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

15	TACCGTCGTT TTATGyGTCA AATTTTACAG TAAATTTTGC TTCATCAAAA GAAATAACCT	60
	TTAACAAGTA TAGTAATTTT ACATTTTACA ATGTTACAAA ATATAATTTT TTATAATTAG	120
	TTAAATCAC TAAAACGCTT TTATACACTA TCAAAATCAGC ATTTATAAAA ATATGAACCG	180
20	ATATCCTAAA ATGTTAATAA TATTACAAGA TAATAACAA CCACACAAAG CTACTTATTT	240
	TTGATAATAT GGAAATCGTA ATATAAAACn AAAACTTAAT TTAATATATA AATTGTCTTA	300
	ATAATTTTTA AAAGTAGTAA AACATAATTT TAAGGAGGAG TCCCTTTGAA AAAATTAGCA	360
25	TTTGCAATAA CAGCAACATC TGGTGCAGCT GCATTTTAA CGCATCATGA TGCACAAGCT	420
	TCTACACAAC ATACAGTACA ATCTGGTGAA TCATTATGGA GTATTGCTCA AAAATACAAC	480
	ACTTCAGTAG AGAGTATTAA ACAAATAAC CAATTAGATA ACAACTTGGT ATCCCTGGT	540
30	CAAGTTATCT CAGTAGGTGG AAGTGATGCA CAAATACGT CAAACACTTC TCCACAAGCT	600
	GGTTCAGCAT CATCTCATAC TGTACAAGCT GGTGAATCAT TAAATATCAT TGCTAGCAGA	660
	TATGGTGTTT CAGTTGATCA ATTAATGGCA GCCAATAACT TACGTGGTTA TTTAATTATG	720
35	CCTAACCAAA CATTACAAAT TCCTAATGGT GGATCAGGTG GTACAACACC AACAGCTACA	780
	ACAGGTAGCA ATGGCAATGC ATCATCTTTT AATCACCAA ATTTATACAC TGCTGGTCAA	840
	TGTACATGGT ACGTATTTGA CCGTCGTGCT CAAGCTGGTA GTCCAATTAG CACATATTGG	900
40	TCAGACGCTA AGTATTGGGC TGGTAACGCA GCTAATGATG GTTACCAAGT AAACAACACA	960
	CCATCAGTTG GTTCAATTAT GCAAAGCACA CCTGGTCCAT ATGGTCATGT TGCTTATGTT	1020
	GAACGTGTCA ATGGTGATGG TAGTATCTTG ATTTCTGAAA TGAATTACAC ATATGGTCCA	1080
45	TACAATATGA ACTACCGTAC AATTCAGCT TCAGAAGTTT CTAGCTATGC ATTCATCCAT	1140
	TAATTAAATA AATTGTACTG ATATATACTA GCAATTCACA TCATGTGAGA TTGCTAGTTT	1200
50	TTTATTTTGT AAAAAAATTT TCATTTTGGT AAAAAAATT ATCTCACCTT TCCCTATCAT	1260
	ACATATTTAT ATTTTGTATG AATGGTAGTT AGGTAAAAAT TAACAACCTA CCTATTTGAT	1320

	ATTTAATTTG TTATACCAGT ATTTTACGCT TTTTCGTCTA CATATACAAA TTTATATTAA	1440
	ATAAAGCCCA ATACAATTTA GGTAAATTAA ACAAGTTGAT AACTATTTAA TTATTCCTTC	1500
5	ATTGAAGAAT ATAAACTATT AAATCATTAT TTTGCTCTTA CATATATTTT AATGACCTAA	1560
	CTGaTTATGT TCCATGGAAT ACATTTATAA TATAGCCTCC TAATTAAATAT GCyTTGTCTT	1620
	GGTCATTCTA CGTAAATTCT ATAAAATATG TTATCTACTT ACATAAATAT CTGrACTTCA	1680
10	ATACCACCAT ATGTTTGTGA TACTGAAGTT CAGTTTAGTT TTATTTTCAA TTAGAAAAAT	1740
	AAGTTAAGTA TATAGAATAG TAAACCTGCT AACAAATGCTG AAATAGGTAA TGTAAATCACC	1800
	CATGTAATGA TCATTTCGTTG CGCAGTGCTC CATTTTACAC CTTTAGCTCG GTTAGAAGCA	1860
15	CCAACACCTA AGATTGATGA TGACACAACG TGAGTTGTTG ATAATGGGAA ATGTAGCGAT	1920
	GATGCAACAA AAATTGTTAA TGCAGATGAT AAATCGGCCG CAGCACCATT TGCTGGACGT	1980
20	ATTTTCATAA TATTACCACC TACAGTTTTG ATAATTTTCC AGCCACCAAT TGCAGTACCA	2040
	AGCCCCATTG CTGTGCGACA GGCAAATTTT ACCCATAACT GTGGTTCAAC ACTGCCATCA	2100
	TTCTGTACAT TAGCGACAAT CAATGCCAAC GTAATAATAC CCATTGATTT TTGCGCATCA	2160
25	TTCGTACCGT GAGAGAATGA TTGTAACGCT GCTGTGAAAA TTTGGAAAAA TCTAAAGTTA	2220
	CGATTGCTC TTGTTAAATT TGCATTTTTA AAGATAACTT TAAAAATTGA ATACATCAAG	2280
	AAACCAACAC AAAATGCGAT AATCGGTGAA ACGATTAAATA CAATAATAAT TTTTGTGAAA	2340
30	CCTTGGAAT GTAACACTCC AAATGAGCCT TCAGATGCGA TTGCTGCACC CGCAATTGAA	2400
	CCTATAAGTG CATGTGAAGA CGAACTTGA ATTCCGTAAA ACCAAGTAGC TAAATTCCAA	2460
	ATAATAGCCG CAAGTATTGC AGCTAACACA ACAACTAATC CATTTTCCAA TTAAATGGA	2520
35	TCGACAATGT CTTTAGTAAT GGTGCCTGCA ACGCCCGTAA ATGTTAAAGC ACCTATAAAG	2580
	TTCAATCACTG CTGCCATTAA AATTGCCGTT TTAGGgTTAA CGCTCTAGTA GATACAGCAG	2640
	TAGCTACTGC ATTGGcTGTA TCATGGAaTC CcATTGATAA AGTCAAATAT CAGCGAGAAA	2700
40	ATAACTACAG CTATAGTGAC GATGATTATA TATGACATAA ATATATACTC CCCTTAGCTA	2760
	TTTTTCATAA TAATAGTTTC AAAATTATTT GCTACGATTT GACATTTATC AGCGATTTCT	2820
	TCCATGCTTT CATAAAATATC TTTTATTTTA ATTAAAGTGA TTGGATCTGT TTCGCTATTG	2880
45	AAAATATGTT TAATTGACTG TCTTAAAATA CCATCACAGT TTGTTTCAA TTCTTTAATA	2940
	TTAATTGAAT GAATACGCAT ATGTGATAAT TTTTATCGA CTAATAAGCC GACAGCAAGT	3000
50	TTCAATTCTG CAACTGCTTT TTGAATGTTA TCAACAACT CAGCCATATA TTCATCTGTG	3060
	TATTGATTG AATACATTTT AAACATgGCT GCCGTTTCTT CAATTGCATC TAAAACATCA	3120

	TTTAAATCAG TAATTACTTG ATGTACTAAt tCGcACCATG TGACTCATAA GTTTTAATGT	3240
	TGTCTGAGTA TGCTTTTAAA TCTAAATGTG TATTGAAATC CATTTTACCG AATTCAATAG	3300
5	CAGCACGATC CAGATTGAAA ACCATCTCTT CTAATTGAAC CATAAACTTA TCTTTTTTCT	3360
	TACTAAACAT TTAAAATCCT CCATTTAAGC GATTGTCACC AATCACATTC AGTTATAATT	3420
	TGTTTCAAAT TAAGACAAGT GAATTTACAA ACTAATGATA CAAATTTGTT ATTATCAATC	3480
10	GTCAGTATAA TTTTAGTGTA CTGATATTAA TTTCAAAAAT GCCTCACAGT AAACAATTTA	3540
	CTGTATTTGC CCTTATA	3557

(2) INFORMATION FOR SEQ ID NO: 258:

15	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1631 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
20	(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

25	AACTATACAT TTCGAAAAAT TCTTCTAGTG AACCTGCGCC ACCAGGAGCC ATGACAAATG	60
	CATCTGCAAG TTCTGCCATT TTATTTTnAC GTTCATGCAT AGAATCAACT AAAATTAATT	120
	CAGTTAAACG TTGGCTTG TG ATTTCATGTT CATCTAACAT TTTAGGCATG ACGCCAATAG	180
30	CTTTGCCGCC ATGATCTAAT ACACCATCTT GaATGGCACC CATAATGCCA ATTGACCCTG	240
	CACCAAATAC TAATTCATAA CCTTGTTTCAG CAAAATATTT ACCTAAATCG TATGCTTTTT	300
	GTACATATGA AGGGTCATGA CCTTTGCTTG CACCACAATA AACTGCGATT CGTTTCATGT	360
35	TAATCCAGCT CCTTAATTCG ATGAATGACT TTAAATAGTG ATTGTTCAAA CACTTTTGA	420
	TCTTGCTTTG TAAAGGTGG GGGACCTTTG TGGCGACCAC CTTGTTTTCT AATTTGTGCA	480
	TTCAATATATC GTTTATCTAA TAGTTGTTGA ATATTTTGG AATTGTATAT CTTCCCATTA	540
40	TGATGCATGA CAATTAAGAC TTTGTCGACT AATAAACTTG CGAGTCCATA ATCTTGAGTG	600
	ACTACGATAT CATCCTTCGT TGATAATTGA ACAATTTTGT AATCAACTGC ATCTGGTCCA	660
	TCATCAACAT ATAATGTTGA TACATGTGGA GGATATAATT GGTTGAAAA ATGGCTGAAG	720
45	CTCCGAATAA TTGTCACAAA AATGCCTGTC TCAGTTGTTA AATCTATAAT AGAATCAACA	780
	ACAGGACAAG CATCTCCATC AATAATAATA TGTGTCACAA TTATGCCTCT GTATTGTTTT	840
50	CTTTATTTTG TTGAGAGGCG CTTTGGCAA CATAATCTTT ATATTTTTTA AATGACTTGA	900
	TGCGTGCTTT ATCAGCTTCT TGTGGCGTT TTTGTTCTTC TTTGTGTCGT TTTCAATAT	960

55

CGCCTTTTT CTCAGTTTT TCATCTAATT TATTAGGTGT TAAGCCTGCT TTTCTTCGT 1080
 ATTTTGTGA TTTTTCATA TCTTAAATAC GTTGTATTTC ATTCTTTTCG CGGGCTTTTT 1140
 5 GCTCTTCITT ATGACGCTTT TCGATATTTT TTTGAAGTAT TTTATTCATT TTATCAGCGT 1200
 CTTTACGATT TTGTTTAGCT AATTTTTCGC CTTTTTCTC AATATAGGCA GGATCATGTT 1260
 CTCTAGCAAA CTTTTTAAGT TCACGTTTAT TTTCAAAATC TTGTTTTTTA TCGCCGACAT 1320
 10 ATTCTTTAAC ATCACTCGCT GTGTTACTGA TTGCTGCAGA TGTTTTTGAA GCAACTTTAC 1380
 TTGTAGCATC TGTAACTTTT TGTACGTCGG GATGTTGTTT GATACGTTTA CGTTCAACAA 1440
 TTAACGGTAC CAATACAATT GGTAAATACAT TAATCATAAA TTTGATGACT TTTTCTTAT 1500
 15 CCATAGATCT TGcCTCCaTA ATTACTTTAT TAAtTTTACa TACCctATGa TACATCAATA 1560
 TAAACGATGA TAGTAGTGAA TCACTATTAA GTATTTTCAAG TGTTTTTTAA AAgAAGaCCC 1620
 20 AATTAnAAAA A 1631

(2) INFORMATION FOR SEQ ID NO: 259:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6645 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

CGAAATCATG ATTTAATGCT TTTTCATATA AGCTTTTCCA ATTAATCTTT CGTCCATGAT 60
 ATTCTTCAAC TGTGCTAGA TATTGTGCAA TTTTAGTTAC TTTAAAGGAG TGTGCTGCAA 120
 35 CaTTGTGkTC mAAATATTTA AATTTTCCaG GLaATCTTAT AAGTCTTTCC aTATCTGATA 180
 ATCTTTTAAA ATATTGATGT ACACCCATTT CAATTACCTC CTCCATTAAT TAATCATAAA 240
 TTATACTTTC TTTTACATA TCAATCAATT AAATATCATT TaAATATCTT CTTTaTATAA 300
 40 CTCTGATTAA ATGATACCAA AAAATCCTCT CAACCTGTTA CTTAAACAGG CTAAGAGGAT 360
 AGTCTTGTCT TGATATATTA CTTAGTGGAT GTAATTATAT TTTCTGGAT TTAATAATTGT 420
 TCTTGAAGAT TTAACATTAA ATCCAGCATA GTTCATTTCA GAAACAGTAA TTGTTCATT 480
 45 AGGGTTTACA GATTCAACAA CACCAACATG TCCATATGGA CCAGCAGCTG TTTGGAAAAT 540
 AGCGCCAAC TCTGGTGTTC TATCTACTTT AAATCCTGCA ACTTTTGCTG CGTAATTCCA 600
 50 GTTATTTGCA TTGCCCCATA AACTTCCTAT ACTTCTACCT AATTGTGCAC GACGATCGAA 660
 AGCATAATAT GTGCAGTTTC CATAAGCATA TAAGTTTCCT CTGTTAGCAA CTGATTTATT 720

	TACATTAACT GTCTTAGTTA CTGCTTGCTT AGGTGCTTGC TTAAC TACTA CTTTTTTAGA	840
	TGCTTGTTGT ACAGGTTGTT TTACTACCTT TTTAGCTTGG CTTGCTTTTC TTACTGGTGA	900
5	TTTAACCGCT TTAGTTTGTT TCACTTTATT TTGAGGCACA AGTGAAATCA CGTCACCAGG	960
	AAAAATTAAA GGTGTTACAC CAGGATTGTA TGAATATAAT TGATTCAACG TTAAGTGATG	1020
	CTCTAAAGCA ATCTTATATA ATGAATCGCC AGCAACTACT GTATAAGTTG TCGGTGATTG	1080
10	CGTTTGTGCT TGAACATTTG ATACATAATT ATGTTGAACA GGTGTTTTTA CTTGTGTGCC	1140
	ATGTTGTTGT GCATGTGCTG CATTATTTAA AGCTAAAAA GCTAACACTG ACGAAACCGT	1200
	CACTGTAAGA GATTTTTTCA TCTTGCTGTC ATTCCCTTGC TGTTAGTATT TTAAGTATGC	1260
15	AAATACTATA GCACAATACA TTTTGTCAAA AGCTATTGTT ATAACGATGT AATCAAATGG	1320
	TTAACAATAT AAAAAGAATA CAACCTTTTA TCATAGTGTA AAATGTATTC ATACCATGTA	1380
20	ATTGAGAACG TTTTCAATAA TTAATTC AAT ACCTTGAAAA TCGCCATAGG TAATATTACT	1440
	AAATGCACAC TGCATATGTT GTTTTAAACA ACACAACCTT TAAAAAATAT ATTCTAACTC	1500
	TATCTACCGA ATTGTACTTA AATATTCATA AACAAATCAT ATTCCAAAAT CTAATTTACA	1560
25	ATTTATTTAG CTACCTTTAA AAAAACC AAA AACCGACGCC CTTT TAGAGC CTCGGTTTTA	1620
	AAATATATAT TAATCGTGCG ACATTGTCTG TCTTAAATAT GATTGATAA ATGGTCCAAT	1680
	GTCTCCATCC ATCACTGCAT CAACCTTACC TGTTTCTTCG TTCGTACGAT GATCTTTCAC	1740
30	CATTGAGTAT GGATGGAAAA CATATGATCT AATTTGGCTT CCCCAGCCGA TTTCTTTTTG	1800
	TTCCGCCACGA ATTT CAGCCA TTTCACGTGC CTGCTCTTCC AATTTTAATT GATATAATTT	1860
	AGACTTTAAC ATTTTCATAG CTGCTTCACG GTTTTAAATT TGAGAACGTT CATTTTGGTT	1920
35	ATTAACA AACT ATACCTGAGG GGTGGTGGGT AATTCGTATT GCCGATTCAG TTTTGTTAAT	1980
	ATGCTGACCA CCTGCACCAG AAGCTCTGAA TGTATCAACT GTAATATCAT CCGGATTGAT	2040
	TTCAATCTCT ATTTTCATCAT TATTAAATC TGGAATAACG TCGCATGATG CAAATGATGT	2100
40	ATGACGACGT CCTGATGAAT CAAATGGAGA AATTCGTACT AGTCGGTGTA CACCTTTTTC	2160
	AGCTTTTAAA TAACCATAAG CATTATGCCC TTTGATGAGC AATGTTACAC TTTTAATCCC	2220
45	CGCTTCATCC CCAGGTAGAT AATCAACAGT TTCAACTTTA AAGCCTTTCT TCTCACAATA	2280
	ACGTTGATAC ATTCTAAATA GCATATTAGC CCAATCTTGA GACTCCGTGC CACCTGCACC	2340
	AGGATGTAAC TCTAGAATTG CGTTATTGGC ATCGTGAGGC CCATCTAATA ATAATTGCAA	2400
50	TTCGTATTCA TCCACTTTAG CCTTAA AATT AATGACCTCT TGCTCTAAGT CTTCTTTCAT	2460
	TTCTTCATCA AATTCTTCTT GTAATAAATC CCAAGTAGCA TCCATGTCAT CTACTTCTGC	2520

55

	TTGCGCTTTC GTTTGGTTAT CCCAAAAATT AGGTTCTGCC ATCATTTCCTT CATATTCTTG	2640
	AATATTAGTT TCTTTGTTCT CTAAGTCAAA GAGACCCCTT AATTTGTGTT AAATCTTGAT	2700
5	TATACTTATC TATATTTTCTG TTGATTTCTG ATAATTCCAT AGCATTGCTT CCTATTTATA	2760
	TTTCAATTCA AGTCATTGAT TTGCATCTTT TATAATGCTA AATTTTAACA TAATTTTGTT	2820
	AAATAACAAT GTTAAGAAAT ATAAGCACAC TGACAATTAG TTTATGCATT TATTGTTAAA	2880
10	AATCAGTACA TTTATCATCG ACATATGCCT AAACCGATT TTTAAACTA AGTACATAAC	2940
	AACGTTTAAC AACTTCTTCA CATTTTTTAA AGTATTTAAC GCTTGTAAAA TAAAAAGACT	3000
	CCTCCCATAA CACAAACTAT AGGTGTTTAA TTGGAAGGAG TTATTTTATA TCATTTATTT	3060
15	TCCATGGCAA TTTTGAATT TTTTACCACT ACCACATGGA CAATCATCGT TACGACCAAC	3120
	TTGATCGCCT TTAACGATTG GTTTCGGTTT CACTTTTTCT TTACCATCTT CAGCTGAAAC	3180
20	GTGCTTCGCT TCACCAAAC CTGTTGTTTT TTCACGTTCA ATATTATCTT CAACTTGTAC	3240
	TACAGATTTT AAAATGAATT TACAAGTATC TTCTTCAATA TTTTGCATCA TGATATCAAA	3300
	TAATTCATGA CCTTCATTTT GATAGTCACG TAATGGATTT TGTTGTGCAT AAGAACGTAA	3360
25	GTGAATACCT TGACGTAATT GATCCATTGT GTCGATATGA TCAGTCCAAT GGCTATCAAT	3420
	AGAACGAAGT AAAATCATAC GCTCAAACCT ATTCATTTGT TCTTCTAAGA TATCTTTTTG	3480
	ACTTTGATAT GCTGCTTCAA TCTTAGCCCA AACGACTTCG AAAATATCTT CAGCATCTTT	3540
30	ACCTTTGATA TCATCCTCTG TAATGTCACC TTCTTGTAAG AAGATGTCAT TAATGTAGTC	3600
	GATGAATGGT TGATATTGAG GCTCGTCATC TGCTGTATTA ATATAGTAAT TGATACTACG	3660
	TTGTAACGTT GAACGTAGCA TTGCATCTAC AACTTGAGAG CTGTCTTCTT CATCAATAAT	3720
35	ACTATTTCTT TCGTTATAGA TAATTTACAG TTGTTTACGT AATACTTCAT CGTATTCTAA	3780
	GATAAGTTTA CGCGCGTCGA AGTTATTACC TTCTACACGT TTTTGTGCTG ATTCTACAGC	3840
	TCTTGATACC ATTTTGTGATT CAATTGGTGT AGAGTCATCT AAACCTAGTC GGCTCATCAT	3900
40	TTTCTGTAAA CGTTCAGAAC CAAAACGAAT CATTAATTCA TCTTGTAAAG ATAAATAGAA	3960
	GCGACTATCC CCTTTATCAC CTTGACGTCC AGAACGACCA CGTAACTGGT CATCAATACG	4020
45	ACGAGATTCA TGTCGCTCTG TACCTATTAC TGCTAAACCG CCTAATTCCT CTACGCCTTC	4080
	ACCTAATTG ATATCTGTAC CACGACCAGC CATGTTAGTG GCAATAGTAA CGGCACCTTT	4140
	TTGTCCAGCG CCTGCAACAA TTTGAGCTTC ACGTTCATGA TTTTTCGCAT TTAACACATC	4200
50	ATGACGGATA CCACGTTTTT TAAGTAAATT TGAATATAT TCAGAAGTCT CAACTGCAAC	4260
	AGTACCTAAT AGCACTGGTT GCCCTGCCTT GTGTTTTTCA ACAACATCTT CTACTACTGC	4320

	TTTATTGTGC GGAATTTGAG TTACTGTCAT GTTATAAATA TTTCTAAATT CTTCTTCTTC	4440
	AGTTTTAGCT GTACCTGTCA TACCCGCAAG TTTATTGTAC ATTCTGAAAT AGTTTTGGAA	4500
5	TGTAATAGAC GCCATAGTTT TAGATTCAIT TTGAATTTGA ACGCCTTCCT TCGCTTCAAT	4560
	AGCTTGGTGT AAACCTTCCG AGAAACGACG GCCTGGCATT GTACGTCCTG TAAATTGATC	4620
	GACAATTAAT ACTTCGCCAT CAACAACCAT ATAGTCTACG TCACGTTGTA ATGTAACGTG	4680
10	CGCACGTAAA GCTGTGTTGA TATGACTAAT AACATCAACA TTTTGTACAT CATATAAGTT	4740
	TTCAACTTTG AACATACGTT CAGCTTTATC CGCACCTTGT TCTGTAAAT GTACAGCTTT	4800
	CGTTTTTTCA TCGTATTTAT AATCTTCGTC CTGTTTTAAC ATTTTCGCAA AACATTTGC	4860
15	TTGTGTATAA AGTGACGTTG ACTTTTCAGC TTCACCAGAA ATAATTAATG GCGTACGTGC	4920
	CTCGTCGATT AAAATTGAGT CAACCTCATC AATGATTGCA AAATGTAATG GACGCATTAC	4980
20	TCTATCTTCA GAATAATTCA CCATGTTATC TCGTAAGTAA TCAAAACCTA GCTCATTATT	5040
	AGTACTGTAA GTAATGTCCT GTGCGTATGC TTCACGTTTT TCTTCTGTCTG TCTTACTGTT	5100
	TAAGTTTAAT CCGACAGTCA AACCTAAGAA GTTATATAAC TCAGCCATTT CTTCACTTTG	5160
25	AACACTTGAT AAGTATTCAT TGACTGTAAT AACGTGAACA CCTCTACCAG CTAATGCATT	5220
	TAAGTATGTT GGCATTGTCT CTGTAAATGT TTTACCTTCA CCTGTTCTCA TCTCAGCGAT	5280
	ATCACCTTTA TGAATTGCAA TACCACCCAT AATTTGAACT TTATATGGTG TCATATTGAA	5340
30	TACACGTTTA GAGCCTTCTC TAACAAGTGC ATATGCTTCT GGTAATAATT TATCTAAATA	5400
	ATCATTTTGC TTTTGTACAT TATCAATGTC AGCTAATCT GTTTGGAATT GTTTCGTTTT	5460
	ATTACGAATT TCTTCATCAG TTAATAATGC CGTTTTTCT TCTAAAGCGA TTACTTTATC	5520
35	AGCAAGTTTA CCTAACTGTT TAATTTCTTT ATTATTGCCA TCAAGAATTT TTGATAAAAA	5580
	TCCCATTTCG TCGCTCCTT TAGCTAAAAA ACTGTTTGGC CTACAACAAT ATATCTTATC	5640
	ATTATAGTT AGAAAATTAT ACTTATTTAC TCATTGTAG AATCAATATA AATATATTTA	5700
40	TGACATACTT CATTACATT CTGTTGTCAA CAAGTTTATC ACTAATAAAT ATATTCTCAA	5760
	TACGCAATTA TACTTCCTAA TAAATTATAT TATAAATATT TTACGATTTT CGACTCGGAC	5820
	TATACAATAG ACTGACATAC TATTATTAAC TTAACATTCA AATATATACA TCCATTAACA	5880
45	TTAGCATAGT CACTATGTTT CATTCAACAA ATTACATTAT CGAACTATGA AATAGTCATA	5940
	ATTGCTTTT GGAGTATAAA AAAGCACTTG TGCAAAAACA CAAGTGCTTT AAACTTAATT	6000
50	TATTGTTTAC TAGTTTGAAT CAAGCCATAT TTACCGTCTT TACGGCGGTA AACGATACTT	6060
	GTTCATCAG TTTCTCTGTC TGTGAATACA AAGAAGTCAT GACCTAATAG ATTCATTTGT	6120

55

ATCTCGTTAT CATCGTAAGC GTCATTATCA ACTTGTGTTT CTTGCATTTC TTGTAATTCG 6240
 GCAACAAACA CTTCTTGATC TCCTCGATCA CGGCTCTTAC GATTAAATACG TGTTTTATAT 6300
 5 TTTTCGAACTT GTCTTTCAAG TTTATTATTA ATTAATCAA TACCTGCGTA TAAATCATCG 6360
 TTTTCGCTCTT CAGCTCTTAA CGTAACATTT TTCAATGGAA TTGTTACTTC AATTTTAGTA 6420
 GCTGAATTTG AATAAGTTTT AACTTTAACA TGCGCCACTG CATTGTTGAC GTCATTAAAA 6480
 10 TAACGTTCCA ACTTACCAAT TTTTTCCKCA ATATAGTTGC GAATAGCATC TGTGATAGTG 6540
 AGGTTATCTC CATGAATTTT AAATCTAATC ATAGTAAATC TCTCCTTAAA CCTCTTTATn 6600
 GGNAACTCnT TATTATATTT AACATTTTTA CGCCAATCGT GCAAA 6645
 15

(2) INFORMATION FOR SEQ ID NO: 260:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7430 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

CAGTTCCAGC ACATCTATTG GGGATCAACA AACTAGGGAA AATGCTAATT ATCAACGTGA 60
 AAACGGTGTT GACGAACAGC AACATACTGA AAATTTAACT AAGAACTTGC ATAATGATAA 120
 30 AACAAATATCA GAAGAAAATC ATCGTAAAC AGATGATTTG AATAAAGATC AACTAAAGGA 180
 TGATAAAAAA TCATCGCTTA ATAATAAAAA TATTCAACGT GATACAACAA AAAATAACAA 240
 TGCTAATCCT AGCGATGTAA ATCAAGGGTT AGAACAGGCT ATTAATGATG GTAAACAAAG 300
 35 TAAAGTGGCG TCACAGCAAC AGTCAAAGA GGCAGATAAT AGTCAAGATT CAAACGCTAA 360
 TAACAACTTA CCTTCACAAA GTCGAATAAA GGAAGCACCA TCATTAAATA AGTTAGATCA 420
 AACAACTCAA CGAGAAATTG TTAATGAGAC AGAAATAGAG AAAGTACAAC CACAACAAA 480
 40 TAATCAAGCG AATGATAAAA TTACTAATA CAATTTTAAC AATGAACAAG AAGTGAAACC 540
 TCAAAAAGAC GAAAAACAC TATCAGTTTC AGATTTAAAA AACAAATCAA AATCACCAGT 600
 AGAACCAACA AAGGACAATG ACAAGAAAAA TGGATTAAAT TTATTAAAAA GTAGTGCAGT 660
 45 AGCAACGTTA CCAAACAAAG GGACAAAGGA ACTTACTGCA AAAGCGAAAG ATGATCAAAC 720
 GAATAAAGTT GCCAAACAAG GGCAGTATAA AAATCAGGAT CCTATCGTTT TAGTGCAATG 780
 50 TTTCAATGGG TTTACAGATG ATATTAATCC TTCAGTGTTA GCTCATTATT GGGGCGGTAA 840
 TAAAATGAAC ATTCGCCAAG ATTTAGAAGA AAATGGTTAC AAAGCTTATG AAGCAAGTAT 900

55

	TCGTGTAGAT TATGGTGCAG CACATGCAGC AAAATATGGA CATGAACGTT ATGGAAAAAC	1020
	ATACGAaGGA ATTTACAAAG ACTGGAAACC AGGACAGAAG GTACACCTAG TTGGACATAG	1080
5	TATGGGCGGT CAAACGATAC GTCAACTAGA AGAATTACTG CGTAATGGTA ATCGTGAAGA	1140
	AATAGAGTAT CAAAAGAAAC ATGGTGGcGA AATTTCTCCA CTATTCAAAG GTAATCATGA	1200
	CAATATGATT TCATCAATTA CTACTTTAGG AACACCACAT AATGGTACAC ACGCATCAGA	1260
10	TTTAGCTGGT AATGAAGCTT TAGTGAGACA AATCGTATTT GATATCGGTA AAATGTTTGG	1320
	TAATAAAAAT TCAAGAGTAG ACTTCGGGTT GGCTCAATGG GGTCTAAAAC AGAAGCCAAA	1380
	TGAATCATAT ATTGATTATG TCAAACGCGT TAAACAATCT AATTTATGGA AATCAAAAGA	1440
15	TAATGGATTT TACGATCTGA CGCGTGAGGG TGCAaCAGAT TTAAATCGTA AAACGTCGTT	1500
	GAACCCTAAC ATTGTGTATA AAACATACAC TGGTGAAGCA ACGCACAAG CATTAAATAG	1560
	CGATAGACAA AAAGCAGACT TAAATATGTT TTTCCCATTT GTGATTACTG GTAACCTAAT	1620
20	CGGTAAAGCT ACTGAAAAAG AATGGCGAGA AAACGATGGT TTAGTATCCG TTATTTCTTC	1680
	TCAACATCCA TTTAATCAAG CTTATACAAA AGCGACAGAT AAAATTCAAA AAGGCATTTG	1740
25	GCAAGTGACG CCTACAAAAC ATGATTGGGA TCATGTTGAC TTTGTAGGAC AAGACAGTTC	1800
	TGATACAGTG CGCACAAGAG AAGAATTACA AGATTTTTGG CATCATTTAG CAGACGATTT	1860
	AGTGAAAAC TAAAAAGCTGA CTGATACTAA GCAAGCATAA TTTATAAAGT AAAGGGAGGA	1920
30	ATTAATAATG ACTGCAGACT TCTTTCAATT AATCGGATCA TTATTTAGAA TTCTAAAAGA	1980
	ATTATTCAAG TAAACATTG GCGAGGCCCC AACATAAGA ATTTGAAAA GAAATTCTAC	2040
	AAACAATGCA AGTTGGCGGG GCCCAACAA AGAAGCTGGC GGAAAGTCAG CTTACAATAA	2100
35	TGTGCAAGTT GCGGGGGCCC CAACATAGAA GCTGGCGGAA AGTCAGGTTA CAATAATGTG	2160
	CAAGTTGGGG TGGGACGACG AAATAAATTT TGCGAAAATA TCATTTCTGT CCCACTCCCA	2220
	TTGGCATTTA CGAAGTTTAA ATGTGCAATT AGAATATATG TATAACAATA TTAAACACGC	2280
40	GGTAAACGA AGTCAGTCAA TTCAAACCTGA TTTCGCCAC CGCGTGT TTTT TAACATAGCT	2340
	TAATAATTAA TAAGCATTAA TGTTCAATTT ATATGGTTGT TTTCCAATAA TAAACCTAA	2400
	GATATAGAAT TCACGCAATA TCATGCCGAC ACCTATACAT AATCCTAAAA TGAATAGTAG	2460
45	TGATATCGCT AGAAAGACCA TTGTATTATC CTCAAATATA TTTGTATATG CAAACAATGA	2520
	GTCTAGAATG ATTGGATGTA ATAAATAAAT AAAGAATGAG AAAGCACTAA TCATTTGAAT	2580
	CGTATTAAAT AACATTGTTT TAAAATGCGT GCAAAATACCC AAGATAACAA TAAACATAAT	2640
50	ACTATTATAT GGTGTTAATG AATATGAAAA GCTGGTAACG TTCCAATAGT CTCCaTTTGT	2700

55

	TCTAAGAAAT TTAATACACG TTCGTAGTTA TAACCCATAT ATGCACCTAA GAAGAAATAA	2820
	AAAATCCATC CGAATATTAT AGTATTTTCA CTTAATGGAT AATAGTGTAG CACGGTATCG	2880
5	TGAAACGCTG TGTGTTTCGT AAAGTAATAT AAAAATGATT GCTGTAAAAT AAAAGATAAC	2940
	AATAATAATA TTTTACTGTT GAATAGGTTA TAGTTAATTT TAAAAATGAT ATAACTCAAA	3000
	ATAAAGAATT GCATGATAAC AACGATAAAA TAGCCATACC ATTGACCTAA TAGGACATTT	3060
10	TCAATGAATT GTTTATTGAA ACTTGAATCT GTTAATAATG ATTCATAATA ACTGTAAAAC	3120
	AATCCCATTA ATATGTAAGG AATAAGTATA TATTTTACGC GTGTAGTTAA GTATCTATAG	3180
	GTGACTTTTT GGTAAATCAA GGTGTGTCAGT AACTGTGACA AGATAATAAA GCAAGGTGTA	3240
15	CCAAAAATCA CAATATTACG AATGTAAAAT TGTAACACTA AGGATCCACC CTCCATATTT	3300
	TCATGTTTTA AAGTAATTTG TGTAAGTAAA TGTGTGATAA TAATAATTGC ACATATAATA	3360
	GCACGTAAAT ATACGAGTTC AAGTCTAATC TTTTTCATGG AATCCGTCCC ATCTCTTAAT	3420
20	TAAATGCTCA AAAGCATCAT CACTAATTAA TATTCTAGGG ATGTAATAAT CATTGGAGTT	3480
	CGGAGTGACT GCTTTTTCCT CTAATGAAAA ACCGTATTTT AACCAGCTT TTTTGATTAC	3540
25	CGGTAATTTA TCGTCATTCA TCAAGCCATA AGGATAGGCT ATAGTTTCTT GCGACTTTTT	3600
	AAAGTTTTTA GTTAGATATT TTTCACTTTT GTTTAAATCT TTTATGATTG TAGCTTCAGA	3660
	AGCTTTCATT AATTTTGACT TATTATTTTT AGATAAGTTA TGCAAATCGT GGGTATGTGT	3720
30	TTCAAATTC CATAACCCAG TTTTATACAT TTCTTTTAGT TCTTTTTTAC TAATCATATC	3780
	GAGGTTGTGA AAGTTTTCTT CCCCACATG ACCTGTGATA ATAAACCCAG TTGCCGGTAT	3840
	TTTATATTTT TTTAAGATTG GATAAGCATT TTCATAAATA GTTTCATCCA TATCATCAAA	3900
35	GTTAATCCAT ACACTTCGTT TTGGAAACTT ACCTTTTTTC TTGLAATATA AAAATTCTTT	3960
	CAAGTTAAA AATTTAGCAT CATGTGATTT TAGCCATTTT ATTTGAGATT CAAATTGTGA	4020
	TTGACTAACA CTATAATTTT TAATTTCTTT ACTACTAGAA AAGAAGTAAA TAAAATTATT	4080
40	CAGAAAATTC GCTTTTCTTA CACGGTGATA ATTTAATGCC AGAGCACTAT TTTCTTTATA	4140
	TTTCAGTTTT TTAGGTGAAT CGTCATCTGC ATTTGCAATA TGATGACCAT CCAGTGTGCT	4200
	TACAGGCAAT ATGATCAAGA TACTCAACAC TAAAATTATA AATTTTCTAT ACTTCACGAT	4260
45	TCTCTTCCTC TCTGCCATTT TTGAATCAAT ATGCTAATTG TAAAAAATAC AAAAATGATA	4320
	ATCGCGAAAA TGCCCATAGT TTCAAATATA TCTAAATTT CAGTATTTTC AATGTTTAAA	4380
50	GCAACACGTA TTGTATTGAT ACTTTCGTCA TGAATTTCAA ATATAGTACC AATATAAACG	4440
	AGTAGAACAA CTAAACAATA TATCCAAAAG ACACAAGATA TAGCGATAAG TGCTGTTTCT	4500

55

	TTGCGTAACC	ACCTTTCTTA	CGTTTTAATG	CTTTTGGAAA	TGCGACAAGA	ACTACTGCTG	4620
	CGTTAATAAT	CCAGTATACT	GTCGGATACC	AACTTACAAA	TATGAGTCCA	GCCATATTCT	4680
5	TTTTCTCGTA	GCGACTATCA	ATAAAGAGTG	CGACTGTAAA	TTGAATAACG	TTTATAAAAAG	4740
	TCATAGTAAA	TGATGATAGT	AGAAATATTG	AAAACTATA	TGTCATAAAT	GTATAGTCTA	4800
	AGAAGTTTGC	TGTTATGAAC	AAATAGCCTA	AATATAGAAG	CACTATATAT	ACCCATAAAA	4860
10	TCGAGATGAT	TTGCTCAAAC	ATCAAAATAT	ATAAAGGAAA	CCTTTTCGTT	TTCATTGTGC	4920
	TAAAAAAGTC	TCGTAGTAAT	ACTTCGTGTC	CCCCTTGAGC	CCATCTCAGC	CGTTGCTTCC	4980
	AAAGACCTCC	CAATGTTTCT	GGAACCAACA	TCCAACACAT	GGCAAGCGGT	TCATACTTAA	5040
15	TACGATATCC	ACGTAAATGC	AATTTCCAAG	AAACTGCAAT	ATCTTCGGTA	ATCATATCAG	5100
	TATCCCGAGTA	GCCAACGTCG	ACAACTGCAC	TTTTTTTAAA	TAGAGTGAAG	ACACCCGAAA	5160
20	TAGTATTGAC	TGCGCCAGCA	AGTGTCTGAC	TTCGCTTAAT	ACAGCCAATT	AAACTTGCAT	5220
	ATTCTATCGT	TTGAATTTTA	CCTAAAATAG	AACTCTTATT	TCGAATTCTA	GGATTACCTG	5280
	TAACTGCACC	AAGTTTTTGA	TCATGTTTGA	AATTCTCAAT	CATATAATAT	GGTGCATCTT	5340
25	GATCAACGAT	AGTATCTGCA	TCCAAGCACA	TTACATAATC	ATATGAAGCC	TGTTTAATGC	5400
	CTTGATTGAG	TGCGTTGGCT	TTACCTCTGT	TTTCTTGTA	ATCGACGAAA	ATAAAGTCAT	5460
	TATTTTCTTT	GATTTTATAG	ATGAGTTCTG	CTGTATTATC	TGAACCTCCA	TCATTAAATGA	5520
30	TAATAATTTT	TTTCTTCTCG	TATTTGAGTG	CAAGAACATT	AGACAACGTA	TCTTCAATCG	5580
	TTTCACTTTC	GTTATAACAG	GCAAGTAAAA	ATGTAATGCC	TTCTAATTCA	TCCACATTTA	5640
	TGTCAGGCTT	CTTGTTCAAT	GAATATCTAA	TTTCTCTGGT	AAAATAGAAA	TAAATTGAAC	5700
35	CGACAATCCA	GTAAATAGAC	ATAAATACAG	GATAAAAAAG	CAAAAAGTTA	AAAAATTGCA	5760
	ATTCTTTTAC	CTACCTTTTCG	TTAGTTAGGT	TGTAAGCCAT	ATGGTAATTG	ATAGTATTTT	5820
	AATTTGCAAT	AGATTGTTGT	TATAATTAAA	CGGAAATATT	TGTAATTGCA	ACTTAATTTT	5880
40	CCTGTAACAT	AGTGTGATTA	ATTTTCAGTA	GGGGGTATA	AAAATTGAAG	GATAAGATTA	5940
	TTGATAACGC	AATAACCTTA	TTTTCAGAGA	AGGGGTATGA	CGGTACAACA	CTTGATGATA	6000
	TAGCTAAAAG	TGTAAATATA	AAGAAAGCGA	GTTTATATTA	CCATTTTGAC	TCGAAAAAAA	6060
45	GTATTTACGA	ACAAAGTGTT	AAATGTTGTT	TTGATTACCT	TAATAATATT	ATTATGATGA	6120
	ATCAAAATAA	ATCGAACTAT	TCAATTGATG	CTTTATATCA	ATTCTTATTT	GAGTTTATTT	6180
50	TCGACATCGA	AGAAAGGTAT	ATTAGAATGT	ACGTTCAATT	ATCTAATACG	CCTGAGGAAT	6240
	TTTCTGGAAA	TATTTACGGA	CAAATACAAG	ATTTAAATCA	ATCATTAAAGT	AAAGAGATAG	6300

55

5
 10
 15
 20
 25
 30
 35

TGCTGTTTCT	TGAAAGTTGG	TATTTGAAAG	CATCCTTTTC	GCAAAAATTT	GGAGCAGTGG	6420
AAGAAAGTAA	AAGTCAATTC	AAAGATGAAG	TGTATTCGCT	ACTAAATATA	TTTTTGAAGA	6480
AATAATTTTT	GTTACTAGTT	TGTAATAATT	AACTTACTTT	TGTAAAAAAA	GACATGAGAT	6540
TATTTTTTTA	AATCTATATA	AAGTTGACAA	TACAAATCGA	TATTGAGAAT	ATTAAGATGT	6600
ATATGAATTT	TATAAATTAA	ATGCAATACA	TTAATATAAA	TATCAATTGT	TGCAAAATAC	6660
GATTTGTTCA	ATGATTTGAT	AATATTATTC	TTTATATTTG	TGAATGGTTA	AGTTTGTCTT	6720
TGAACATATT	ATAAAAGTGT	AATGTTCCCC	TGAAAAGAAT	AAGTTGTCAT	CTAATTACAG	6780
GAAATCCGCA	TAAATTAGAT	GAAATGGAAA	GTAATAAGTA	ATAATTTATT	GATAAGCGCC	6840
TATGTGATGG	TAAATCATGA	CATAGGCGCT	TTTTTTTATA	AGTTAAAAAT	GTAAATAAAA	6900
ATTATATAAA	TTACCCACAT	CTTTTTAAAA	GGTGTGGGCT	TTATTATCAT	TAACCCAACT	6960
CACAGTGACG	GGTTACGCAA	GGTATTGAAT	TACCGAGTAC	GGGCACGCTC	GGTGTGTGTA	7020
AGAGCAAATA	ATCAAGTAAT	GATGATGCTT	CTACTCGATT	ATAAGAAAGC	CATGATAGAG	7080
TACGATGGTA	TCTAGTTTTA	TTATTAATAG	GTTTGGATAT	TTAAAGTTGG	ACAATATTAT	7140
ATCTTGTGCA	AAAATATAAA	TAAGTTATAC	ATAATGGTAG	AGAATCATGA	TATAATTTTA	7200
AACGATAAAA	TATTTATATA	AATAATTAGA	GaAAATGTAG	TTGTGTATGT	yTTGTGGtCG	7260
TTAAACTAGA	TATAATTGTC	CGATTTATAA	AACATACATA	ATGAATACaA	TGATTGATTA	7320
TGTGGAGGAA	ACCATGAmAG	AmAAGTTTgA	TTTAGTAAAA	CTATTAAATA	TTCTAAAGAA	7380
GAATATTAAA	TTATTGCTTA	TTTTACCGGC	AATATGTCTT	GTAGTAAGTG		7430

(2) INFORMATION FOR SEQ ID NO: 261:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4082 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

45
 50
 55

ATTGTTACTC	ATTATTTTTT	CATATTCACA	CAATGATCT	TGTTTATATT	TAGCTAATTG	60
ATTTTTATCT	AGCATTTTAT	CCTCCTGCTG	AGTTTGTAAC	CTTTAATAAT	TTATTTTCTA	120
TAAAAACTTA	GTATTCCAGT	TGCTTATTAT	ATCATTGATG	AAAGGCTGAA	ATAAAACATA	180
AACTGTTCCG	ACCATTAACG	CTGTAGCTAA	AGATAAGTCT	ACAAGTCCAC	CTGTTTTTAA	240
TTGAATCGGT	GTCTTCACAT	TAAACGGTAA	gGAKnAAAAAT	AATTTCACGC	CTTTTGGTGT	300

	AGGCGTCTGA ATAATTTGCA ATAAAAATGC TATGATTGCG ATAAATAATA TTGAATGCGT	420
	AAAGGTTTCGA TGTCCAAAGA TCAATCTCAC AAAAAAATA ATTACCTTAA ACCTTCTGCC	480
5	AATCTTACTT TGAGTGTGAC ATATATCGGG TAATAAGCTA GCTAGAGTTG CTAGAATGAT	540
	AACCGTAACC GACGAAAAAA TATCCGTTTG AAAATATTGT GTTGTTAGCG CTCCAACGAG	600
	CATGCCGCAT GAAGCATGTG TTTTACCTGT CATATTTGTT CTCCTTTAAT ACTCACATTT	660
10	TACCACATCC CTAACAAAAA CACGAACATA TTTTCGGGTT AAAATTCATT AGTATGACAC	720
	AATTTAAAAA AGTATCACAT AACTCTTGAA AACGATTACA AAATCGTTTA TGATGTATTT	780
	ACAAAATATT TAAAGGATGT GTTTGAATAA TGGCAATGAC AGTAAAAAAG GATAATAATG	840
15	AAGTGCGTAT TCAATGGAGA GTTGCTGATA TCAAAATTCC TACAAGTGAA ATTA AAAATA	900
	TTACACAAGA CCAAGATATT CATGCAGTTC CTAAATTAGA CAGCAAAGAT GTATCTAGAA	960
20	TCGGCTCAAC GTTTGGTAAA ACGAATCGCG TTATTATCGA TACTGAAGAC CACGAATACA	1020
	TTATTTATAC TCAAAATGAT CAAAAGGTTT ACAATGAATT AACTAAATAA ATTGTATAAA	1080
	AAATCATTTC ATGGTGAGGG CTTCATGAAT GATTTTTTTA ATTGATTCAA CACCCAGCAT	1140
25	AAACAAATAC AAAAGGACAA CTGTTCCCAT AATTTTAACA GTTGTCCCTT TTCACATATA	1200
	TTTATAACAA AAGATGTGCC ATCAAAGAAA TAATTGGTAG TGTAATGATT GTTCTAATCA	1260
	AGAAAATCAT AAACAATTTG CCGATGCTTA CAGGAATCTT CGAACCAAGT ATGACGCCAC	1320
30	CTACTTCAGA CAAGTATATT AACTGCGATA TACTAAGTGC CCCAATAACA AAACGAGTTA	1380
	TATCATTTTG TACACCTTCA ATTAATATAG AAGGTAAAAA CATATCGGCA AAACCGATAA	1440
	TAATCGTTTG AGAAGCCTGT GCCGCTTCAG GTATTTGCAT TAACTCTAAA AATGGAACAA	1500
35	AAGGTTTACC CAATATGACA AAAAAGGGCG TGTAGTTCGC AATAATGGTA GCAATAGTAC	1560
	CAATACTCAT TACTACAGGC AAAATAACAA ACCACATATC AATGACTGTT TTTAATCCTG	1620
40	ACTTAAAAAA GTCAATAACG CCCGGTGCTT TAATACCTAC TTCTGTTGCA GTATCAAAGC	1680
	CATGTCTCAA TGCCGTCTTT CTTTCTGGCA ATGCCTCAGT ACGCGCACTT TCAGGTACCT	1740
	CCTTAGCATA CTCATCAGGA ATTTTATTTA AAGGCCAAAT TCTTGGCATA ATGACTGCTG	1800
45	CAACGAGGCA GGATACTATC ACTGATAAAT AGAAAGCAAA AAATTGATTT TGCATGTGCA	1860
	CTGTTTCAGC AACTACAATT GCAAAGGTGA TAGAAACTAC ACTAAATGTC GTTGAAATAA	1920
	CTGTTGCCTC ACGACGAGAA TAATATCCTT CACCATATTG TCTACTTGTA ATTAAGACAC	1980
50	CAACAGTTCC GTCTCCAATA AATGATGCTA AATTATCTAC CGTCGAACGT CCTGGCAATG	2040
	TAAATAAAGG TCTCATAACC GGTCTAAAAA TAGGACCCAA CATCTCTAAC AAACCGTATT	2100

55

	AACTTGAGAA CAATAATCCA CCCGTTTCAT CTGAGTAAAT AACCTTTGAA CCAATTTCGTA	2220
	AAAATGTCAT CCATGCAAAA ACAACTGCTA ATATTTCGTAA AATTAACCAA CCAATTCTAA	2280
5	CGTTAAAAGC ATTGTTTCATT AGCCCGTCAG GTTTC AATTT ATCTTTTAAA ATAGTTGAAC	2340
	AAATCAGAGT TATGATACCC GATAAAGTAA TTATCGTCAC AATTAAAAAT GGCATTACGC	2400
	CACCTAATAC ATCTTTAAGC ACGCCTGCTA AAAATGCCAC GGGCAACGTT GTTTGCTTCT	2460
10	GTCCATCTTG TTCGACTGGA ATTGGTACTA AAAATAATAA GATACCAATT AAAGACATCG	2520
	TAATAAACTT AAGTCTCCCA ATAACATCT CTTCCTTGA AAAGCTATCC ATAAATCAA	2580
	TCCATTTCTC TATGTATTCTG TTTTAAGTAT ATACAGAATT CTATTTCAGTT AACAAACATA	2640
15	TTCTTATCA TTCTATCTTT CAAAATGTTT ATGTATGCAA AATAATGAAT AATTACAGTT	2700
	ATTAAATATA CGCTATTTCT TGTAATTTTT CAAGATGAAT TCAAAAAAGG TTAAGTACAA	2760
20	TTACTGATTT CGTACTTAAC CTTTTTTAAA CTCTAATCAT ATGTTAGTTA TTTTCATTCTT	2820
	CGTAATAATA TTAAGAAGTA TGGTGCACCG ATAATTGCAA TGATAACCCC AACAGGAATA	2880
	TCCAGTGGCG GATGAATGCC ACGGGCTAAA CCATCTCCAA ATGTTAACAA TATAGCACCA	2940
25	ATTAACCCCG ACATGATAAT AACGTGTAAT GTTTTATTTT CTATTAATTG TCTCGCAATA	3000
	TGAGGTGCAA TTAATCCTAA AAAGCTAATA CCACCGACAA CTGAAATTGC GGATCCTGCT	3060
	AATATTACTG CTAAAATTAA CAATAGCATT TTAATAGTTT TAACTTTTAA ACCGAGTGCG	3120
30	GTTGCAACAG CATCACCTAG ATTCAATACA TCTAATTGAT AACTCAATAA AATGATGATT	3180
	GGTATCGTTA TTAAAAACCA AGGTAATATA GTATAAATAT TCGACATATC ATGTCCATAT	3240
35	AGACTACCTG TCAACCAAAC AAGCGCTTTG TTTGCTTCCA GTGGATTTCT GATTAATAAG	3300
	AACTGCACAA TCGCCGTACA TATTGCGCCT ATTGCTAAAC CAATTAAGGC AAGCTTTGAA	3360
	CCTTTAACAT CATATTTTGA AATTAAAAAT GATAAAAATA AACTTACTGC AAAGGCACCT	3420
40	AAGAATGAAC CTATAGGTAA TACAAACAAT GGTGCTGTTG GAAAGGTCAT AATAATAATC	3480
	ACAGCAGCTA AACTGGCACC TTTAGAAATA CCTATAACAT CAGGTGAGGC TAACGGGTTT	3540
	CTTATTACAG CTTGTATAAT TGCACCTGAA ATAGCCAAGC TACTACCGAT AATAATACCA	3600
45	AGTAATGTTT TAGGTATACG ATACTCATTT AAAATAAAAT CATCTTGTGT AAAGATTCCC	3660
	TTAATAGCAT CAATCGGATG AATCATGACA GACCCTACAC ATAAACTTAT GAATATACTC	3720
50	ACAATTAAAA GGATTGTGAT TAAACTATAA CGACGTATAA TTTTCGTTGT CATCATATTC	3780
	TTTTCACCCC TTTAATCGTT ATAAATAAGA AGTAAAGTGC ACCTACGAAT GATGTAACAA	3840
	TCCCTACTGG TGATTCAATA GGATATGTAA TTAAACGACT TAATACATCT GATAGTAGTA	3900
55		

AGCGTTTGAC TATATGCGGT ACGATTAAGC CAACAAATCC AATTGGTCCT GCCACTGACA 4020
 CCGACATACC TGTAAGAATA ATGACTAATA GTCCAATGAT AATTCTAACT TTATTTATAT 4080
 TT 4082

(2) INFORMATION FOR SEQ ID NO: 262:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1145 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

TATTAGAAGG TCGTTCGGAT GAACAATTAA AAAATTTAGT TAGCGAAGTA ACTGACGCCG 60
 TAGAAAAAAC AACGGGGGCA AATAGACAAG CAATTCACGT TGTATAGAA GAAATGAAAC 120
 CAAACCATTA TGGTGTGGCT GCGTAAGAA AGTCAGATCA ATAATCTTC ATAAGATGCA 180
 TGCCAATTAA TTCTTTGAAA ACGAACAAGG CGACTTCTAT CTGAGTATGA TAGAAATCGC 240
 CTTGTTTATT TTTAATCTTC ATCTAAAAAG TCTTTAATAG CTTGTTTATT TGTGTTTTTA 300
 TTAATCTGTA ATGCACTACC ATCAGTATTT GTATTGACAT CTTGCTATGA GTTCTTGATT 360
 GGCACAGTCA ATGACTTAAC ATCTTTTTCA CCTCGGATAC CAAAACTCAA ACCTGTTTGG 420
 AAAATCCCTG AATCAGGAAT GTTGTATTC ACATAGCCTC TTAAATACC TGCAACTTTT 480
 GGTAATTTAA CAACTGTTCT AAAATTAACC ATTTCTTTT TCAATGTTG CATCACTTGT 540
 TGCTGACGTC GCACGCGTCC GAAGTCACCT TCAGGGTCGT GACGGAATCT TGCATAACCA 600
 AGTAATCTTT TACCATTCAA CCTATGGTTA CCCTTTTTCA AAGATACACC AATATTTTTC 660
 GACATATCTT TTTGACATT AATTGGTACA CCTTCAGGCA TTAATTCATC AATCATTTTC 720
 TCAAATCCAG TAAATCAAC TACTGCATAA TATTCAGGAT TAATTCCTAA ATTTTATCA 780
 AGTGTTTTTT TAAGTAGCTC TGGACCACCT AAAGCGTATG CTGAATTAAT TTTGTGTTTT 840
 CCATATCCTG GAATATCTGC ATAAATATCA CGCATGACAG ACATCATTTT CATCTTTTTA 900
 TTGATAAGT CATATTGAAC AACCATGATA GAATCTGTTT TTGATTGTCC ACCTTGTGCT 960
 TTATCTGCAC CGAGTACAAG AATAGAAATT ATACCATCAT TTTTACTGG TCCATTAAAT 1020
 TGATGTACTT TAACATCTTT CGCATGTTT TTGGCATATT CTACACCGCT ATTGTAAC TA 1080
 TGTACAATAT ATACAATAA TGCCGATAAG TAAATTAACA ACAATCAGAA GAATGATAGG 1140
 TAATT 1145

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7075 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

10 TATGGCTCAT CATTAAATGCA CGTATCGGGT AGCGTTTACC ATTGATAAGT GCTTCATGTT 60
 TAGCACGAGT TCTTAAAT CCATCGCCAT AACCGATATC AACTACAGCT AATTTTGTAT 120
 15 TGTTTTTAGT CACTTCAAAG GCAAAGCTAT AACCGCAATA ATCACCAGCT TGTACTTCGC 180
 GCACCTGAAT AACATGTGCT TTTAAAGTTA ATGACTGAAC TATATCATGT TGATTCAGTG 240
 AACTATATGG TCTTGAACCG TATAACGCAA TACCTACACG CGCATGTGTA TGGTGGGGTA 300
 20 GTAATayyyg TCCTTCCCGA TAAAACTCG CACTATTTTG AGCATGGATT AGGTGGAAGT 360
 GATAACCTTC AGATAAAAGT GCTTCAACAA TTTCCATCCA TTGTGAACGT TCAACATTAT 420
 AAtCTGACAC ATCGAATTCA TCAGCATATC CAAAATGGGk CCATAAACCA CTAATAATCA 480
 25 TTTTTCGATT TTGATTATGA TGGTGATCTT TCAATACTTC TTTAATTTTCG TTTAAATCTT 540
 TAAATCCAGA CCGATGTAAT AAATTTTCAA ATTCTAAGTG AACATGAATA CCAGCTAAAT 600
 CATTTTTATG GTTATAGTAA TATGTCAACG ACGGCAAAGT CATGTGTATT TGATGTTTAC 660
 30 GGACTAAATC AAACCTCGTAA ACTGCATTCA TTAAAAAGAT TGTTGCATCT GGAGCAAGTT 720
 GTCTAATTTG AATTGCTTCT CGTAGTGATG TTGTGCTAAA TGTATCTATA CCTGCATGGA 780
 TAAACTGAGT TACAGCAAAT TCTAGGTCAT AGTGATATGC ATTaTTTTTA ACAACTGCCA 840
 35 TTAATGGCTG ATTGTTTTTG ACTGTGATTG CATTTTGTAA AAATATTTTC TTATTTACAG 900
 ACCtTGTGTC TGTCAATGTA TTACACCTCT TTGTAATTAT TTAATAAATT TTCGTAAAAA 960
 40 TTAACCACGT TTATTAACAC TTTTTCATCA AAATTTAAAT GTGATGTGTG CAAACCAGTT 1020
 AAAAAACCTT TATCTTCATT TCGTGTTCCT ATAAAAACAA AGTAAGCTGG AGCTAGTTGT 1080
 TGACCATAAA AACTAAAATC TTCCCCAAAT AAGAATGGCG TTGGTTTGTC ATAGACATTT 1140
 45 AAATCAGCTT TTATTAAGGC GTCCTCTATT TGAGTACGTA ATTTCCGACT ATTGATTGTA 1200
 GGGGGATAAC CTTCTGCAAA TTTAACTTCA CAATCTACAT TAAACAGAAG CTTGACACTT 1260
 TCTGCTATCT TGTGCATTTG ATTTTAAACG ATTGTTAAAT CATCAATATC ATATGTACGA 1320
 50 ATAGTACCTT CTAAATAGCC ATTACTTGGT ACAGTGTTAA TCGCTTCACC AGCTTTAAAA 1380
 TGACCAATAT GAACAATATT TCGTTTCAAA CCGTTAAGGT GAAATTGTTG AATTTGTGAT 1440

	ACATGACTTG ACAGGCCTGT TAAGAAAAAG CGATACTCTG TTGCGCTGGC CGTAATTTCT	1560
	TCATCTCTTA TCACTGCAAT GCCTTCATCA GCAAATGGGT TAACATGAAT ACCAAATACC	1620
5	GCTTCAATTG GATACTTATC AAAGGCACCG GCTTTTATTA ATCGATTGTC ACCGCCACCA	1680
	GTTTCTTCTG CAGGTTGGAA AATGAAAACG ACATTTTGCG GTAATTGACC TGCATCTTGC	1740
10	ATGTCTTTGC AACGTTGTAC AAAAAGCATT AATGCAGTTG TATGACCATC ATGTCCACAA	1800
	GCATGCATCA CATGATCAGA TTGACTGCGA TAAGGCACAT CATTTCCTC TAAATAGGT	1860
	AACGCATCAA TATCAGCTCT ATACGCTATC GTATGTGAGC CATTACCTC TAAGTATGCA	1920
15	ATGACGCCAG TTTCCAATGG GCAATCGTAT TTAATATTTA AACTATCTAA AAACGCTTTA	1980
	ATATAAGCAG TTGTTTCAA TTCATGTAAG CTTAATTCAG GATGTTGATG TAAATGACGG	2040
	CGATGTTTCG TAACAAATTC TAATTCATTC ATAATTATCA ATCCTTTGTG TTAAATTACT	2100
20	ATATAAATAG TGTAACGtAT TTCGAAATTT GTGATCATAA GTTTATTCAA TGCTAAACAA	2160
	TAAGGTTGAG ACATAATCGT ATCTCAACCT TGAATTATT ATACGTTGAC GTCAGTAGTC	2220
	ATTCAGTTTT CTTAATGCTG CTACAATCTC TTTTTAGTA TCTGTACTT CAGAAGCTTG	2280
25	CTTAATCACT TTTGCAGGTG TACCAGCAAC AACTGCACCA GCTGGTACAT CTTGTGTCAC	2340
	AATCGCGCCA GCTGcAACAA TAGCACCTTT ACCAACACGT ACACCTTCTA AAATAACTGC	2400
30	ATTTGCACCG ATTAATACAT CATCCTCGAT TATAACCGGT GAAGCACTAG GGGGTTCAAT	2460
	CACACCTGCT AATACTGCGC CAGCCCcTAC ATGTACATTT TTACCAGTTG TAGCACGACC	2520
	ACCGAGAGTA GCATTCATAT CAATCATTGT ACCTTCGCCA ACGACTGCGC CAATATTAAT	2580
35	TGTTGCGCCC ATCATAACGA CAGCACCATC TTCAATAATG GCTTGTCTC TAATAAACGC	2640
	ACCTGGTTCA ATTCGTGCAT TCGTATTTGT TAAGTCTTTT AATGGAATAG CAGAATTGCG	2700
	ACGATCCATT TCAATTTCTA TATCTTCGAA TTGACTACCA TATGCTTCGT AAAAAGGTTT	2760
40	CCAATCATCC GCTTCACAAA AGATTACTTT AGATTGTTCT GAACCAAATA CTTTAAACT	2820
	TTCTGGATAT GTGATGCCTT CAAAATTACC ATTTAAATAT ACTTTTATTG GTGTAGACTT	2880
	TTTAGCATCA CTTATATATT GAATAATTTT TTCAGCTGTT AAATGTTGTA CCATAAAATA	2940
45	ATCGATCTCC TTAAATATGT TTATAAGTTG TCAAACGTAT AAAAGCCGTT TGGTTTATTA	3000
	ACTAAGCGTT CTGCTGCTTG TATTGCACCA TTCGCAAAA TATCTTTTGA TTGTGCACGA	3060
50	TGCGTGATTT GAATCGTTTC ATCAGTGCCA GCAAATAGAA CTTCATGTTT ACCGACAATC	3120
	GTACCTCCAC GAATAGAATG TATACCAATA TCTTGTGGCT GCGGTTTTTC ATTTAATTCA	3180
	TGTCTATCAT ACACAGGTGT TACATTTTCT TTCAAAGATA CGATCACATC ATACAATTTT	3240

55

TCGAAATCAT CAAGTAGGGG AACAGCAGCT GCTAAAATTT TAGTCAATGC ATGAACGCCA 3360
 TAACTCATGT TCGCGCTGAA AAACACAGGC ATATTTTGAC TCAATTCATC TAACTTATTA 3420
 5 AGTAGTTTTT CTTTCTCGnC CAGTgTTGCC ACAACTAATG GCAAATGAAA ATCTTCATCT 3480
 AATAAAGGGA AAAGCAGATT TGGATTGAA AAATCTATTG CAACATCGGC ACCTTTAACA 3540
 10 TCTGCAATAT GTTGATATTG TTGATATGGC GTTGTGCTT TCGGTGTATT TTCAATGACC 3600
 CCAACGATT CATGTCCTTT TTCTTCTGCT AATCTAGCAA CGCGTTGATT CATTGCGCCA 3660
 TAGCCAATTA GTAATATTTT CACTCATTTT CACCCGCTTT AAATGTGTCA TATGTTTCAC 3720
 15 GAAGCACTTT AGTATCTGTA TCTTCTAGGC TAACCAATGG TAGACGTAAT TCATAATTTT 3780
 CAAATCCTAA ATAACCTGTT AGAGCTTTAA TAGGAATTGG GTTAATATCA ACTGATAAAG 3840
 CTGATAACAG TGTGCCGATT GGTTTAAATT GATCTTGAAT ATCTAATCCA CTTTGTGAG 3900
 20 CATCGTATAA CGCTTGAAAT TCTTTAGGAA TGACATTGGC AATAACAGAG ATAACCCCTT 3960
 GACCGCCACG TTGATAGTAT TCGACGACGT TGTCAATCATT GCCACTATAT AATGCAAATG 4020
 AATTTGTATC AATGCGCTTT TTCACCTCTT CTAAATACTC AAAATCATTG GTAGCATCTT 4080
 25 TTAAAGCAAC TATATAAGGA TGTTGACTTA ATATTTCTAC AGTTTCTGGT TCAATTGTCA 4140
 TGTTCTGTTCT TGAAGGAACA TTGTACAGCA CGACTGGTAA TTTCACAGCA TCTGCAATCG 4200
 CTTCAAAGTG TTTGACTAAA CCACGTTGGT TCGTTTTGTT GTAGTAGGGC GTAATTAACA 4260
 30 TAATTGCATC AGCCCCTAAG GCTTTAGCTT GGATTGAAGC TTGGATTGAC TTTTCAGTAT 4320
 CATTAGTGCC AGTTCCTGCT ATGACAGGAA CACGTTTATC TACAAGATCA ATAAGTGTTC 4380
 35 TTAGAATGCG TTCTTTTCA TCTGTTGTTA AAGTAGGGCT CTCAGCAGTA GTTCCATTAA 4440
 CGATGATTGC TTGGGCATTA TTTTCTAGTA AAAAATTAAC GTGTGTTTTT AAAGCTTCAA 4500
 TATFAACTTT GTTATTTGTA AAAGGGGTG TAAGTGCaAC SCCAACACCC TCAAATAAAT 4560
 40 GTGTCATTTT AATTCGCTCC TTTTAAACGC ATAACCTGTT CCAATACTG TACAGCATTT 4620
 AATGCAGCAC CTTTTAATAA ATTGTCTGAT GTACACCATA CATGGAAAGT ATTTTCTAAT 4680
 GAATCATCTC TACGTATACG GCCAACAAAC ACTTCATCTT TATTAGTAGA ATTGATTGCC 4740
 45 ATTGGATATT CATTGTTCTC TGGATTGTCT ACTAAAACAA CGCGGTCATC TTGATCAAAT 4800
 AACGCTTAA TATCTTCTGC TGTGTTTCT TTGTCAAGCG TTACATCAAT TTCAACACTA 4860
 TGACTATCTT GAACAGGCAC ACGTGCGCAT GTTGCTGTTA CTTTTAAGTC TGGCGCATTT 4920
 50 AAAATTTTTC TCGTCTCATC AATCATTTT TGTCTTCTT TTGTATATCC GTTTTCTAAA 4980
 AACACATCAA TATGCGGTAA CACATTATTA TAAATTGGAT GTGGATATGC TTCTGGTGCT 5040

55

	TGATATGTTG TATATGCCAC TCGTTTTAAA CCATAAGCAT CTTGCAATAC TTTTAGAGGT	5160
	ACAACAGATT GAATCGTAGA GCAGTTTGGA TTGGCAATGA TACCTCTTGT AAATGTAGGT	5220
5	TCATTGACTT CCGGAACGAT TAAATCAATA TCTTCTGCCA TACGCCATTG ACTTGAATTG	5280
	TCTATAACGA TTGCACCAGC TTTTTCAAAA AGTGGGGCAA AGTGTTTCGCT TGTACCGCCA	5340
10	CCAGCACTCA TTAATACATA ATCGAAATGT TCACTTGACAC GAGCATCAGT TAATTCTTGA	5400
	ACTGTATATG TTTTTCCTTG AAATTCAACT TCTTGCCCTG CAGAACGTGC TGATGAAAAT	5460
	AATACTAATT CATCGAAAGG AATATTTTAA CGATTTAATG TCTCCAACAT TTTTGTACCT	5520
15	ACTAATCCTG TTGCACCCAC AACTGCTAAC TTTGTCATAA CTTGTCCTC CATTTTATAA	5580
	TAATTTCCaA TTTTTAGAAT ATTTTAACAA TCATTTTACC ATTAAATGTT AAATGCGTCA	5640
	TATAGTTTTT CTACCGCTTG TTGCCCATTA AAATCATCAA TGACGTATGA AATACTTATT	5700
20	TCAGATGTTG TTGTTTGGA GAAAGGTATA TTATTTTCAA TTAATGTCAA AAATGCTTTT	5760
	GATGCCACAC CTGACATATC ACGCATGCCT GAGCCAATTA ATGAAATTTT GACATAATGC	5820
	TCATTGATT TATAAGCTAA TGCTTCATAT TGATTCTTTA ATGTTTCAAG AATCATAGAA	5880
25	ATTTGATGAA AATCACTATC TTTAATCGTG AAGGATAGTT GTAGCCCATC CAAGTTGACG	5940
	ATTTGTGAAA TCATATCAAC ATTTACAGCA CCTTCTTCAA GTTCCGTAAA TAGTTGGGTA	6000
30	AGTAGCTGAT TGTCAAGTAG GGGATAACTA ATTGTTACAT GCATCATATG TTTATCCAAA	6060
	GCCACACCAG TAACTGCTTT TTTCTCTAAT ATTTCTTCAT TTGACATAAT CCATGTTTCT	6120
	TTCACGTTCTG ATAAAGTTTT TCCTAAATAT AAAGGGATAT TATAGTTTTT AGCTAATTCA	6180
35	ACACTTCTTG TTTCAAGTAC ACCAGCACCT AAAGCGCTCA TTTCCATCAT TTCTTCATAT	6240
	GAGACGATGT CTAGTCGTTT AGCCTTTGGT AAAAGTCTTG GGTCAGTGGC ATACACACCA	6300
	TCAACGTCGG TATAAATTTC ACAAGGTATT TGATTACTAA CAGCAAGTGC CACAGCGGTC	6360
40	GATCAGAAC CACCTCTGCC TAAAGTTGTT AATTCCTGAT GTTCATTGAT GCCTTGAAAT	6420
	CCAGCAACTA CTAAATATC GTTTTCTTGA AAGGCTTGTT CAAATGTTTG AGGATTAATT	6480
	TGAGCAATT TACTTTTTTAA ATGATGGCCA ATGGTTTTTAA TACCCGCTTG ATAGCCAGTC	6540
45	ATTGCTTTGG CATTCAATAC GATATCATT AATACCATTG ATAAATAAGA TACAGTTTGT	6600
	TGCTCTCCGG TTGTCAATAA TAATGCCAGT TCTTGTGTGTT TTGGTGCTTT AGTCAAGGTT	6660
50	GATACATTCG TCATTAATTG ATCTGTTGTG TTACCCATAG CACTTACAAC GACAATTAAa	6720
	TTGTTTCTCT TGATTGACTC GCTCCTTTAA CATTTTCAGCG ATCCTTTTTTA TTTTGTAAa	6780
	ATCACTGACG GATGATCCGC CAAATTTCAA CACACTTCTT GTTACCATAT AATCCTCCTA	6840

55

TAATCTATAT ACAAGTGATG CACTCCATTA TTTTAAATA ATGACAACT CTCAGCTCTT 6960
 AACCAAAAAG TCCAACAAAT TATAACTGCT ATTATAATTG CTTCGGCATC GCACCCTTTC 7020
 5 AAATTTAGCT GTTAGCAGAC AGTAATCTAa ACTTTACTCA TGATTGATGC GCCTC 7075

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5171 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

AGACGTACTT TGTGATTmCG AAGyrCGTAC TmAGCACTT GTCGACGTTG ATGTACTTGT 60
 20 TGAACCTGAT TGACTAGTAC TTTGTGATAA TGACTTACTA TCAGAATCAG ATGTACTTTG 120
 TGAATCACTT AATGATTCTG ATGTACTACC TGACTGAGAC GTGCTCATTG AACTACTTAC 180
 GGACATTGAT TTAGTGTCTG ATGCAGATAA TGACCCACTT GTACTGATAG AGTCACTTAC 240
 25 TATCTCTGAA GTACTCATCG AGTCTGATGT ACTTGTTGAG AACTTTGTG ATGCTGCTAT 300
 GCTTAGTGAT CCAGAAACAG AACCCTTGT GCTCGTCGAA TCGCTCAATG ATTCTGATGT 360
 ACTCATCGAT TTTGAATCAC TTGTACTTAA TGATATTGAT GTACTTTGTG AATCTGATTT 420
 30 GCTTGTTGAC GCACTTTGAG AGTTGGCTAT GCTATTTGAA ATACTGATAG AGTCCGAGGT 480
 GCTAGCTGAC TCGCTCAATG ATGTTGATGT ACTAATTGCA TTCGATGTAC TGCTACTTAA 540
 35 TGATGCTGAT GTACTAGACG ACCCTGATAT ACTCGTTGAT AAGCTTTGTG ACTTAGACAA 600
 GCTTCCTGAT GTACTCATAC TTAATGAGTC ACTGAGTGAT GTTGATGTAC GCAATGAATC 660
 AGATSTACTT GTTGATAGAC TTTCCGATTT TTCAGTACTG CTAGAGTTTG AAATAGAATC 720
 40 GCTTAATGAT GTTGATTTAC TAGCTGAATC CGACATGCTT GATGATACAC TTTGTGAATT 780
 CACTAAACTT GTGCTTGTTG AGCTTGATAC ACTATTACTT TCAGATGTGC TTAATGACTT 840
 AGATGCACTC ACAGAATCAG ATAGGCTTAC ACTTGTCGAT TTCGAGGTAC TAGCTGATGT 900
 45 AGATACCACA ATCGATCCTG ATGTACTCGT TGATGCACTT TGTGAGTCAG CTTTACTTGT 960
 TGACACACTT TGAGATTGTT GTGTACTTCC TGATGTTGAT ACGGAATCAC TCATGCTATT 1020
 50 TCTTGTTACT TCATATTTAA AAGTTGTCGT CGTTTTGTTA CCGCTCGCAT CTGTAGAAAC 1080
 GATTGATATA GTACTTGTAC CAATGTTTGT TGGTGTACCA CTAATAGTAT TATTTGTACT 1140
 ATCAAATGTT AGTCCGGATG GCAATCCAGT CACTGTATTC GTCACCGCAT TTCCACTGTT 1200

	ATTGGTGTCA CTGTTGGTGC TGTCGTATCC ACAACATTTA TTGTAAAAGT TGTCGTCGAT	1320
	TTGTTATTTG CTTGGTCAGT AGACACAAC TGCACGTGTG ATTGACCAAT TTTTGTGGT	1380
5	GTCCCAATGA TTGAATTCGT TGCACATCG TAACTTAATC CGCTTGGTAA TCCTGTAAC T	1440
	GTATTTGTCA CAGTCCCAGT ACCATTATCC GTTGTAGTCA ATACAATAGG ATTCATTGTT	1500
10	TTACCCACTT CTATGGTTTG ATTGCCTACA GTTACAGTTG GTGCTTTTAC ATCAGTAAAA	1560
	TAATATGTCA CTGATTGTCC AGCATTCTGC ATTTTTACAG TTTTATTTGT ATCATTATAA	1620
	GTTGACGCAT ATGAACTATC GACGGACGTG TAGTTATATC CTTTAGCAGT CAATGCAGAT	1680
15	TGCTGATTAT CGATTGTCA CACTTGATCA ACATTTCTCG AATATGTTTT TGGTGGAATA	1740
	ATATCTTTAC CTGTTGTTAC ATCAACGTAT CTCACTTGTG TAAACAGCAGA CTCTGTATAT	1800
	TCGAATGTTT CAAATTGTAC TTGTTGTAAA TTTGTGCGC CACCTGTTGA GGCTGTCATT	1860
20	GATAATGAAA AGTTGGTCGT ACCACTTTTC GCAATCCAAT CTGAAATATT ACGTGTCCAT	1920
	GTTTGACCTG CATATTTGAC AGTCATAACC TTTGTATCAC CATTATAGTT AATATCAAAA	1980
	TCTTGGAACG TGTTATTTGT AGGTTGACA TTTAACTTCG CAGCATTATC AGCTGTTGAA	2040
25	CTTGATGTAT ACGTTGTCGC AACACCATAA CTATCTGTTG TTACAAATGC ACCAAACGCA	2100
	CCTCCACCAG CTACATTAGA TGGGTGAGCA TTCGCCTTTG CAGCTGAATT TGGTTTAGAT	2160
	GTATTGTGAT ACGTATCCAA TTTGAAGCCA AATGCGTTAC TTAAGCCACC AATACCTACT	2220
30	GCGGCACCGT TTAACCTGT TTCACCTAAT ACACCTGGTG AAAAGGCAAA ACCGATACCA	2280
	TCTCCACCAT TTCCATGCCC TTCATATTG TTACCTAAAT TTAATTTTCC AGAAAAATGA	2340
	AAACTCTTAT TAGAGTCAAT ACGTGTTCCT AATGTAATAG CACCTTTTTG GCTGTATGCA	2400
35	TCCTGTGTTA ACGTCACAAT ACCGGTACTT TGATCATAGG TAGCATTACC TGACGTTGTC	2460
	ATATATTGTT TTAAGTTATC TTTATTAACT GTAATTGTAT TAGCAGTTAC TGCGGTTGTC	2520
40	GTCGCTGCTG ACGCAAATGT TGACATAGCT AAGCGACTGA AAGTTCGAAG TTTTACTGGT	2580
	GCGGTGCTAG TTGACGTTGT GCTAGTTTGT TTTAAGTTGA CCGAAGATGG CGTTGTGCTT	2640
	TGTGAAGTGT TATTTGATGC AGTACTTTGA TTTGTTGATG TATTAAATTG TTGTTCTGTA	2700
45	CTTGAAGTTG AAGCTACAGA TTTAGTATCA GAACTTGATG TAGTATTCTT TGAGGATGTT	2760
	GATTCTGATG TAGATGTCAA TTTCTCTTGT TGATTGCTTG TACTATTAGT TGTCGAAGTG	2820
	ACCTTTTCAG ACTTTTCACT TGAGACTGTG TCACTATTTG ATGTTTGTAC CGAACTACTA	2880
50	TTTTTCGTTA CACTTGTTGA ATCGGCTGTT GATGTTGATG CTTGATTGT CGTTGAGTTT	2940
	TGATTACCTA CTGTTTCACT TTGTGTGTTT AATTCAGAAG TTAATGGTGC ATCAGAAGCC	3000

55

	GTTCAGTC CGTATCCCGT CATTTTTTTA CTAATGCTTT GATTATCTTG ACTCACTAAA	3120
	CTATGACTAA TAAATGGTAG CCCCATAAAT TTGAACATTT CTATTTCTTT AATTCCGGAT	3180
5	TTTACCCAAT TTTTCCAGA TTTATAAAGT CTTACTCTTG TTTTTCGTT TGCTAAGCTG	3240
	TCATGAAATG CTTTCTGTCT TTTACTCATG TAATAACTCC TTGTATTATC TTTACATTCA	3300
10	TTAGATTATA ATATATGCCA CTATTCAATT TAATACAACT CTTTTTGAT ACAAAAATAC	3360
	TCATTTTGTT AAAATTTGTA AAAATTCaTT TTTATTCGTC TAAATGTAAT CGTTTTCATA	3420
	TTTTTAAAAT TACTTTTTCT CGTTTATGCG TATAATCTTT TTTTATATAA ATTTGGCTAA	3480
15	TTGGCTTTAT GTTTAATCAT TATAATTGTT TCGTTTTTAA AATAATTATT GTATTAATAT	3540
	ATCTATACCA TCCACCTTTT ATTTATAAAT AGTTAATTIA CAACTAAACG ATAAATATTA	3600
	TATGCAAAAT ACATCTTTAA TATTAAAGTA ATACCAATAT TTTTCAATA AACCTAGTGT	3660
20	AATATATGTG TAATTCTAAA AGATTCTTCT TTAAAAATAT AAATACCACG ACATATTGCT	3720
	TTAACATTTT CATTTATAAA GCGAAAAAAT GCATCGCTAC TAAGTTGAAT GTTTAGTAAG	3780
	GATGCATTGA ATTCACTAAA ATGATTAAAT TACTTATATC TTTTCATCTG ATTGATTATC	3840
25	GAAATTTCTT CCTTCTAAAC CTGCTAACTC TTCTTTAGAA GCTGCAGGTG CTTTCATTTT	3900
	AAATATCTCA TTCACTACTG TGTAATCGTA ATATCCTAAT CTGGCAATAG GTTTAATCGA	3960
30	CTTAATGTCC AATTTACCAT TATCAAGAAT AACCTTATCG TCAATATGAA CTTGGGCAAC	4020
	TCTTCCTATA ACAATATCTA CGGTAGATAC TGGATCTCCA GTTGAATAC GAATCGTTTG	4080
	AACGTACTCA CATTCAAAAT GAACTGGCGA TTCTTTTACA CGATATCCTG GAGCTTCTAT	4140
35	ACATTTTCC TTTGTTACAC CTGCAAAAT AAATTCATCC TCTTCTGGTG GCAATGCTTT	4200
	CGATGATAAA TTAAC TGCTT CTCTTAAATC ATACGTTGCC ATATTCCACA CAAACCAACC	4260
	TGTCTCTTCA GCATTTTCA CTGTATCTTT ACGTTCGTGA TCACCAAGAA CGGATTGATT	4320
40	TGCTGCGAAC ATAACCATAG GCGGATCCCA AGTTAAGTTT TGATACTGAC TATAAGGCGC	4380
	TAAATTATCT TTCCCATCTT TCGATACAGT AGAGATCCAC CCTATTGGAC GTGGTACTGT	4440
	ACTACTTTTA AATGGGTCGT GCGGTAAACC ATGACTTCTT ACACCTTGTT TTGGCGAATA	4500
45	ATTCATACTA TCTTCACCCC TTATAAGTAA TTACATTTAA GGTTACGCCC TCTTTACATA	4560
	AGCGTCTAAT ATAAATAAAC AATTTATTTA TAAGTAGAAA CTATATATGA CGTGGTTGCT	4620
50	TATAATTTGC GTTCTTGATT CGAAAAATC AGATAAGGAT TTATACAATT AATATTTATG	4680
	ATATCTTTTG TAAATTTAAT TAATTATAGT TACTTCAATC ATGATTAGTT TATAATAATA	4740
	AAGTGAAATT GAAAAAGACA GCTATTATGC GATGAGCGAA AACTTCAAG TAAAACAAGA	4800
55		

TAAGTTCAAA AAAGAATTCA AACCTGTTAT GCACTTAAAA GGTGATGCAT TCAATCAACA 4920
 GTTACAATCT TTGATTAACA AATATCCACA AATACAAAAA AATATGAAAT CAGAGTTCAT 4980
 5 TGCTTATTAT GATAAAGAAA AAAATAGAGA AACAGTAAAA AACTATGCTT GGAACCTTCA 5040
 AAAATCTATA AATGACATTA TGCAATCATA TCCTAGCACA AAATTGTAC AGTTTTATAA 5100
 AAGATGATGT TTCCCGTCA ATGGTAGATG GAAATGGCCG TTTAAATCG GGATACTAAT 5160
 10 GTATTTCCAT C 5171

(2) INFORMATION FOR SEQ ID NO: 265:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3589 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

CTACACACTA AACCTATTTT AGTTATGGGT GGTACAGTGA TTCTCTTTTC ATTTTAAATA 60
 25 GGTATTTGGA TTGGTCATCC TATTGAAACA GAAATCAAAC CACTTATTAT TGGTGCAT 120
 ATTATGTACG TACTTGGGCT TGTAGATGAT ATCTACGATT TGAAACCGTA TATAAAATTG 180
 GCTGGTCAAA TTCCCGCTGC CTTAGTAGTT GCTTTTATG GTGTGACTAT TGATTTTATT 240
 30 TCGTTGCCAA TGGGTACAAC GATTCATTTT GGATTTCTTA GTATTCCAAT TACTGTGATT 300
 TGGATTGTTG CTATTACAAA TGCAATTAAC TTAATTGATG GACTCGATGG TTTGGCGTCG 360
 35 GGTGTTTCKG CAATCGGACT CATTACAATA GGGTTCATTG CAATTTTACA AGCTAATATT 420
 TTCATAACGA TGATTTGTTG TGTTTTATTA GGCTCTTTAA TTGGGTTTTT ATTTTACAAT 480
 TTCCATCCTG CCAAAATATT TTTAGGTGAT AGTGGGGCTT TAATGATTGG ATTTATCATC 540
 40 GGATTCCTTT CTTTACTCGG ATTCAAAAAT ATTACAATTA TTGCATTGTT CTTCCCAATT 600
 GTTATCTTAG CAGTTCCATT CATTGATACT TTGTTGCAA TGATTGACG TGTGAAAAAA 660
 GGGCAGCATA TAATGCAAGC TGATAAATCG CATTTGCATC ATAACTATT AGCTTTAGGC 720
 45 TACACACATA GACAAACAGT ATTATTAATC TATTCAATCT CTATTTTATT TAGTCTTTG 780
 AGCATTATTT TGTATGTATC GCCACCATA GGTGTTGTAT TAATGTTTGT ATTAATCATA 840
 TTTAGTATTG AATTAATTGT TGAATTTACA GGATTAATAG ATAACAATA CCGACCAATA 900
 50 TTAATTTAA TTAGTCGTAA GTCATCTCAT AAAGAGGAAT AGGGAATGAA AGCATAGCTG 960
 TATGGGATAA TTTGTATTAT ATGGCTTTAC TCTTTACAAT TTTTTTGTAT TAAATTTCAA 1020

55

	ATTACCGTC TTATGATAGT GCTTTTATT TTTATTCAGT TGGTATATCG AAAGGTAAC	1140
	GCTTTGGAGT TTCTTCAGTC AAATCGAAAT TTCCTGCAGT CATTTGATTT AAAAAGTTAA	1200
5	TAAACGCTTC ATAGTCACTT TTAACGACAT CGATATAGTA GCTTACCTTA TCAGTGTAAG	1260
	TTTGTTTTCT TAACATAAAA TGAGTTGAAG CTAATTCATA TTCAAATTTA CCAGTTTGAT	1320
10	CATAATTCAG TGTTACTATA CATGGTACTG CTTCTCGTAG TTCGACACGC CCGATATCAT	1380
	AAATGACGTC TCTAACAGCA CCGCTATAGG CGCGAATTAA ACCGCCACCA CCTAATTTAA	1440
	TACCACCAAA ATATCTTGTT ACTACGACAC ACGCATTATG aACATCGrGC TTTTTTaATA	1500
15	TGTCTAACAT TGGGaCACCG GcAGTTCctG TCGGTTcACC ATCATCATTc GChTTTTGAA	1560
	TATTCATTtC AGGTCCAATA GTATATGCAG AACAATTATG AGTGGCATCT TTATGTTCTT	1620
	TTTTTATTGC AGCAATAAAT GCTTTaGCTT CATCTTCATT TTGAACAGGT TTGATATGAG	1680
20	CAATGAATCT TGATTTACTA ATCACAATTT CAATAATGTG TTCTTTTTTA ACAGTAATGA	1740
	TATTTTGTGT CATAATAACT CCTTAATTCA TAAGCTTAAG ATTATTTAAT CTTCAATTATA	1800
	CACTGAAAAT GACATGACTA TAAATCGTTT GATTGCCATT TTCTTTTTTA CTGAAATATT	1860
25	GTATCATTGC TATGAGTATA TTTTAGGAGG ACGACTATGA AAATTGCTGT GATGACCGAT	1920
	TCTACAAGTT ATCTGTCGCA GGACTTAATC GATAAATATA ATATTcAAAT AGCGCCATTA	1980
30	AGTGTGACTT TTGAAGATGG CAAGATTATA CCAGAAGAAA AAGTTCGTAC TAAAAAGCGT	2040
	GCCATTCAAA CATTAGAAAA GAAAGTATTA GATATTGTAA AAGACTTTGA AGAAGTAACT	2100
	TTATTTGTCA TAAATGGAGA TCATTTGAA GATGGTCAAG CGTTATACAA AAAGTTACAA	2160
35	GATGATTGTC CTTCAGCTTA TCAAGTAGCA TACTCTGAGT TTGGTCCAGT TGTTGCAGCA	2220
	CATTAGGTT CTGGTGGATT AGGTTTAGGC TATGTTGGCA GAAAAATAAG ATTAACATAA	2280
	TTATAAAATT TTAATAAAAG AGTCTATATT GTAATTGGAA ATTATCTCTC GTATACATGG	2340
40	CTTTAAATGT TCATCATTTG AAAGCCAAAA TGCTAAAGAT ATAAGAAAAT CATTATAATA	2400
	TTAGGCTCTT TTTTACGTTG AAATGAGGTT TTAAGCATTa AACATTACGG GAAATTAATT	2460
	CATCCTCATA CTTCACTTAC TAATGAAAAA ATTAAAAAAG AAGTAACAGG TGTCATCAAA	2520
45	CAAAATTCAA ACTATTATTG TGTTCAATGT GAAAGTACAA ATCCAAAGCA TTTTATCAG	2580
	TATGATTCTT CAGTACATTc CAAGAAAATT GTATATTGCA GAAATTGTAT ATCACTGGGT	2640
50	CGAATGGATA ATGTAACAAG ATATAAAATA ACAGAGAGTT CGCAAAGTTC ATCACAAGCA	2700
	TATTATCATC TCTCATTTGA ATTGTCGGAA CAGCAGTCTT ATGCCTCAGA ACATATTGTT	2760
	CGAGCCATTA GAAAGAGACA AACGATTTTG TTATATGCCG TAACAGGTGC AGGTAAGACA	2820

55

TCACCACGTG TAGATGTTGT TGTAGAAATT AGTAAACGTA TTAAAGACGC ATTTCTTAAT 2940
 GAAGATATAG ACATACTACA CCAGCAATCA AGACAACAAT TTGAAGGGCA TTTGTGTGTA 3000
 5 TGCACAGTGC ATCAACTTTA CCGATTCAAA CAGCACTTTG ATACTATTTT TATTGATGAA 3060
 GTCGATGCCT TTCCTTTATC AATGGATAAA AATTTACAAC AAGCATTGAA GTCATCTTCT 3120
 AAAGTTGAAC ATGCAACAAT TTATATGACA GCAACACCAC CGAAACAAC TCTGTCAGAG 3180
 10 ATTCCCCACG AAAATATAAT TAAATTGCCA GCTCGCTTTC ATAAAAATC ACTTCCAGTT 3240
 CCTAAATATC GTTATTTCAA ACTTAATAAT AAGAAGATTC AGAAAATGTT ATACCGAATT 3300
 TTACAAGATC AAATTAATAA TCAACGTTAT ACACTGGTGT TTTTAAACAA TATAGAAACA 3360
 15 ATGATTAAAA CATTTTCGGT TTATAAGCAG AAAATTACTA AATTAACATA CGTCCATAGC 3420
 GAGGATGTTT TTCGCTTTGA AAAAGTTGAA CAATTAAGGA ATGGACATTT CGATGTCATT 3480
 20 TTTACTACGA CAATATTAGA ACGTGGATTT ACAATGGCAA ATTTGGATGT TGTGTTTATC 3540
 GATGCACATC AATATACTCA AGAGGCTTTA ATACAAATTG CTGGACGTG 3589

(2) INFORMATION FOR SEQ ID NO: 266:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1017 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

35 TTTCCAAGAA GGcTTgAAAA AtGTTwCaAG TGGTGCGAmC CCAGTTGGTT TACGACAAGG 60
 TATCGACAAA GCAGTTAAAG TTGCTGTTGA AGCGTTACAT GAAAATTCTC AAAAAGTTGA 120
 AAATAAAAAT GAAATTGCGC AAGTAGGTGC GATTTGAGCA GCAGATGAAG AAATTGGACG 180
 40 TTATATTTCT GAAGCTATGG AAAAAGTAGG TAACGATGGT GTCATTACAA TTGAAGAATC 240
 AAATGGACTA AACACTGAAC TAGAAGTGGT TGAAGGTATG CAATTTGATC GTGGTTATCA 300
 ATCACCGTAT ATGGTTACTG ATTCAGATAA AATGGTTGCT GAATTAGAAC GCCCATACAT 360
 45 TTTAGTAACA GATAAGAAAA TCTCGTCTTT CCAAGATATC TTACCTTTAT TAGAACAAGT 420
 GGTTCAATCT AATCGTCCAA TCTTAATTGT AGCTGATGAA GTTGAAGGCG ATGCATTAAC 480
 AAATATCGTG CTAAACCGTA TCGTGGCAC ATTTACAGCT GTTGCAGTAA AAGCACCTGG 540
 50 TTTTGGTGAT CGTAGAAAAG CGATGCTTGA AGATTTAGCT ATTTTAACTG GTGCGCAAGT 600
 GATTACTGAT GATTTAGGCT TAGATTTAAA AGATGCATCA ATTGATATGT TAGGTACTGC 660

55

CAGCATTGAT GCACGTGTTA GCCAATTGAA ATCTCAAATT GAAGAAACTG AATCTGACTT 780
 TGATCGTGAA AAATTACAAG AGCGCTTAGC TAAATTAGCA GGTGGTGTG CAGTTATCAA 840
 5 AgTAGGTGCA GCAAGTGAAA CAGAGCTTAA AGAACGTAAA TTACGTATTG AAGATGCATT 900
 AAATTCTACA CGTGCAGCAG TTGAAGAAGG TATTGTTGCA GGTGGTGGTA CTGCATTAGT 960
 10 AAATGTTTAC CAAAAAGTAA GTGAAATTGA AGCTGAAGGT GACATTGAAA CAGGTGT 1017

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1409 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

TTAATCCAGC GTTAACTGTA TTTGCATTTA TTATGATTAT TCGATTCTT TTAGCGTATG 60
 TATTTAAATG GCTTGGATTA GTGGATGATG TGTATTAAAT GGTCATTATC ATTTCAACTA 120
 25 TTTCCTTAGG CGTAGTTGTT CCAACTTTAA AAGAAATGAA TATTATGAGA ACAACTATAG 180
 GGCAATTTAT CCTATTAGTA GCACTACTTG CGGACTTAGT AACTATGATT TTATTAACGG 240
 TCTATGGCGC AATCAATGGT CAAGGCGGCA GTACAATATG GTTAATAGGT ATATTAGTTG 300
 30 TTTTCACAGC AATTTTCATAT ATTTTAGGTG TTCAATTTAA AAGAATGTCA TTTTACAAA 360
 AATTGATGGA TGGTACGACG CAAATCGGTA TTCGTGCGGT ATTTGCATTA ATAATATTAT 420
 TAGTAGCCCT AGCAGAGGGA GTTGGCGCAG AAAATATATT AGGTGCATTC TTAGCAGGTG 480
 35 TCGTTGTTTC ATTATTAAAT CCAGATGAAG AAATGGTTGA AAAGTTAGAC TCATTGTTT 540
 ATGGTTCTT TATTCCTATT TTCTTTATAA TGGnTGGTGT AGATTTAAAC ATACCTTCAT 600
 40 TAATTAAAGA ACCGAAATTA CTAATTATCA TACCGATTTT AATCGTnGCA TTTATCATTT 660
 CAAAATTAAT TCCAGTCATG TTTATTCGAC GTTGGTTTGA TATGAAAACA ACGATTGCAT 720
 CAGCATTTTT ATTAACATCA ACATTATCGC TCGTGATAGC TGCAGCCAAA ATTCAGAAA 780
 45 GATTAAATGC TATTTAGCT GAAACGTCAG GTATATTAAT TTAAAGCGCA GTCATTACAT 840
 GTGTATTTCG TCCGATTATT TTCAAAAAAC TGTTTCCAGT TCCAGATGAG TTTAACCGTA 900
 AAATTGAAGT TAGTTTAATT GGTAAAAATC AATTAACGAT TCCTATAGCG CAAAATTTAA 960
 50 CATCTCAGTT ATATGACGTG ACATTATATT ATCGCAAAGA CTTGAGTGAT CGTCGTCAAT 1020
 TGTCAGATGA TATCACGATG ATAGAAATTG CTGATTATGA ACAAGATGTT TTAGAACGAC 1080

AAGTTGCTAA ATTAGCCAAA GCACATCAAG TTGAGCGTGT CATTTCGAGA CTTGAAAGCA 1200
 CAACGGACGA TACAGAGTTA GTTGATTGAG GTATTGAAAT TTTCAGTAGC TACTTAAAGTA 1260
 5 ATAAAATCTT ATTAAAAGGT TTAATTGAAA CACCTAACAT GTTGAATTTA TTAAGTAATG 1320
 TTGAAACGTC ACTATATGAA ATTCAAATGT TAAATTATAA ATATGAAAAT ATTCAATTAC 1380
 GTAATTTCCC ATTCGGAGGA GACATCATC 1409

10

(2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4702 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

20

AAAGAGGGTT TTTTACTACT ATAATCCATC CTTTAATGGA ATTCCATTG TCCCTTTGGC 60
 CCGTTCATAA CCATAGGAAA ATATATGCnG NAATCATTGA TAGTAAATAC ATCAATAATA 120
 25 CGAATAATGT CCGTGAAATC ACGTCCGTAT TACCATTATT AGCTAAAACA TTTTCCAAGT 180
 TTTCTTTTGA ACCGGATACT CATAAATGCT TTTAATGChT GGTTTTGTCT GTCGCCATCT 240
 TTAGAATTTG TAATACTTTT TTAAATTCTT TAGCGAACAA CTCATTATCT TTATCGTTTT 300
 30 TAGCCATTTG ACGATTCAAA TCGTTAGCTC TTACGCCTAA TGCTCGACCT GTCGCATGTA 360
 GCTTATTCAC ATTATTGTCG ACTTGATTGA ATTGGCCACT AACAGAATCT GCAATTGATT 420
 TTGATTCTTG TGTATCTGAT AGCAATTGCG TACTCTTTTC AGAAATTCTA CTAATTTCTT 480
 35 TATCTAAATT TGAAGACATC GTATTAAATT CATCATTTTT GCCTTTATCA ATTTTTGGTT 540
 CTTGTGGCTC TTCAGCAAAA GTCTTTTTTAA CGTTTTCTAA CTGATCAATC AGCTTGGA 600
 40 TATCTTCTTk ATTTTTTGTT gTATTCTTTT TGTTATTTAA AATGTCATCA ATCAGTTTGT 660
 CTGAGTTTTT TCATTGAA TCAATTTGaT GTAACACAGC TACTTTATCG TCTTTGAAAC 720
 TTTCCATGTC ATTGATAACT TGGTCAACCA TCATATCAAT TAAACGTTTG TTGTCGAATG 780
 45 GTTTATCTTC tCTGCCTTTT GTATCTGTGT ACATTTTATA ATGCGCATCA AACCTAGATA 840
 ATGCACTCAA TTGCTGGCTT AATGCATCTT TCGATAAGCG ACCATCaAGG TTATGATTCA 900
 ATGTTACATC CACAACACTC GTTGCTTTCT TATCATTGGG TTCATCTTGA CGATTGCTT 960
 50 GTCCAAATAA CAATTGTAAA TGCATTGTTT TATCTTTTAA GAAATCTTTC TCAGCATCCT 1020
 TTTTCAATTT AGCAACGCCA TTGACTTCAA CTTTATATTC CTTGTTAGAT GTATCGAGTT 1080

55

	TAAAATGAGG AtCTGTTGCA ACAGTTAATT GATTAATATC ATTACTTTTA ATCGTTTCAG	1200
	TACGTTGCAC TTTGACACCA TCATTAATCA AACTACTTGT GTCTTGCGCA ACTTGTTTAT	1260
5	CGTAATCTGT TAAATTAATG TGTTCTGCTA ACGGTTTTTT CAAATTATAT TCATTTTTAT	1320
	AACGTTTTGC TTCTTTGACA ATTGCTTCGT ATTTATTAGC TTCATCCTCA TTTAAACCTG	1380
	CAGCTATAAA GTCTTGTTTA GACATGTTAT AGATAAATGT TGTATCTGTA TCAGTTCTTT	1440
10	TGACAAATATC ATCATGAAGT TGTTTCTCTA AGTTTTTCAGC GAATTGAGCA TTGTTCAATT	1500
	TAATGCTATT TAGCGCATCT TGTAAGTCTT TGTTATTTGC AAGCTCATCT TGCAGTGATT	1560
	CTGTTAATTG CTTACGATAT TCTTCAATCA TACCTTTTGA AAATGGTGAC TCTTGTGATT	1620
15	GAATGATTTT TCTTAATTTA TCTAAGTTTT CTTTAACAGT TTGTTTATAT TCTTCTTTAC	1680
	CTGTATCTTG CATACTTGAT TGTTGATCAA TTTGGCTGTC CATCTGTTTT AATGCATTGA	1740
20	TATAGTTATC AAGTTCCACG CTATCTTTTT GCGATTTATA ATCTTGTAAC ATTTTATCCA	1800
	TCGCTGTATT GTGCTCGTCA AATAATGAAT TTTGTTTTTC AATTAAAGTC GAAACATTAT	1860
	AATCTGTGTT CACTCTGAAC GTATCTGAAT TCGCACTCAA TAATGATTTA TTGTATGTTT	1920
25	GGAACCATTT TGTAATGTCT TTGTTTGCAG AAATTGAATT TACAAGCGTA TCTGTAAATA	1980
	ATTCCGGGAA GTCGTTAATT GGATTTAATA AGTAATTCGA GAATTTACTA TTCACACCAT	2040
	GTTCAACGGT CATAATAGCG CCAACATTTT TTTGTGCATT ATGTAAATTA TCAATGATGC	2100
30	TTGTTAAATA AATTCGACT AAGTTTTTGT TAAAGTCGTT AAGTACATTA CTTACAACCTT	2160
	TTTCTGTGTT TTTAGCTACT TCTTCTTTTT GTCCTACAGC TGTTTTATAC TGTAGCGATA	2220
	TTTTCGATGG TGTTTTAGCG TCTAATTGCA TTGCCAATTT TGAAAAGTTT TCTGGGATAA	2280
35	CAATCATGAC TTGGTATCCA CCATTTTTC AACCAGACTC AGCAACGTTT CTTGTTACTG	2340
	TTTCAAATTT ATAGTTTTTC TCATTTGCTA ACCTTTTAAT AAATGCTTGA CCCAGCTCAA	2400
40	CTTTTTTACC GTTATATGTC GTTGGTTGAT CCTCGTTAAC AATTGCGATA TGTATTTTAT	2460
	TATTTTTTATT ACTTACACTT TGGGATCCTT TTTCTGATTG ATCTCCATAT TTTGTTTGAA	2520
	CAAAAAATAT CATACTAACT ATGGCAATTA TAATAATTAA AGTGACAATT AATGCATAAA	2580
45	TCCAATTTTT CTTTTTCATG CTTATTTCTT TTCAGTTGTT TTCTTAAAAA AATGATAAGC	2640
	AAAGCCACAT TAGAAAATGT GACTTTGCCA ATTTTCAGAAT GCTTATTGCA AACCGAAATT	2700
	ATTAGAAAGT TGTTGGTCTT GTTCTTGAAC GGCATCAGCA GTGCTATTCA ATTGTTGTTT	2760
50	AATTTCTTCT AATAATTGTG CAAATTTTTT TACTTTAGGA CTAAGTTGTT GGAATTGCTC	2820
	TTCGAAACGG CTGAAAGCTT GACCTTCCCA GTTCGCTGCA ATTTACCTT GTGCACGKGT	2880

55

	TCTGATTTCC	TCTGGACTCA	TCTTAATCAT	TGCCATAACT	AGAAACCTCC	TGAATATTTT	3000
	AAGTTTATCa	AAACTTTTTA	GGGACACTAT	TTTTTGAAAA	AGTGCTCCTT	ACTCAAATAA	3060
5	TATATAAATT	ATTAGTATAT	GTATATAGTt	TTTTAAGTAT	TTTtagCTTT	TTTAAATAAA	3120
	ATATATTGAA	TATAACCATA	TATTTTAAAT	TAACCATTCA	TTTTTGTAAT	ATAAATGTGT	3180
	ATACTAAAT	TAAATTAAAT	ACATAAAGGA	TTAAATGGTT	ATTATGAAGA	AAACAATTTT	3240
10	ACTGACGATG	ACAACTCTTA	CTTTATTTAG	TATGTCGCCT	AACTCGGCTC	AAGCATATAC	3300
	GAATGATAGC	AAAACATTAG	AAGAAGCAAA	GAAAGCACAC	CCAAACGCAC	AGTTCAAAGT	3360
15	GAATAAAGAC	ACCGGCGCGT	ATACTTATAC	ATATGACAAA	AACAACACGC	CAAACAACAA	3420
	TCATCAAAAC	CAGTCACGTA	CAAACGACAA	TCATCAACAC	GCAAATCAAC	GTGATCTTAA	3480
	CAACAATCAG	TACCATTCTT	CATTAAGTGG	TCAGTATACG	CACATTAATG	ACGCAATTGA	3540
20	TTCACACACA	CCGCCTCAAA	CGTCACCAAG	CAATCCTTTG	ACACCAGCAA	TACCGAATGT	3600
	CGAAGACAAT	GACGATGAAT	TAAATAACGC	TTTTTCAAAA	GATAACAAAG	GGCTTATTAC	3660
	AGGCATCGAT	TTAGACGAAT	TGTATGACGA	ATTACAAATC	GCCGAATTTA	ATGACAAAGC	3720
25	AAAGACCGCT	GACGGTAAAC	CTTTAGCATT	AGGTAACGGT	AAAATCATTG	ATCAGCCTCT	3780
	TATCACAAGT	AAGAACAAC	TATATACTGC	TGGACAATGT	ACATGGTATG	TCTTTGATAA	3840
	ACGTGCCAAA	GATGGACACA	CGATTAGTAC	ATTTTGGGGA	GATGCTAAAA	ACTGGGCAGG	3900
30	CCAAGCTTCA	AGCAATGGCT	TCAAAGTAGA	TAGACACCCA	ACACGAGGAT	CAATTTTACA	3960
	AACAGTAAAT	GGTCCATTTG	GTCATGTAGC	CTACGTTGrA	AAAGTTAATA	TTGATGGAAG	4020
35	TATTCTAATT	TCAGAAATGA	ACTGGATTGG	TGAATATATC	GTTTCATCAA	GAACCATCTC	4080
	TGCTTCAGAA	GTTTCATCAT	ATAATTACAT	CCATTAAATT	AATCATGACA	TCAATAAAAA	4140
	GCGAECAGTT	CGCAGTTTAC	AATTCGTAAC	ACTGCAAAAT	TGGTCGCTTT	ATTTTGTATG	4200
40	TTATTCGATT	ATAAAATTAC	AAAGAAATGT	TCTCTACATT	CCCCATTAAT	CAAAATCGTT	4260
	TACGAAAGTA	TAATTGTAGC	TATAATAATC	CAAGTCGTAA	CAACTAGTGG	CACTATCGTC	4320
	TTGAATAAGA	ATATACCGTA	TTTTTCTTG	CGATATATAT	CCAGTACTAG	CCAAATTAAA	4380
45	ATGATTATAA	CACCAACAAA	AATAAATACA	GGATTCATCG	ATATAGCATC	TGCCTGTAAC	4440
	TCAGGTTGCA	TTCTTAATTT	AGTGATAATT	AACATCACTA	CTGAAATAAT	GAAAAAGTAG	4500
	ATACCTCTTA	TCTTTGATGT	CTGTAAATCT	AATTCCTGCT	CTTCAATGAC	CTCTTTAGAT	4560
50	TCACCCAATT	CTTTTGCAAT	CAAATAATTT	ACTACCTTAG	GTTTCACCCA	TAAACACTTA	4620
	ATTGCAAAGT	ACATAAAAAAT	ATATGATCCA	GATCCATAA	ACATTAAAAA	GTTGCTTAAA	4680
55							

(2) INFORMATION FOR SEQ ID NO: 269:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2004 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

ACAAAAAATT CACCCTCATT AATATTGAAA CTAATGTTAT CGACAGCAAC ATGTTTGCCA 60
 TAACGCTTAG TTACATTTTT AACTTAATC ACTTGCCAC CTCTTTTTTT CTCATAGCAT 120
 AAAACCGAGA TTATATGTAT GTATTCCCTA TTTAACCACG TTTATTACAA TTTTCAAATT 180
 TAAATGATTT ATCCTTGAAC TTTTTAATA AAATAATGAA TAAWAGGWAA TCWCCAGTTA 240
 AGAAATAGTG TTATTTTACC TTGAATTCAA AAAACACCCC AGTAAAACAA GGAATGCTTA 300
 CTAGGTGTCT TCACTATACT TTGGCTTTAT AATTTTGAAT CGTTCTAAA AATGCTGGAC 360
 AATAATGTTT TAATTTGTAA CTACCTACGC CATCATATATT AATCATATCT TGTTCGAAG 420
 CAGGCTTACG TTTAGCAAAT TCCTCCAACG TGTAATCAGA AAATATACTT ACAGGTGCTA 480
 TCGTTAATTT GTCACCTAAC TTTTACGAA CTTCTACCAA CTGACTGAAT AATACTCGGT 540
 CAACCCCTTC AACCGTATTT ATAAATACTT TTTCACTCGC TTTTGTCTTA AATGGTGTTG 600
 TGAATACTTC TACTTCATTA CTGAGTAATT TTTTAATTGA AGTATCACAC ATTAATATTT 660
 CGTCATTTTC ATTAAAGAAC CCTTTGAATC TTAATTCATC TATTAAGTGA CTTAATTCTG 720
 ATGTTGTGTA ACCTTTCATT AAACCATGGG TTGAAATTTG GTCATAACCT TTATACTTAA 780
 TATAATCTGK TGACTCTCCT CTTAACACTT GAATGATAAC ACTATAACTC TCTTGTGTGT 840
 TCATACGAGC GATGCAACTA ATAATCATCT TAGCTTCTTG TGTATATTA TATGATTTAT 900
 CTTGTTGAAC ACAATTACTA CATGTTTCAC ATTCTTCTAA TTTTTCATTC GGTTCAAAAAT 960
 AATGGACAAAT TGTGCTTCT AGACATTTTT TTGTTTTTGT ATATTGAATC ATTTTAGTTA 1020
 ACTTTTCGCC CATTTTATCT TTATAGTCAT CATCAGCTTG AGAGACTGTT ATAAAATACT 1080
 CGTGTAATTT GATAACGCGT TCGCTAAATA ACAAATACA TTCACTTTTC AACCCGTCAC 1140
 GACCTGCACG ACCCGCTTCT TGATAATAAG ATTCTAAATC TCCAGGCATA TTATAATGAA 1200
 TAACAAAGCG TACATTGGAT TTATCAATAC CCATACCAAA AGCATTTGTA GCAACGACTA 1260
 CTTTAACACG ATCAAATAAG AAATCATTCT GCGCTTCTTC TCTTCTTTA TTGCTCAAAC 1320
 CTGCATGATA TATAACACTT TCAATTTTCT GACTTCTTAA GGCTTCTTGA AGCTCTTCAA 1380

CCAATATAAA TTTTGACGT TGATAAGTAG GATTTACTTT AAAAATTAAG TTTCTACGCT 1500
 TAGTACTCGT TTTAATTTGA TCAGTTTGAG CGATATTTAA CTTTCTCTA ATATCTTGCT 1560
 5 GTACTTCAAC CGTGGCAGTT GCTGTCAACG CTATTATTGT AAAATCTTGA GGTAACGTAA 1620
 AACTTTTGA AATAACATTT TGTAACCTCG GCCTGAAATC ATGACCCCAT TTAGAAATAC 1680
 AATGCGCTTC ATCAAACGCG ACTAAGTGAA TCTTTATACG CTGAAGCATA TTTAAAAAAT 1740
 10 ATCGGTTTTTC AAATCGTTCT GGTGCAACAT ACAAAAATTG AATTTCTCCA TTTGATAATG 1800
 CTTTTTCAAT ACGTTGTTGC TCTTTTGTAG TCAAACTACT ATTTAAAAAA GCAGCTTGAA 1860
 TTCCCATCGC TTTAATTGA TCCACTTGAT CTTTCATTAA TGATATTAGT GGACTTATTA 1920
 15 CAATTGTTGT ACCACCTAAC mATAAACCTG GTACTTGtTA GcmtATAGAC yTACCTCCAC 1980
 CAGTtGGkAA GrCACCAAGC ACAT 2004

20 (2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2244 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

30 AAAGATTGCT TGCCTTGAGG GTTTATATAT CTGACTCAAT TGCCACATTT TTATCAAGAG 60
 TAGTTGATAA TACTCATCAT AATTATAGCT AATATTATAT TTTTTTAAAA GATAGTGTAT 120
 GATTTTCTGG TGTGTGTTGT ATACGTCATT AAATTTCAAG TAGTCATTCT CCAAGTTATA 180
 35 CGTATAACAA ATATTTCCCG ATAAAGTTAG AATAAAATAT TTAGAAAAAT CATTCAATTG 240
 CGTAAATCGCT AAATTAAGTG TTAAATATAA GACATAAGTA ATTAATTTAA TGTGATATGA 300
 TGTATTATTA CTTTGCTAAA TAGTAGATAG AACAAAATTT GTAATCGGGA GGTAACAATG 360
 40 GATTACGCAC ATTTAAATTT AGAACATTTT TTTGCACGAA ACGACGATTT AGATGTTATA 420
 AGAGATCGCG CTGATTTCTG GATGATAAAT AACTTCACTA ATGAAATGAT GTATCGTGAT 480
 45 GGTCAAATTG AAGGCACGAT TGATTTAAAT CAGTACTATT ATAAAAATAG ATCAAATGCA 540
 GCAAGTTTTA TTATGATGGA TTATAAAAAA GAACTAAGT AAACGAACAA AAGAATTTTT 600
 TGTTTTTTAA TACGTGAATA ATAAGATTAT TGATATAAAG GTTTTCAAAG GTTATACAAA 660
 50 AAGATAAAAC ATTTATGATT CGTAGATCAA CGTAAAGTAA TGTTGATAAA TGGTTTAAAA 720
 CGTTTCATTT ACATTACTGT TTATTTATGA ATATGTAACA ATGCATAGAT AAAATGTGTA 780

ACCTAAGAGG TGTGGATATG AATAAACACA AGAAAGGTTT TATTTTTTGA ATAATAGGAC 900
 TTGTTGTCAT ATTTGCTGTT GTCyCaTTTT TATTTTTCTC AATGATATCC GATCAGATAT 960
 5 TTTTCAAACA TGTAAATCC GACATTAAGA TTGAAAAGTT AAATGTTACA TTAAACGATG 1020
 CAGCAAAGAA ACAAATAAAT AATTATACGA GTCAACAGGT ATCAAATAAA AAGAATGATG 1080
 CATGGAGAGA TGCATCTGCA ACTGAAATTA AAAGTGCAAT GGATAGCGGT ACTTTTATCG 1140
 10 ATAATGAAAA GCAAAATAT CAATTTTTAG ATTTATCAAA GTATCAAGGG ATTGATAAAA 1200
 ATAGAATTAA ACGTATGTTA GTAGATAGAC CAACGTTATT GAAACATACG GATGATTTCT 1260
 TAAAAGCTGC TAAAGATAAG CACGTTAAGC AAGTTTATTT AATTTACAT GCATTATTAG 1320
 15 AAACCTGGCGC AGTTAAAAGT GAATTAGCTA ATGGAGTCGA AATTGATGGC AAAAAGTACT 1380
 ACAATTTCTA TGGAGTAGGA GCCCTTGATA AAGACCCAAT TAAACAGGT GCAGAATATG 1440
 20 CTAAAAAGCA TGGTTGGGAT ACACCTGAAA AAGCTATTTT AGGCGGTGCT GATTTTCATTC 1500
 ATAAGCACTT CTTATCAAGC ACAGATCAAA ATACATTGTA TAGTATGAGA TGGAAATCCAA 1560
 AAAATCCAGG AGAACATCAA TATGCTACAG ATATTAAAGTG GGCAGAAAGT AATGCAACAA 1620
 25 TTATCGCTGA CTTTTATAAG AACATGAAGA CTGAAGGAAA ATACTTCAAA TACTTTGTGT 1680
 ATAAAGATGA CAGTAAACAT TTGAATAAGT AATTTGATAA GCTACGAGTT GTTTTTATGA 1740
 CTCGGACATA CTAAAAAGAC GCTTTCTATC TTGTTTTGAT AGAAAGCGTC TTTTTGCATT 1800
 30 AGAGAAAACA CATTGATkGA TAAtCCCaCC aATGCAAgTG GGGcAGGACa TCGATAAAGA 1860
 ATTACTTTTT CTTTAGAAAT TAGTATTTCT TATGCATGAG TTTTACTCAT GTATTCCTAT 1920
 TTTTAAGTAC ACATTAGTTA TAGCTAATGA TAAAGAACCA CTACATAATA AATCATTAGT 1980
 35 GTTTTTTTAT CATTTCTGTC CCaCTCTCAT CGTGATTGA AATTTTCAAT TCGGATTTTA 2040
 ATTTTCATCTC TTACACGTTG GAACTCTGAC CAAGGCTTGC CTGCAGGATC ATCAAATCCC 2100
 40 CAATGTTCTT TCTTAACATT TGTGGTAAA GAAGGGCAAT TTACGTCTGC ATCACTACAT 2160
 AATGTAACAA CTAAATTTGA ATTTnTAATA ATATTATTAT CGGATTAAAA TCTGATGGAT 2220
 GATTTGATAT ATCAATGCCT ACTT 2244

45 (2) INFORMATION FOR SEQ ID NO: 271:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1371 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

ATAAGCAATT TAATTTTGAG TCTACAATGG AGGAATTATC ATCTTTATCA GAGACTTGCC 60
 AACTTGAAGT GTTGGGTCAA ATTACTCAAA ACAGAGATCG TGTAGATCGC AAATATTATG 120
 5 TTGGTAAAGG TAAAATTGAA GAAATTCAAG CATTTATTGA GTTCAAAGAT ATTGATGTAG 180
 TCATCACAAA TGATGAATTA ACGACTGCAC AATCCAAATC ACTAAATGAA GCTTTAGGTG 240
 TAAAAATTAT TGATAGAACT CAGTTGATTC TTGAAATATT TGCATTAAGA GCAAGAAGTA 300
 10 AAGAAGGTAA ATTGCAAGTA GAGCTAGCAC AACTTGATTA TTTATTACCT AGATTGCAAG 360
 GCCATGGTAA AAGCCTTTCT CGTTTAGGTG GCGGTATTGG AACTAGAGGC CCTGGTGAAA 420
 CGAAGTTAGA GATGGATCGC AGACATATTC GAAC TCGTAT GAATGAAATT AAACATCAAT 480
 15 TGCGGACGGT AGAAGAACAT CGCGAAAGAT ATCGAAATAA AAGAAATCAA AATCAGGTGT 540
 TTCAAGTAGC TTTAGTTGGT TATACAAATG CTGGTAAATC ATCATGGTTT AATGTTTTAG 600
 20 CAAATGAAGA GACGTATGAA AAAGATCAAT TATTTGCAAC GTTAGATCCT AAAACACGAC 660
 AAATTCAAAT AAATGATGGA TTTAATTAA TTATTTTACA TACTGTTGGT TTTATACAGA 720
 AACTACCTAC GACGTTAATT GCAGCTTTTA AATCAACTTT AGAAGAGGCT AAAGGTGCAG 780
 25 ATTTATTAGT ACATGTCGTA GATAGTAGCC ATCCTGAATA CCGTACGCAG TATGACACAG 840
 TTAATGATTT AATCAAACAA TTAGATATGA GTCATATTTT TCAAATAGTT ATTTTAAATA 900
 AAAAGGACTT ATGTGATCAT GCATCAAATC GTCCAGCAAG TGATTTGCCT AATGTTTTTG 960
 30 TTTCTTCTAA AAATGATGGT GATAAATTAC TTGTTAAGAC GTTATTTATT GATGAAATCA 1020
 AAAGGCAATT AACTTATTAT GATGAGACAA TTGCGACGAA TAATGCAGAT CGATTATATT 1080
 TTCTAAAACA ACATACATTA GTGACTGAAC TTAAATATGA TGAAATTGAA AATGTTTATC 1140
 35 GTATAAAAGG ATTTAAAAAA TAATAAAAGG ACGAAATTCA AATGAAAGAT ATAAGTAAGA 1200
 TAGTAGCTGA CGTCGAATCA ACGTTAGCAC CATATTTTAA AGAAATTGAA GAAACAGCAT 1260
 ATATTAATCA AGAAAAAGTA TTAAATGCAT TTCATCATGT CAAAGCAACC GAAAGTGATC 1320
 40 TACAAGGATC AACAGGATAC GGGTATGATG ACTTTGGACG TGATCATTTA G 1371

(2) INFORMATION FOR SEQ ID NO: 272:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6035 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

55

	CGTAAAcCTA TGCgTTTTAA TATTCTGAAG TTACTTAGTt CATCCTCAGT TTCATCCATT	120
	TGTTTAATAT AAATAATACA TCCAGCTGCT ACTAAAAATG CTAATCCTAA AAATGATGTA	180
5	ACAAATATTA GAATACCGTt AGTAGCATcG ACCTCTTTTT TCATTTTCATC ATACGTGATG	240
	ACTTTGTCTC CAAACTGTTT TGCAATTGCT TGAGCTTTTT CCTTTTGTGA TGTTTGTtTA	300
	ATATCATATC CATAAAAAGT ATGAACGTTA TTTTGTGTtT TCAACTGCTG ATACTTTTCA	360
10	GGACTTACTT CGATGACAGG TGAGTTGAAG CTTAGATTtTA AAGGATAAAC CTTACCTTTG	420
	TCTTCTTGtG TAcACGGAAA GTTTCATTCT TAGTACCTtT TACTACTAAA TCTTTGTtTA	480
	AATGGATATT AATAATGTtTA GGCAGCGATT TTGTATTtGT AATGATGGCA TTGTTCGCTG	540
15	TtAACTTGtT ATTTGCACTT AAAATAGAAT TCGTGCGACC TGAATCACTA CCATTtTTCCA	600
	AAGTAATAAC CTGATCATTa ACATTATCTA CAGTAATAGT TTCGTAAGCA TTTTtTAGAA	660
20	ATGTAATtTG TTGTtGGCTT AGTTTtAGTTT CAAATTGTtT AGCATCTTGA GTAGCGACCA	720
	CGTTAAATtC ATTTGGTGCC ATAGATGTAA GGGTTTGATC TGTATTtTGAT TTAGATAATG	780
	CCGCAAAACA CAATACAGTt ACTGTAACTG CAGAAATAAT TGCAATGATA GTTAAAGACA	840
25	TGGCATTtTTT CTtCAttCTG TACATAATAG ACGATGTGAA TACAACATCG GTAAATAGATA	900
	CGCGTCCATT TTTTGATTtT TTCAATGTtT TAAAAATAAG TGACACGGAA CTtCTGAAGA	960
	ATAAATAGGC GCCTACAACC GTTAAAAATA AAATGATAAA CGGTGATGTC ATAGCCATAG	1020
30	TTAGTGCTtT GAACGTACCA AACATTtCTG TCGCCATATA ATAGCCTAGT GCAATCATAG	1080
	CAATACCTAA TACGCCTGAA ATAACCTCTG CAGTCGTTAC TTTAGCAGTt GTGGCATCAG	1140
	TTTTAATTGA ATCTtTCATC ATTGATAAGA TACTACGTCT TTTTAGAAAT AAAGCACTtT	1200
35	GAAATAAAAT CAGTACATAC GCAATAATTa GCATGAAAAT AGTTAAAACA AGGGCCATAG	1260
	GTTTCGAAATG TATCGATAAG TTAATCGATA ACGACATCAA TTTAGATACT ATGGAAAGCA	1320
40	ATAaTTGTGC ACCTGCAATG CCACATAATA CACCGACAAC ACCTGTGATT AAAAATACGA	1380
	TCATTtGtTC AAGTGCTAAC ATTTtncAAA TGTTTTGTcG TGTTAAACCA ATCaACTGAA	1440
	ATAGCGCAAA TTCACGTGTA CGGCGTTTTa CGTmTAAATG ATTGGCATAc ATTAAAAAGA	1500
45	TGACAATAAT GATAAATAAA AATATTGATC CGACTAAAGC ACCTTTCTTA ATGATGGCCA	1560
	TCGAGTCGTC ATTATTtTACA CCTTTAGTAA ACTGTAAAGT TGTAAACTG AAATATAAGA	1620
	CGATGCTAAA AAATAATGAA AATAAATACA TTGCATAATG TTTTAAGTTT TGTCGTAAGT	1680
50	TTTTGAAAAC GATATGGTTA AATGTCAttT GAGACACCAC CTAATACTGA TTGAAGATGT	1740
	ACAATGTCTT CATAAAAGGC CTGTTTAGAA CGTCCTTCCT GATAAAGTTG TGTATGAATT	1800

55

	ACCATGACAA TAGTTGTATC AAACGATTTA TTCATTTCTT CCAAACGTTG TAATAGGTCA	1920
	TTTGCACTTT TCGAGTCGAG TCGCCTGTT GGCTCATCTG CAAATATGAT TTGTGGTTTG	1980
5	TGAACAAATG CTCTCGCTGC TGCAGTTCTT TGTGTGTGAC CACCAGATAA TTCGCTAGGG	2040
	TATTTATTTT CTAGTCATA AATACCTAAT GCTGTCTGTA TCGCTTTATA ATTTTCTTCC	2100
	ATTGTTGCC TCGACATTTT TTGAACAGAT AAAGGTAACA TAATGTTTTT TTTAACGGTT	2160
10	AATGTCGGCA GAATACTGTA ATCTTGAAG ATGAAACCTA ATGATTCTTT GCGGAATTTG	2220
	GCAAGTGCTT TTTGATTAAG TTTATTAAGC TCTTGTCCGT TAGCAATCCg cTACCGCTAG	2280
	AAATTTGGTC AATTGAACTT AGTACATTTA ATAAGGTTGT CTTACCTGAT CCAGAAGGCC	2340
15	CCATAATCGC AACGAATTCG CCTTTTGTGA TGTCAAAGTT AATATCTTTA AGTGCTTGAA	2400
	ATGTGTGCTT TTTACCGTAT GTTTTGTGAA CATGTgCACT GATAATATCG TCATAGTCTC	2460
20	ACTCCTTtTG TATTTAATTT CATTTTAAAT AATGTTTGGG GTAGTAGCCT TTATCTAAAC	2520
	TTACAATTCA ATGAATGAAC CTTACAGAGT TGAAArCTAT CGCTACTTAG TAGATTTTTG	2580
	AGTGAGGATA CAGATTCATC GTACATATTA GACAAAAGCA ATGGTGCTTT CTAAGTGATG	2640
25	ATGTTTGTGT AAATTGAGAA AAGGGAATTT AATTATTGTA TAATAAATTT TTTGTAAAAA	2700
	TTAAAGAGG GTTTTATTTG AAAGGAATTG ATTGTTATGG AAAAAGGAAA TCAAGGTATT	2760
	AAATGGTCTA GTTTAATAAT GGGTGTATTA TTATTAATGT TGGCAGTCGT TATTTTTACA	2820
30	TTTCCAATG AAAATTTTTA TGCTATTACC TGGTTGATTG GACTGTTTGT ATTAATTAAC	2880
	GGTGTGATT C AAATCGTTTA CCGTAGAAAA GCAAAGCTT TAGTAGGTGG TAACCAAAT	2940
	TGGATTCTGT TTATGGGGAT TGTAATATT CTATTTGGTC TATTAGTTAT TTTAATGTT	3000
35	GGCGCAAGTT CAGCATCTT TATTTATATG TTTGCTTTTT GGTTTATTTT TAGTCTATC	3060
	TCTGGATTAT TTACGTTTTT GGGTAGTGGT AGCTTAAAC TAATTCAGT GATTTTTAAT	3120
	TTATTAGGTA TTGTTTTCGG TGTCATTTTA TTATTTAATC CATTAAATGG TATCGTCTTT	3180
40	ATTCGACGA TGATTGCTAT TGCATTTGTA TTCGTAGGTG TCATTTATGT TGTAAGTCA	3240
	CTTGCTTAAG TAAAATGAAG CGGTTCAAAA GAAGGGTGTG ACATGAAGTT TGTGTCATAT	3300
45	CCTTTTTGTT GTGTTTATGA AGCATAAAAA AGGGGCGCTA CCTACAATAA GTAAGATACG	3360
	CCCATATTTT TATATTTTAC TATTATTGTT TTTCAATACG ATTAATAGTT ACATTTAGTC	3420
	CAAAATATTT TTCTAAAAA TGTTTATAGT TATCTTTAGT GACATCAAAT TTTTCTGAGC	3480
50	TACCATTCCCT TGTAAAGTT AAATGATTTT CAGACATTGT AGCACGGCCA AATGATTGTG	3540
	GCATTGTAAT TAATAAATGC TGTACAAATA TTGAATCTGG ATGCGTTTGA TTATATTCGA	3600

55

	AATGATCATT TTCGAATTTT TGAACATAGA AAATATCCTT GTCTTCGTTG TTAAAAATAG	3720
	CGCGGAATGT ACCACTGATA TCAGTAATTG GTTGTGTATG CTCAGATGAA GTAATAGGAA	3780
5	TGGCATGTAG AGGTAAGTCT CCAAAGCCAA CATCAGTTAC ATAGAATACA TCATTTATAG	3840
	AAACAACAAG TGAAGCATGT GAACCGTTCA GACTACGACC GCCACCGGGw GTGTGAATAG	3900
	TAGCTGACAT TAATTCAGGA TTAAATCCTT TTTGTTGTAA ATAGGCTTTG AAAAATGTAT	3960
10	TTAATTCATA ACAAAAACCA CCACGTTTAT CATGAACAAT TTTATTAAAA AGTGCATCGA	4020
	TATTTATAGA TATCGGCTTA CTATTTTGAA CATCAATATT TTCAAAAGGT ACAGTTAACA	4080
	TAAAACGTGT TGCATAATAA TTTAATGCTT CAATACTCGG TCGATTATAA CGAGATGAAT	4140
15	CAATTTGTAA ATAATTCTCT AACTTCGCAA TATTCATAAG CATAGCGCCT CCTGTATTAA	4200
	AGATTATAAT TAAATTTTAA ACAGAAATAC TGAAATTTTA AATTCGAAAG CATTGAATTT	4260
20	TGGATAAATA CATTTTAAAT AGAAAAATAC GCTCTCAAAA TGAAGTCATC TCTAAAAGAA	4320
	ACGATTTAAA GATGACTACT GAGAGCGTAg CATAATGGAA GAAGTGTGCA GGGTGTCTAA	4380
	AAATGCAACA ATACAAAGGT AGTTGCAAGA CAAGTTGCCT TATCTAGACC ATTTGTGTTT	4440
25	TATGCGACCA AACTTCCAAA TTAAACTTGA AATAAGCCAA GTAATTAAAA ATAATGCAAC	4500
	TAAAATATAG CCTAAATAAT CAAATTCGAT CGAACCAATG AATGCCCAAA ACGCACCATG	4560
	TAAATCTAAC TTATCAGCAA GAATTTGTAG CAATTCAATC ATCCCAATCA CTAATGCTGC	4620
30	CATGACTGAT ATCGCAGTAA TCGTTATATT GTAATAGATT TTGCGAATAG GATTGAAGAA	4680
	TGCCCAATTA TAGGCATACT TCATTACAAC ACCATCTAAT GTATCCAATA AACTCATACC	4740
	TGATGCGAAT AAAATTGGTA AAGATAAGAT TCCGATAAAT GAAATGGCTT GTTGTGATGC	4800
35	GCCTGAAGAA AGAGCGAGTA ACGCAATTC ACTAGCTGTA TCAAAACCAA GTCCAAATAA	4860
	AAAGCCAAGT GGCAATACGT GCCAACTACG CGTGATTAAT TTGAAATAAG GTCCTACAAA	4920
40	TCGAGAAACC AATCCTCTAG ATTCAAGTAA TGCATCGACT TCAGCTTCTT CAATGTGTTC	4980
	ACGACGTAAT TTAGCGAACA AGTTAATTAA AGAGATTAAA ATAATTAGAT TCAACACACC	5040
	GATAAGCACT AAAAAGAAAC CTGAACTAG TGTACCAATC GTTCCACCAA TATCTTGAA	5100
45	ATGCGGTAAT TCATCTTTAG CCCATTTTAC AGATACCCCT AAAAAACAG CCATTAAAAA	5160
	TACGACAGAT GAATGTCCAA TTGAAAAATA GAAACCCACA CCAGATGGAT CTTTGCGTTG	5220
	CTGTAATAAT TTGCGaACCG TATTATCTAT TGCAGCAATG TGATCTGCAT CAAATGCATG	5280
50	ACGCAAACCT AATGTATATG CAAGAATCCC CaTACCAAAT AAGATATGAT GGTCTTTTCC	5340
	AGCAATCCAT AAAAACTAA ACCCAATAAC GTGTAACAAA ATGACAATAG CTATGTATGG	5400

55

ATATTTAATC ATACTGTATG TTCAATGGGC ACTCTAGTAA TAAGTGTTCA TATAACAAAA 5520
 ATGTTATGCC AAATTATTTG TTATATAAAA ATATACATGT AACCACAAAA GATTTTTTGC 5580
 5 GATATATATA ATTTGATAAA TTAACCAACA ACAATGTAAG ATGTCACTTT GCTTAACTTG 5640
 GCATCCTTTT TATGATTTTC AAATTCAAAA AATGAGCAA AATGAATCTC TTTACcAGTT 5700
 TTTAATATTT CaATACCATG CATGGAACCT AAGCACCCAT GTGTGATGCT GGAATGGATA 5760
 10 TTGAGACTAG CAACCTGATT GTAATGATTA GATAGTTCTT GAATTAATAT TTGAGGTCCG 5820
 TATATGTCAA AGCGGCCAGG GACAGACCAA ATAAATTCTG TTGTAACCAG TGAACGTAAT 5880
 AATTCAATAT CTAATGCTGC TGTAACAACT ATAAAATCTA TCATTTGTTG ACGTTTAGGC 5940
 15 GCATGATTGC ATGACACATC TCCTGTTAAC TTAAGGTA ATGATGACTG AACTTCCGTT 6000
 TTAAATGTA GTTGGTGCTG AAATAAGCT TGTC 6035

(2) INFORMATION FOR SEQ ID NO: 273:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1039 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

30 TTTTGAACAG CCATATTTAT TCACCCTCAA CATCATTATA ATGGTATTAG TCGCATTACC 60
 TTCACTTGTT TTAGCTATAT ATGATTATAT GAGTTTTAGA ATTTCTTCTG CTATTTTACA 120
 ATTTCTAGGG GCTATCTCTT GGTCTTTTTT ATCATTGATA TTATCGCTCA CACAATTTAC 180
 35 ACCTTTTACA TTAGCGTCAT TTATAACTTC AATTATTTTG TTCACAAGCA CAATTATCAC 240
 ATTAGCCATT GGTGGTAAGT CTGTTGAAAA GAATGATTCC CCTTAAATTC CAAATGAAAA 300
 AAAGGTTCTG AAGGCCGCTA TAAACACAG TTTTTCAGAA CCTCTATACT TCTATTCAAT 360
 40 GATATATGGT TTGCAATTTT CTACCTTTAA ATCCACAGCT TCTGCCCTTG AAACTTTGTT 420
 AAAATAAACC ATCAACAAC GAATGACAAC TTGATGTGCA ACAATGACAA TATCATCTTT 480
 45 TTGTGTATCT TCATTGACAA CATGATTCAT AAAATGTTCT ACGCGTTGAT ATACATCTTC 540
 ATAACTTTCT CCTTCAGGCG CTTTTGTGA AAAACTATGA CGAAAGTCTT TAAAGTTTGG 600
 ATCATTGAAA TATTTTTCAT ATTTCCGATT CGCACTGATT TCATCTTTAT ATTCACCCTC 660
 50 AAATACGCCA AGTGAACGTT CTCTTAATAG AGGGGTAGTC GTTGATGCAA TGTCAATG 720
 AAAAATATGT TCAAACGTTT GCTGTGTTCT TAATAAGTCT GAAACATATA CATGTTTAAT 780

CGGCACATCT AATTGTCCAC AAAAATAAGA TCGAAAATGT TTATTATCAT AATTCGATTT 900
 TGATTCGCCA TGTCTAACTA AATAAATCGT CATAATATTA CTCCTTACCT TATGTATTTT 960
 5 ATATCTACCA TAACACTTTG ACTACTAATT CGATATCAAT CTTAATATTC TATTCTAAAA 1020
 AAAGAATTAA TTCATATnT 1039

(2) INFORMATION FOR SEQ ID NO: 274:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1496 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 274:

GAGAGAATnT GCAATTAGTT ATTCAATTAG TTGATTTAAG ACATGATCCA ACACAAGATG 60
 20 ATATCTTAAT GTACAATTAT TTGAAACATT TTGATATTCC TACTTTAGTT ATATGCACTA 120
 ArGAaGACAA AATTCCaAAA GGTAAGGTYC AAAAGCATAT TAAAAATATT AAGACACAAT 180
 25 TAGATATGGA CCCAGACGAT ACAATTGTAA GTTATTCATC AATTCAAAAT AATamaCAAC 240
 AACAAATATG GAATTTAATT GAACCGTATA TTTCATAGTT TTTGTACGTC AAAACTTATA 300
 CAAAAATTTT AAAAATAATG TAAGCAGGAA ACTTTTAATT AGTACACAAT TGATAACATT 360
 30 TTTCAACGTT CATCATTTTG TCAAAAATC AAAAGTAAAT TAGAAAGATT ATAATTTATT 420
 TAAGCATCGT ACTTAATTGG ATTTTAAATT ATGTTATAAT ATTTGTATTG TTAGTATATA 480
 TGGGGGCTTT TCAAATGCAT TTTATTGCAA TTAGTATAAA TCATCGCACA GCTGATGTgC 540
 35 ACTAAGAGAG CAAGTTACTT TTAGAGATGA TGCCTTACGA ATTGCCCATG AAGATTTATA 600
 TGAAACTAAA TCTAyTTTAG AAAATGgtCA TATTaTCAAC ATGTAATCGA ACTGAAGTAT 660
 ATGCTGTTGT TGATCAAATT CACACAGGTC GTTACTATAT TCAACGATTT CTAGCTCGTG 720
 40 CATTTGGATT TGAAGTAGAT GATATTAAAG CAATGTCAGA AGTAAAAGTG GGGGACGAAG 780
 CaGTAGAACA TTTATTGCGT GTCACCTCTG GTTTAGATTC AATCGTACTT GGAGAAACTC 840
 45 AAATTTTAGG TCAAATAAGA GATGCATTTT TCTTAGCGCA AAGCACAGGT ACGACAGGrA 900
 CAATTTTTAA TCATCTATTT AAACAGGCAA TTACTTTTGC AAAAAGAGCA CATAATGAAA 960
 CAGATATAGC TGATAATGCT GTAAGTGTGT CTTATGCTGC GGTCGAGTTG GCGAAAAAAG 1020
 50 TATTTGGCAA ATTGAAAAGT AAGCAAGCTA TCATTATTGG TGCAGGGGAA ATGAGTGAAT 1080
 TATCACTATT AAATCTTCTT GGTCTGGAA TTACTGaTAT TACAGTAGTA AATAGAACAA 1140

TACCAAATTT ACTTGAAAGT GCAGATATTG TGATTAGTTC AACGAGTGCA CAATCTTATA 1260
 TCATTACAAA TGAAATGATA GAAAGAATTG CAGAAAATAG AAAGCAAGAT TCACTAGTAT 1320
 5 TGATTGATAT TGCAGTTCCT CGAGATATTG AACCTGGTAT TAGTGCCATC ACAAACATCT 1380
 TTAATTATGA TGTGTATGAC TTAAAAGGTT TAGTTGATGC AAACCTACGT GAGCGACAAT 1440
 TAGCGGCTGC AACAAATTCG GAACAAATTC CTACAGAAAT ACATGCACAC AATGAG 1496

10 (2) INFORMATION FOR SEQ ID NO: 275:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4826 base pairs

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:

CTTGATTTTT TCCCTTAGT ATTTTCCaTt TGanTGTCGC AGCTTCTAAA TCCTGCTTTG 60
 GTTCTCTAGT GAACTTCATA ATTAAAGCAG CTACAACGAA TGATACAAGT GCAGCAAGGA 120
 25 AGACACCGAG TAACATGTGC AAGAATTCAC CTCTAGGTGC ATTTAAACAG TAAACTATAA 180
 ATGAACCTGG TGACGCGGGA CTTTTAAATC CAAATCCTGT TGCTTGATAA GTTGCAACAC 240
 CAGTCATTCC ACCTAAAATA ACAGCGATAA ATAATAAAGG ACGCATTAAT ACATATGGGA 300
 30 AATAAATTTT ATGAATACCA CCTAAGAAGT GGATAATTCC AGCACCATAT GACGTTGCTT 360
 TTGCAGTGCC TTTTCCAAAA ATCATATAAG CAAGTAAGAT ACCTAAACCT GGTCAGGGT 420
 TAGATTCAAT TGTGTATAAA ATTGATTGAC CAGCTTTTGC AGCTTGATCT GCACCAAGCG 480
 35 GTGTGAATAC ACCATGGTTA ATCGCATTGT TAAAAATAC AATTTTGTGA GGCTCTACTA 540
 AAATACTTAC AAGTGGAAGT AGGTGTGCAT GACTAATGC TTCAACTGCC ACTGATAAAA 600
 TATGCATAAT AAATTTTATA AGTGGTGCTA AAATTTTAAA TCCTGCAATC GTCATGATAA 660
 40 ATCCTAAAAT ACCAGCAGAA AAGTTATTAA ATAACATTTC AAAACCTTGC GGC GTTCTAG 720
 GTTGAATCAA TTGGTCGGTC TTCTTCATTA ACCAACCAAC AAGTGGACCC ATAATCATTG 780
 45 CACCAAGTAA CATTGGTGTA TCAGGTAATG CAACGATGAC CCCCATAGTT GCTGTTGCTG 840
 CGATGATACC ACCACGTAAA TCATAAATTA AACGACCACC ACTAAATGCG ATCAATAATG 900
 GGATTAAATA AGTAATCATT GGTCTGTGTA AAGTAGCTAA ATCTTTGTGA GGTAACCATC 960
 50 CATTATCTAT AAAAATGGCC GCGATAAAAC CCCAAGCGAT GAAAGCGCCA ATGTTTGGCA 1020
 TGATCATACT ACTTAAGAAT GATCCAAATG CTTGAACACG ACGACCAATT CCTTTTTTCT 1080

55

	GAGAGGTTAC TTGTTACTCA ATATAAACAA AAATCAACTT TGTCAAAATA AATGTGACAA	1200
	AATTAAATAA AGTGTGATCA ATGTGACAGT ATAGATATTT TGAAAAAGTA AAACAAAAAA	1260
5	ATTGTTTTAG GATTTTTTAAA ATTTTATTGT GAAAAATTTT GCAAAACAAA ACAACACCGT	1320
	GTACAATAAT GATTAATGGA AAGGGGGAAA GTTCGGCAGT ACAGTTAAAG CGCCTGTGCA	1380
	AATAAATATT TGTATTGAA GATTAAAGGT TAATATATGA GTGGCCTTTA TAGAGTGCAA	1440
10	TATATGTATT TGTAGACGAG GAGGATAGTG ATCGAATAGA TCGGCGGATG CTATCCCGGA	1500
	TGTGGCTCAT TCGTTAGCTT ATTAAGTAAA ACATTAGGGT GACTTAATGG ACAAAGTTAA	1560
	TAAGATCGCC AGAAATTGAA TATAAAAAAT ATTAATATGG AAAGTACAGT GTGAGCAATT	1620
15	TGTATAGTTG TAAAAATAAC TATGCTTAAT TTGTTATGGA TGAATGCGAT GATAGCATGT	1680
	TCCTATTTAT ATTATGAAAG CAGATTGTCA ATCTAAATTA TCGGCAATAA ATCATAATTT	1740
20	ACGCGTACTA TTCCAATATG GAGGAAAATG TCGTTATGTG TGGAATTGTT GGTTATATTG	1800
	GCTATGaTAA TGCCAAAGAA TTATTATTAA AAGGTTTAGA AAAATTAGAA TACAGAGGTT	1860
	ATGACTCTGC AGGTATCGCA GTAGTAAATG ATGATAATAC AACTGTATTT AAAGAAAAAG	1920
25	GTCGTATTGC AGAATTACGT AAAGTTGCTG ATAGTAGCGA TTTTGATGGA CCTGTTGGAA	1980
	TCGGTCACAC ACGTTGGGCA ACACACGGTG TACCGAATCA TGAAAACCTCT CATCCACATC	2040
	AATCATCAAA TGGCCGTTTT ACTCTAGTTC ATAACGGTGT TATTGAAAAC TATGAAGAGT	2100
30	TAAAAGGTGA AACTTTACAA GGTGTATCAT TCATTTGAGA AACAGATACA GAAGTTATCG	2160
	TTCAATTAGT TGAATACTTT TCAAATCAAG GACTTTCAAC TGAAGAAGCA TTTACAAAAG	2220
	TTGTGTCAAT ATTACATGGT TCATATGCAT TAGGTTTATT AGATGCTGAA GACAAAGACA	2280
35	CAATCTATGT TGCTAAAAAT AAATCACCAT TATTATTAGG TGTTGGTGAA GGTTTCAATG	2340
	TTATCGCATC AGACGCACTT GCAATGTTAC AAGTGACAAG CGAATATAAA GAAATCCATG	2400
	ACCATGAAAT CGTTATTGTT AAAAAAGATG AAGTTATTAT TAAAGATGCA GATGGAAACG	2460
40	TTGTAGAACG TGATTCATAT ATTGCTGAAA TTGATGCATC AGATGCTGAA AAAGGTGTTT	2520
	ATGCACACTA CATGTTAAAA GAAATTCATG AACAACCAGC AGTAATGCCG CGTATTATTC	2580
45	AAGAATATCA AGATGCAGAA GGTAACCTGA AAATTGATCA AGACATCATC AATGATGTTA	2640
	AAGAAGCAGA CCGCATTTAC GTTATTGCAG CAGGTACAAG CTACCATGCA GGTTTAGTAG	2700
	GTAAAGAATT TTTAGAAAAA TGGGCTGGCG TACCAACTGA AGTACACGTT GCATCAGAGT	2760
50	TTGTCTACAA CATGCCATTA TTATCTGAAA AACCATTGTT CGTTTATATT TCTCAATCAG	2820
	GTGAAACTGC AGATAGCCGC GCCGTATTAG TTGAAACTAA TAAATTAGGT CATAAATCAT	2880

55

	TACACGCGGG TCCTGAAATC GCAGTTGCAT CTACAAAAGC ATATACTGCA CAAATTGCAG	3000
	TATTATCAAT CTTGTCTCAA ATCGTTGCAA AAGAGCaTGG TCGTGAAGCA GATATTGATT	3060
5	TATTGAGAGA ATTAGCAAAA GTAACAACAG CAATAGAAGC AATTGTTGAC GATGCACCAA	3120
	TTATGGAACA AATTGCTACA GATTTCCTTAG AAACAACACG CAATGCATTG TTTATCGGAC	3180
	GTACTATTGA CTATAACGTA AGTTTAGAAG GTGCGTTAA ACTTAAAGAA ATTTCTTACA	3240
10	TTCaAGCAGA AGGTTTTGCT GGTGGAGAAC TTAAACATGG TACAATTGCC TTAATCGAAG	3300
	AAGGTACACC AGTTGTAGGT TTAGCAACAC AAGAGAAAGT TAATTTATCA ATTCTGGTA	3360
	ACGTTAAAGA GGTAGTAGCA CGTGGTGCAC ATCCATGTAT TATTTCTATG GAGGGTCTTG	3420
15	AAAAAGAAGG CGACACTTAT GTCATTCCTC ATGTACATGA ATTGTTAACG CCATTAGTAT	3480
	CAGTGGTTGC ATTACAATTA ATTCATACT ATGCAGCATT ACACAGAGAT TTAGATGTTG	3540
20	ATAAACCACG TAACCTTGCT AAATCAGTTA CTGTGGAATA ATTCACITTT TTAGAATCAA	3600
	TCATGTATTA AAATTAAAGT ATATGGCACC CTTTTAGATT AATCGACTAG AAGGGTGCTT	3660
	TTTTAGGTG ACTTaGCTTT TACTTCATCT TAATTTGGCA GAAATGCGTa AAAATGAAGT	3720
25	GTTTTATTTA TTAAATAGT CTGACAATTA AGGGTGTTAT GTTAATATGA TTTTATGAGA	3780
	AGTATGGAGT AGCAATAAAG GGGTGACCTC GCATGTTAAT TCAATTAGAT CAAATTGGGC	3840
	GAATGAAGCA AGGAAAAACA ATTTTAAAAA AGATTTCTTG GCAAATTGCT AAAGGTGATA	3900
30	AATGGATATT ATATGGGTTG AATGGTGCTG GCAAGACAAC ACTTCTAAAT ATTTTAAATG	3960
	CGTATGAGCC TGCAACATCT GGAAGTGTTA ACCTTTTCGG TAAATGCCA GGCAAGGTAG	4020
	GGTATTCTGC AGAGACTGTA CGACAACATA TAGGTTTTGT ATCTCATAGT TTAGTGAAA	4080
35	AGTTTCAAGA GGGTGAAAGA GTAATCGATG TGGTGATAAG CGGTGCCTTT AAATCAATTG	4140
	GTGTTTATCA AGATATTGAT GATGAGATAC GTAATGAAGC ACATCAATTA CTTAAATTAG	4200
	TTGGAATGTC TGCTAAAGCG CAACAATATA TTGGTTATTT ATCTACCGGT GAAAAACAAC	4260
40	GAGTGATGAT TGCACGAGCT TTAATGGGGC AACCCAGGT TTTAATTTTA GATGAGCCAG	4320
	CAGCTGGTTT AGACTTTATT GCACGAGAAT CGTTGTTAAG TATACTTGAC TCATTGTCAG	4380
45	ATTCATATCC AACGCTTGCG ATGATTTATG TGACGCACTT TATTGAAGAA ATAAGTCTA	4440
	ACTTTTCCAA AATTTTACTG CTAAAAGATG GCCAAAGTAT TCAACAAGGC GCTGTAGAAG	4500
	ACATATTAAC TTCTGAAAC ATGTCACGAT TTTTCCAGAA AAATGTAGCA GTTCAAAGAT	4560
50	GGAATAATCG ATTTTCTATG GCAATGTTAG AGTAAATATT TTGCAAATAA TAAGTAATAA	4620
	TGACAAAATT TAATTAAGAT AAAATGGACA GTGGAGGGCA ATATGGATAA CGTAAAAGCA	4680

55

AAAGATGTCA TTAATCAATT GAGAGAGAAA GGATATAAAG TATTTTGGC AACAGGACGT 4800

TCGCATTCTG AAAATACATC AACTTG 4826

5

(2) INFORMATION FOR SEQ ID NO: 276:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4846 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

15

GAATAAAAAG TAAAATTACT CGCCTTTGTT ACCTTTACT TTATCAATaA AATCAGTTGC 60

TTTTTCTTTT GCATTTTCAA CGAATTCTTT CGCTTTACCA GAAGCTTTAT CTTCTTTACC 120

20

TTCGTTTTCT AAATTTTAT TATCAGTAAC ATTACCTACT GTTTCTTTAA CATTACCTTT 180

TGCTTGTTCA AATTACTTT CGTCTGCCAT AATAAATGCC TCCTCGGAAT AATTAAATGT 240

TATATATAAT ACTTACCCAC TGAAAAATTA TCTAAACATT TTAATTAAAT AATTTTTGAT 300

25

ATTGATTGA CGTCATTTTA TAACTAGCGA AATAGATTCA TCATTAACCT GAGGGAGTGG 360

GACTGAAATA ATAAAGAATC ACTAATGATT TATGATGTAT TAGTCACTAG CCATGTGAAA 420

TTAAAAATAA GAATAAATGA GTAGCAGCGA TGCATATAGG ATTTTACTTT ATCCGTAATA 480

30

GCATCTCATT CCTAAATATC ATATAAATAC CTGTTTAAAT TAAAAAGCCC AGCAACATCA 540

CGTTACTGAG CCATTAATAT GATTATTTTA GCAGGAATAA CTGCACCATT GTATTTTTCA 600

TTAATGAAGT CTTGAATATC TTTAGATTGT AATACTTCAA TTAATGCTTT GATTTTCTTA 660

35

TCATCTTGAT GTCCTTCTTT AACAGCAATT AAGTTTGCAT AAGGATTATC TTTCGCACTT 720

TCTACAGCAA TAGAATCTTT TTTAGGATTT AGTTTTTGTG CGATTGCAAA GTTCGAATTA 780

40

ATGATAACAG CATCAGCGTC TTCATTTTGA TAAATTTTAG GTAAGAATTC TGCTGATTGT 840

TTATTATTAA ACTTAATATC TTTTTTATTC TCTGTAATAT CACTAAACTT AGCATCTTCA 900

ATTTTACGC CTTTTTTGAT TTTAATTAAA CCTGCATCAA CGAAGAATTT TAAGAAACGT 960

45

CCTTGTTTCAAG CTGGATTATT AGACACATAG ACTGTTGCAC CTTTTGGTAA TTCTTTTAAA 1020

CTTTTATACT TTTTAGAGTA TACAGCCATA GTTTCTAAGT GAACATCACC GGCACCTACG 1080

ATTTTGTAAC CTTTATCCTT TTTCTCTGTG TTTAAATATG GTGTATGTTG GAAATAGTTT 1140

50

GCGTCAATTT CACCTTTGTC TAGTAATTTA TTAGGTGTAG TGTAATCGTT AATTGTTTTA 1200

ATATCTAGTT CATAACCTTT TTTCTCTAAT AATGGTTTTG CTTTTTCTAA AATTTCAGCA 1260

55

	TTACCGCAAG CTGCTAATAC AACTGCAAAT GTTAATACTA AAATAAGACC AAATAATTTT	1380
	TTCATAAAAT GAAACCCCCA ATTTATCGTT TATCAAGTTT ATTTGTAAGC CAATCCCCAA	1440
5	TGAATTGGAT TATAAATACA ATAATTAAAA TAAAACTGT TGATACTAAA ATGACATCAT	1500
	TTTGATTTCG AGTGAAACCT GTTAAGTATG CTAAATTTCC TAAACCACCG GCACCAATTA	1560
	CACCTGCAAC TGCTGTTGAA CCAACTAAAG CGATTGCTGT AACTGTAATG CCAGACACTA	1620
10	GCGCTGGCAT AGCTTCAGGT AAAAGGACTT TACGAATTAC TGTCCAAGTA TTAGCGCCCA	1680
	TTGACCAAGC CGCTTCGATG ACACCTTTAT CAATTTCTTT AAAAGCAATT TCTACGAGCC	1740
	TTGCATAAAA CGGTGctGCG CCAATGATCA AGGCTGGTAA CGCACCTGTC GGACCACTTA	1800
15	TCGTTCCAAG TATCAAACCT GTAAATGGAA TTAATAATAA AATTAAAATA ATAAATGGTA	1860
	TCGCTCTAAA TAAGTTAACA ATGAAAGAAA CGATAGAATA AAATAACCTT GCACCGATAG	1920
20	ACTTACCTTT AGCAGACAAG AATAATAACA CACCTAAAAT AAGACCAAGT ATAAATGCAA	1980
	ATATAGTTGA GACGACTGTC ATGTATAGTG TTTGACTAT TGCAGTCCAA ACTTCTGGCC	2040
	ACTGAATATT AGGCATTGTA ATCATTTCAT TTATAATTTC ACTAAATGAT TTACCCATGT	2100
25	CTTAACACCT CCATTTTAAC TTGTCGCTCA ATTAACCTT TTTGGAATTT TCCGAAATCT	2160
	ACACTTGAAA TATATGGAAT ATGCAGAACT AAAAAGCCGA CTGTTCCATT TTTGTATTT	2220
	TTAATATTTG CTTCTAAAAT ATTAATTTTA ATATCATAGG CAGTTGATAG ACTCGATACA	2280
30	ATAGGCTCGG TTGTTGTTGA ACCAGCGAAA ACTAATCTAA CGATATATGC ATCTTTTCT	2340
	AATGGCTCTA ATTCTGTTAA AGATGTTTCG AAATCATCAT TTAAATCGTC TTTCACAAAT	2400
	CGTTTTGTCA CAGTGTGTTG CGGATTTTCA AAAACCTGTG TCACCGGTCC TTGTTCTATC	2460
35	ACTTTACCAC TTTCCATAAC TGCAACTTCA TCACAAATAC GACGAATGAC ATGCATTTCA	2520
	TGCGTAATTA GTACAATTGT TAAATTTTGT TGTTCTCTAA TTTTAGTAG TAGATCTAAA	2580
40	ATTTATCTG TTGTTTGGG ATCAAGTGCA CTTGTTGCCT CATCACAAAG CAAGACCGTT	2640
	GGATCATTaG TAACGCTCGT GCAATCCCAA CACGTTGCTT TTGTCCACCT GATAACTCTG	2700
	ATGGATAAGC CTTTTCTCTA CCTTTTAAAC CGACGAGTTC GACAAGTTCT AATGCTTTTT	2760
45	GCTTAGCTCT CTTCTAGGG ACACCTGCAA TTTCAAGCGG AACATAATA TTTTTTAAAC	2820
	CAGTCCTTGA CCATAACAAA TTAAAATGTT GGAAGATCAT ACTTACTTTT TGTCTTTTGT	2880
	CTCTTAATCC ATTTTTGGAC AATTGACCTA TATGGTCTCC ATCTATAATA ACTTCACCTG	2940
50	ATGTAGGCGC TTCTAAATGA TTAAACATTC GAATCAAAGT ACTTTTTCCT GCTCCAGAAA	3000
	AACCAATGAC GCCATAAATC GATCCTGCTC GAATCGATAA ATTAACGTGA TCTACAGCAA	3060

55

	TTCTCCCTG TGTGCTTAA TAAATAAAA AATGCTTTCT CAATATCGAT AGAAAAATTG	3180
	AGAAAGCAAT AGTAGTATTG TTTCTCTCAT CTTCAAAAGT TAAAACTTTA TGTGAATTGG	3240
5	CACCATTCT ATATAAGACG GTTGCCGGGC TTCGTAGGGC ACATCCCTCC ACCACTCTCG	3300
	ATAAGAGTTT ACGCATCATT TAATTTGTAT TAATCCTAAC ACCTTAGTAA AATTTCTGCA	3360
10	ATAACTATTT TAAATTTTCT AACAAATCAG TCACCGATT AAATGCATAA ATTCGTTTTA	3420
	CTTCTTTATC TTTATTCATC AACAATAAAA TCGGCGTAGA CATGATTGTC ATATCTTTAC	3480
	AAAACGAGG ATAAAAGTTT AAATCTATTT TCAATAATGG TAACTGCAAT ATTTCAATTAG	3540
15	CAATGTCTAA CATCTTTCT GAAACCTTAC AAGTACCACA CGTTGGTGTA TAACCAAAGA	3600
	TTAAATGTTT GTCTTCTCA TAAATGTAG TTACATCTTT GATGTCTAAT GAATTATTCA	3660
	TTTACTAAAA CTAACCTTTC ATTATTTATA TTCGGTAAAA GAGGTGTTTC TTTCTTACAA	3720
20	GTAAAGCCAT GTTTTGAAAG TACATGCGCC AAATATTGTT TGGGGCAATT CGCAACTGA	3780
	CAGTAAGTTT TATCAATAAA TATATGTTCA CTTTCACTCA AATAACGTTT AAACCAATTT	3840
	CTAATTTCGAT CTCCTTCGTC ATCAGAATCG GCTAATACAA AAACCTGTTT ATCATAAGT	3900
25	GATTCTATCA TATCATCAAG CTTATCTATA CTCATTGTTT CATGAGTACA AATAATATTG	3960
	ACTGGTCTG CAATAACCTG TTGCACCCTT TTTTATCAG ATTTTCCTTC AACAATTATC	4020
	ACTTTATTTA CAATAGCCAT CATCATCACC CTTTAAAATC AATAAACATC TGTCACGTGA	4080
30	TCATTTTACA AAATTGGTAT GAATAAAACA TAAATCACAA AAAATTTAAA CTAGCTTAAT	4140
	ATAATAATTA CAACTCAAT GTTTGACTAG CTGGAACATT TAACATAAGC AGACAAAGGC	4200
	TAAGTCAAAA ATCAACATCC TAAATCTAC AATGTTATAT TAACAATAGT TAACCAAAAG	4260
35	AAAATACACC TATAACAAAC TTTTCAATTA TAGCGGGGCC CCAACACAGA AGCTGATGGT	4320
	AAGTĒAGCTT ACAATAATGT GCAAGTTGGC GGGGCCCAA CATAAAGAAA TACTTTTTCT	4380
40	TTAGAAATTA GTATTTCTTA TGCATGAGTT TTACTCATGT ATTCCTATTT TTAAATACAC	4440
	ATTAGCTGTG GCTTATGAAA ACAGGCTGGG ACATAATCA ATGTTCTATG CTCTACGAAG	4500
	TTATATTGGC AGTAGTTGAC TGAACGAAAA TGCGCTTGTA ACAAGCTTTT TTCAATTCTA	4560
45	GTCAGGGGCC CCAACACAGA GAATTTGAA AAGAAATCT ACAGGCAATG CAAGTTGGGG	4620
	ATGGGCCCCA ACAAGAGAA ATTGGATTCC CAATTTCTAC AGACAATGCA AGTTGGGGTG	4680
	GGACGACGAA ATAAATTTTG CGAAAATATC ATTTATGTCC CACTCCCTAG ATTGATCTAT	4740
50	AGATACTACA CTTATTAAAG TAATATAITT TTATGATTCT CTTAGCTGCA ATCCCATGAA	4800
	TACATGTAAT CATCAAACCT CATAGCCTCA AGGTCAGTAG ATTTCA	4846

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1843 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 277:

10	AACAAAGACA CAATCGAACA TGAACCATCA GTAAAGCTG AAGATATATC AAAAAAGGAG	60
	GATACACCAA AAGAAGTAGC TGATGTTGCT GAAGTTCAGC CGAAATCGTC AGTCACTCAT	120
15	AACGCAGAGA CACCTAAGGT TAGAAAAGCT CGTTCGTGTG ATGAAGGCTC TTTTGATATT	180
	ACAAGAGATT CTAAAAATGT AGTTGAATCT ACCCCAATTA CAATTCAAGG TAAAGAACAT	240
	TTTGAAGGTT ACGGAAGTGT TGATATACAA AAAAAACCAA CAGATTTAGG GGTATCAGAG	300
20	GTAACCAGGT TTAATGTTGG TAATGAAAGT AATGGTTTGA TAGGAGCTTT ACAATTAAAA	360
	AATAAAATAG ATTTTGTAGTAA GGATTTCAAT TTTAAAGTTA GAGTGGCAAA TAACCATCAA	420
	TCAAATACCA CAGGTGCTGA TGGTTGGGGG TTCTTATTTA GTAAAGGAAA TGCAGAAGAA	480
25	TATTTAACTA ATGGTGAAT CCTTGGGGAT AAAGGTCTGG TAAATTCAGG CGGATTTAAA	540
	ATTGATACTG GATACATTTA TACAAGTTCC ATGGACAAAA CTGAAAAGCA AGCTGGACAA	600
	GGTTATAGAG GATACGGAGC TTTTGTGAAA AATGACAGTT CTGGTAATTC ACAAATGGTT	660
30	GGAGAAAATA TTGATAATC AAAAATAAT TTTTAACT ATGCGGACAA TTCAACTAAT	720
	ACATCAGATG GAAAGTTTCA TGGGCAACGT TTAAATGATG TCATCTTAAC TTATGTTGCT	780
35	TCAACTGGTA AAATGAGAGC AGAATATGCT GGTAAACTT GGGAGACTTC AATAACAGAT	840
	TTAGGTTTAT CTAAAAATCA GGCATATAAT TTCTTAATTA CATCTAGTCA AAGATGGGGC	900
	CTTAATCAAG GGATAAATGC AAATGGCTGG ATGAGAACTG ACTTGAAAGG TTCAGAGTTT	960
40	ACTTTTACAC CAGAAGCGCC AAAACAATA ACAGAATTAG AAAAAAAGT TGAAGAGATT	1020
	CCATTCAAGA AAGAACGTAA ATTTAATCCG GATTTAGCAC CAGGGACAGA AAAAGTAACA	1080
45	AGAGAAGGAC AAAAAGGTGA GAAGACAATA ACGACrCCAA CACTAAAAAA TCCATTAACT	1140
	GGAGWAATTA TTAGTAAAGG TGAAYCgAAA GAAGAAATCA CAAAAGATCC GATTAATGAA	1200
	TTAACAGAAT ACGGACCAGA AACGATAACA CCAGGTCATC GAGACGAATT TGATCCGAAG	1260
50	TTACCAACAG GAGAGAAAGA GGAAGTTCCA GGTAAACCAG GAATTAAGAA TCCAGAAACA	1320
	GGAGAYGTAG TTAGACCACC GGTTCGATAGC GTAACAAAAT ATGGACCTGT AAAAGGAGAC	1380
55	TCGATTGTAG AAAAAGAAGA rATTCCATTC rAGAAAGAAC GTAAATTTAA TCCTGATTTA	1440

CCAACACTAA AAAATCCATT AACTGGAGAA ATTATTAGTA AAGGTGAATC GAAAGAAGAA 1560
 ATCACAAAAG ATCCGATTAA TGAATTAACA GAATACGGAC CAGAAACGAT AACACCAGGT 1620
 5 CATCGAGACG AATTTGATCC GAAGTTACCA ACAGGAGAGA AAGAGGAAGT TCCAGGTAAA 1680
 CCAGGAATTa AGAATCCAGA AACAGGAGAT GTAGTTAGAC CACCGGTCGA TaGCGTAACA 1740
 AAATATGGgA CCTGTaAAAAG GAGACTCgAT TgTaGGAAAA AGaGAATTc CaTTCaGGAA 1800
 10 AGAcGTaATT TaTCCTGTTT AGCACCCGGG GCAGAAAAGT TAC 1843

(2) INFORMATION FOR SEQ ID NO: 278:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8536 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:

TAAACAGCGC GTGTACTTGT GATTCCCCCT TCTTCTATTT TACCCACCCG GGAAATAATA 60
 25 CTTTTCGCGA TTCCTTACTT GaACAAGCAA TATTTTATCa GCTGTTTCTT CAACTAAACA 120
 GACACATTTA ATCATCTTTG ACACCCCAAC TTTGTGAAAT CAATTTTTCa AATTATACTG 180
 TACAATTATG TTATCATATA TGAGTAGTTA TAGCGCAAAA CGTTAGCAAT TCAGCGCACC 240
 30 CAACTTTTCA TATAAACAGA AGATACTAGG GGAATTATT ATTATGGCAA AACgTTCCAA 300
 ATCACAACGT TTATCAAGTT TACTAAATGT CGCAGGTTTC ATAGTCGACG GCTACAATGG 360
 CTATAAATAT CATGCTAAAA ATAAAAAATT AGTATATCTT TCATTAGGTT TAAGCACTGT 420
 35 AGGAACCGTG TTAGACTTTT ACATTTCAAT TAAGTCACCA CGTAAGTTCA AAAAAGCAGT 480
 GGCAGTTGTT ACTTTAATAA CAAACGGTGC TAGATTATTT ACAAGCATTc GCAAAGTAAA 540
 40 ACATGAATAC TAATTCAGAA AAGGATTGGT CGAACATAGA ACATGAAGTT CATTcGACCA 600
 ATCCTTTTTTc ATATACAAAA ATTCCTATTA CAACAATCAC GACTTTGATA GCCCCGCCAT 660
 AAAATTTAGG ATTCAATCCA ACTTTTCAGC TTGTGaAATG TAATAGGaAT TCATTATATA 720
 45 TTTATATACG TAAGACTTTA GTGAATATAT CTATAATTAT TTAATTGGTA AGCTGGTACC 780
 GTTCTGTAAG TTAATAATACC TGGTGCAGCT GAATAGTTCA TTTCTGAAAC TAAAATACTA 840
 CCATCGTTAT TTACACGTTc TACAAACATA ACATGACCAT AGTAACCTAC ATCTGTTTGA 900
 50 GCGATAGAAC CTACAGTAGG TCTATTGTCG ATAGTGTAAC CATCTGCTGC CGCTGCGTTA 960
 TCCCAGTTAT TAGCATTCCA CCAATAAGTA CTAATACCTT TACCAATTTc AGCACGACGA 1020

55

EP 0 786 519 A2

	GTATTGTAAC CTCTATTTGT CGTTGTTGCA GATCCTGAGT TCGTAGATGC aTTACCAGTT	1140
	ACTTTCAATT TTTGACCCGG ATATATAAAG AAATTATTTA AGCCATTTAA GCTCATAATT	1200
5	TTTTGATAAG TTGTACCATA TTTTGATGCA ATTAATGACA ATGAGTCACC TGCTTGTA	1260
	GTATAGTATG ATCCGCCACC TGAGTTCGTT GATGGACGGC TACTATTGCT CGCAGCGTTA	1320
	CTTGAGCTAG CAGTACCTGA TACTTTTAAT TTTTGACCTG GATAAATAAA GAAATTATTT	1380
10	AAACCATTAA GTCGCATAAT GTTTTGGTAA GTTGATACCAT ATTTTGATGC GATTAATGAT	1440
	AATGAGTCGC CTGCTGTAC TGTGTAGTAT GATCCGCCAC CTGAGTTCGT TGATGGACGG	1500
	CTACTATTAC TCGTAGAATT ACTTGAGCCA GATACTTTTA GTACTTGGTT TGGGAAAATT	1560
15	AGATTAGATG TTAAATTGTT TAATGACTTT AATTTAGCAA TCGAAATCCC ATACTTATTT	1620
	GAAATGCCC ACACTGATTC ACCCGGTTTT ACTGTGTGAG TTGTAGCCGC ATTTGCTTGA	1680
20	GTTGCCGCAA CAGCGCTAAT CGCGCTTGTC CCAATAATAG CTGCAATTAC TTTTTTTTGC	1740
	ACTTTAAAT CCTCCTCTTG CTTAACCTTC CTAACATTCT TTTATCCGAA TTTATGAATA	1800
	CTACATCATT ATACGATTTT ATTATGTATA ATAGGTTGAT GTTTGATGAC ATTATGaTTA	1860
25	CAAAAAATC ATATACTGTA TCATCAAATT TTATAATTAT CCCTTAAAT TATTACAACT	1920
	TATTAGATTT TACAATATCT AAATTATTAC AATTTCATAA TATTTCACTA TAAAATGATT	1980
	ACAATCCCTT TCTCTATTGG AAATAATTTT ATTCTCCAAC AATAACGCCC TACAAACATA	2040
30	AGCATGAACT TTGCTGTAG GACGCAATAT AATTTATTTT GCTAATGACG TTTCTATTGC	2100
	CTTAATCTCA TCTTTAGATA AATTAACAGG TTTCTCTCCA TCTTTGACAT CTTCTGCAA	2160
35	CGCTTTTGA GCTTCTTTTG AATGATACAA TTCTACGATT TTAGCATATG TTTTGTATC	2220
	CAAGTCTTTG TCATTAACTG CAACAATATT AATATATGGC TTTACTGCAT CTGAATTGA	2280
	TTTTTCTAAA AATATCGGAT CATTTTATAGG ATCTTTACCC GCTTTAGTTG CTACACCGTT	2340
40	ATTAATAACT GCAATATCGA CATCAGATAA AGCACGTGCA GTTTGTTGTG CATCTACTGC	2400
	AGTAATTTTT AAATGTTTTG GATTGACGT TATATCTTC ACCGTGCCTG CTAATCCGAA	2460
	ATCTTTTTTC AGTTTTATTA AACCAGCTGC TTCTAATAGT TTAAGTGCAC GTGCTTGGTT	2520
45	TGACACATCA TTTGGAATGA CAACTTTAGC ACCATCTTTA ACCTTTTGA CATCTTTAAT	2580
	TTTATCTGAG TAAATGCCCA ACGGTGCTAA AACTGTTGTA CTTAATGCTG AAATCTTTGT	2640
	TCCTTTATGC GCCTTTTAT ATTGATCTAA AAATGCAAAA TGTGGAATG CATTCAATC	2700
50	AATATCACCA TCATTTAATG CTTTATTCGG TAAATTGTAA TCTGAGAAGT GCTTAATCTC	2760
	CACATCAATA TCATCTTTTT TAGCTAATTC TTTAACCTTC TCCCAAGCCT TAGTGTCATT	2820

55

EP 0 786 519 A2

	AAGTGCTACG	ATAACTAACC	CAATCAATCT	TTTCATTCTA	TCAATTCCTT	TCAAAATCTT	2940
	CACTATATAT	CATTAATGTC	TACGTATGAA	TCTAGCTAGA	ACATTCCCTA	GCGTTTGAAT	3000
5	CAC TTGGACA	ATAATGACTA	ATACAATAAC	GGTAATAATA	ATGACCGTCG	TATCAAATCT	3060
	TTGATAACCA	TACACTAAAG	CTAAGTCTCC	TATACCACCA	CCGCCAACAG	CTCCTGCCAT	3120
	CGCCGTACTT	CCAATAAGTC	CAATAATCGC	AGTGGAATT	GCTAATACTA	ACGAACCTAA	3180
10	AGCTTCAGGA	ATTAAAAAAT	ATCTAATGAT	TTGTAGTGGT	GAAGCGCCCA	TCGnTTTCGC	3240
	CGCTTCAATA	ATCCCCCTCGT	CTACTTCCAA	TAATGAGTTT	TCAACAAGTC	TTGCAATGTA	3300
	AGGTGCCACA	TATACTGTTA	AAGGCACGAT	GGCAGCAGTC	GTACCAATTG	AAGTACCTAC	3360
15	TACTAATTTT	GTGAATGGCA	CAATCGCAAT	TAACAAAATA	ATAAATGGTA	GTGACCTTAA	3420
	AATATTGATT	AAAGGATTTA	AACTTGATG	TATCACTATA	TTGGGCCATA	TGCCTTGTTT	3480
20	TCGAGTAATT	ACCAATAAGA	CACCTAATGG	AATACCAATC	ACTGCTCCTA	AAAATAAAGC	3540
	AATAGATACC	ATATATAGCG	TTTCGTACAA	TGCTTGTAAT	AACTGTGCAC	TGTCTAAATC	3600
	AGAACCAAAC	ATATGTTAAT	GcACCTCCTC	AAATTGAATA	TTTTTCTCTT	TGAAATATTG	3660
25	ATTTATTGCC	GTGTCTTCAA	ATTGTTGATC	CATATTAAAT	CGAAGCCACA	TATAACATAC	3720
	GGTGTTACCT	TGTATTTCTG	ACATAGATGA	AAATAAAATT	TTAACCTCTC	TGCCACAAAT	3780
	TTGAATCAAG	TCATTTATAA	TCGGTTGTGT	CACCTGAGTT	TCCTCGACGA	AGATTTTATA	3840
30	ATCTTTAAAA	TCGCCAACTT	GTTTCGTCATT	CAATCGACGA	ATCAATGATG	TACTTGCTC	3900
	AGTCTGTATA	ACTGTAGACA	CAAAATTTTG	AGCAATCGTC	GTTTTAGGAT	GACTAAACAC	3960
	CTCTTTAAca	GTTCTGTTT	CAACCACTTT	CCCCTTTTCC	ATTACAGCAA	CACGATTACA	4020
35	AATGTCTTTA	ATAACGCGCA	TTTCATGTGT	AATCATCATA	ATTGTAATGC	CAAAGGTTTG	4080
	ATTGACATTC	TTTAATAACG	TCAATATCGA	AGCAGTCGTT	GCTGGATCCA	ATGCGCTTGT	4140
40	TGCTTCATCG	CATAGGAGTA	TTTTCGGATT	AGTAACAAGC	GCTCTTGCAA	TAGCCACCCT	4200
	TTGCTTCTGC	CCACCAGATA	ATTCATCAGG	AAATTGGTCT	TTTTTATCAC	TCAATCCTAC	4260
	AAATTCAAGC	ATTTCCGTTA	CTCGTTGCTT	AATTTCTGTT	TTGCTTTTCT	TACTTAAAAT	4320
45	GAGTGGCATT	GCTACATTTT	TAAATACGGT	AGCTGAATTT	AATAAATTGA	AATGCTGAAA	4380
	TATCATACCG	ATATCTTTCT	TAATATCCCT	CATCATTTTA	TCGCTATAAT	TCGTAATATC	4440
	ATGTCCATCT	ACAATCACTT	GTCCATTGCA	GGCAGCTTCA	AGATGATTCA	CGAGTCTTAC	4500
50	CAACGTACTT	TTTCTGTCAC	CACTATATCC	AATCACACCA	AAAATATCAT	TGCGATTGAC	4560
	CGTAAATGAT	ACGTCCTTCA	AAGCATCTAT	TTTTTGCTTC	TTTTTATTAA	AGGTCTTACT	4620

55

	TAAGAGCATT ATATGTAAAA TTGCATATAT CGTCAATACA ATTTGCCGAA TTTTCTAAAA	4740
	AATTAAAAAA TAAGTAATTC ATGTGACAAT GACGAATTGT GAGACTACTA TGACATTTAT	4800
5	CAAAATTAAAT CCATAAAAAT GTCCACCAAT CCTCCACAAC GCAATTACTA AATATTAACA	4860
	TGCGACAAAA AAGCACTAGC ATATTCAAGA ACAACAAACG TTGAACTCAA AATATATGCC	4920
	AGTGCTGCTA TTATTTATAA AGTATCTAGT GCTTGTTTTA AATCATCGAC TAAATCTTCA	4980
10	GTATCTTCAA TACCTACAGA AATTCCTTACA AGTCCGCTCTG TAATACCTTC TTTAGCTCGA	5040
	ATATCTGCTG GAATGGATGC ATGTGTCATC AATGCAGGTA CTGAAATTAA ACTTTCCACT	5100
	GCACCTAAAC TTTCAGCTAA TGTGTAATAC GATGTTGCTT TAATCAATTG TTTGGCACTT	5160
15	TCTGTATTTT TCACTTCAAA TGCAATCACA CCTGTATGGC CATCCGCTTG AGCCATATGG	5220
	ACATCATGAT TTAAATGACT TTCAATACTT GGATGGAACA CTTGTTGCAC AGCTGGATGT	5280
20	GCTTGTAACA TTTTAATAAT TTCAATAACG CTGCGATTAA TTTGTTCCAT ACGTAAACCT	5340
	AATGTTTTAA TACCCCTCAC AAGTAAATAG CTATCTTGAG GTCCTAAAAT GCCACCTGTT	5400
	GAATTTGAAA TAAATGCTAA ACGTTCTGCA AGCTTGTCAT CCGATGTTGC AACTAAACCA	5460
25	GCAACGACAT CACTATGTCC ACCTAAATAT TTCGTTGCAG AATGTAAGAC AATATCGATA	5520
	CCTAAATCTA ATGGATTCTG ATAATAAGGT GTCATAAATG TGTATCAAC AACTGAAATC	5580
	AAACCGTGTT CTTTCGCAAT TTCAGCAGAC TTTTAAATGT CAGTAACACG TAATAATGGA	5640
30	TTAGAAGGTG TTTCAATAAA CAACATCTTT GTTGTGGGC GTATCGCTTG TACAATTGAA	5700
	TCTGTATGCG TTGTATCTAC AAAATCCACT TCAATGCCAA ATCGTGTAAT TACTTTTGTC	5760
	AATGCGCGAT AAGTACCGCC GTATACATCT GAATTTAAAA TAATATGATC TCCTTTGTCC	5820
35	AACAGCATAA CAACTGCACT GATTGCTGCA ACACCTGAAC TAAATGCAAA GCCATGTTTG	5880
	CCATTTTCTA ATGTCGCAAT AACGCTTTCT ACAGAACTTC TTGTTGGATT CGCAGTACGA	5940
40	GAATATTCAT ATCCTTGACG TAAATCACCA ATATCATCTT GTAAATATGT ACTTGTTTGA	6000
	TAAATTGGTG TTGTAACGGC ACCTGTATAA TCGTCTGTTG TGTGCCCACC ATGAATTAAT	6060
	TTAGTTTTCT TGTTCATTAT TATTCTCCTC ATAATTAAAT ATTTGCTTAG ACATATATCG	6120
45	ATCACTACCA TCTGGAAATA CGACAACAAT CGTACCTTCA GATAATTGCG CTTTTAAATT	6180
	CAATGCACCT TGTAATGCTG CACCTGAGA ACTGCCTACT AACAAACCTT CATTTATAGC	6240
	CAAACCTTTG ACATTTTCGAA AGGCATCTTG ATCTTTAATC GTAAATATCC CATCTACAAG	6300
50	ACGTCTCTCT AAAAATATCG GCCATTTCTC AGAACCGATA CCTTCAGTGT CATGTGCATG	6360
	AGCTGGCCCT CCATTTAACA CGGACCCTTC TGGCTCAACG GCATAACATT GCACGTGATG	6420
55		

ATAATCAATT TGCTGTAATG CTGAAGTCAA TTCGGGTCCC AATGTATGAA AATATGTATC 6540
 CGGATTATGT TCGGATTCAA ATTGATTCAT ATAAACGGCA CCATATTTTT CAGCATAGGA 6600
 5 ACGTGCAGCT AATTGTGCCC CATGCATACC TTCAGACTGA CTCGTCCTTG AAACCTCTGC 6660
 ACCAAGCGCT ATCATAATAT TAATCTTTTC TTCTGAAAAA CCATACGGCG CAAAGATCTT 6720
 ACATTTCAAA TGATGTCTAT TCGCTGCAAT AGCTAACCCCT ATGCCTGTAT TACCAGCAGT 6780
 10 CGCTTCAACA ATAGTTTGAC CTGCACGCAC ACGCCCTTCT TGAATTGCCT TCTCTACTAA 6840
 ATATTTCCCG AGTCTGTCTT TAACACTGCC TCCAGGATTC CATGTGTTCAA GCTTGGCATA 6900
 AATTTTAACT TTATCATCAC TATAATGTTT TAACAGTACT AATGGTGTAT TGCCaATTAA 6960
 15 ATCATAAGTA ATCATAGATG CACCCTCATC TGACATGCCG ATCAAATGAA TGAAACCTTT 7020
 CTTTCATGTCT CAATTTTAAT TCTTACTTTT CAGATAAGAA TTATAAACGA CATTTTGTTA 7080
 20 TTTTGCAATT ATCTAAGTTT CGATTAATTC AGAACAGTA CTAAATTTTC AATTCCAAAC 7140
 AAAAAACAC CTGAGCAACA CAAATACTTG TGTGTCAGAT GCTTCTATAT ATTAACTAAA 7200
 TAATTGCACG ATAAAGACTA AAATAATAAC GACAGGCATC GCATACTTAA TTAAGTAATA 7260
 25 CCAACCACTG AATAATCTAA ATCGATCTTT ACCAAAATAT TGTGTGAATA ATTTTTTATC 7320
 TAATAATTGT CCTACGACAA GCGTAGTACC TAATGCGCCT AATGGCATCA ATACATTGCA 7380
 AACGATGAAA TCCATATTAT CAAAATCGT TCCCGCACC G AATCTTACAT CTTTTAAGAT 7440
 30 ACCAAAAGAT AAGGTTGCTG GAATACTAAT GATAAATACT AAAATACTAC CGATCACTGC 7500
 GACTTTTTTA CGTTTTGTAT TGTCACTTCT CGTGAAGTTA GAAACATTTA ATTCTAATAA 7560
 AGAAATAGAT GACGTTAAAG CCGCAAATAA GAACAGCACT AAGAATCCCA AATAGAATAA 7620
 35 TGTGCCTAGA TGCATTTGAC TAAAGACCAT TGGCAGTACT TTAAATAATA ATCCAGGCCC 7680
 TTCTGTGGT TCATAGCCAA AACTATGTAA AGCCGGAAAT ATAGCTAGAC CTGCCAATAC 7740
 40 AGATACAAAG ATATTCATAA CAACGATAGA AATAGCTGAT GACTTAATCG TCATGTCTTT 7800
 AGAGGCATAA CTCGCATAAG TAATCATACC TGTAGTTCCT AATGATAACG TAAAGAATGA 7860
 TTGACCTAGC GCAAACAAGA TGCCATCAGC AGTAATCTCT GATACTCTTG GTTGTAATAA 7920
 45 AAATTTCAAC CCTTCTAAGA CGCCATCTAA TGTAAAGAC TTAATCACAA TGACGATTAA 7980
 AAAGACAAAC AGCAATGGCA TCATAACTTT CGATGCCTTT TCTAATCCTT TTTCAACACC 8040
 TAACATGACA ATAATCATCG TAGCGAATAT GAATATACCT TGCCCTAGAA CGGTTAACCA 8100
 50 AGGATTTGAT ATTACCGCTT CAAAATTCAT TTCTTGAGA TGATTGATGC GTTGAAATAT 8160
 AACTAATTGC CATAATACTT GTCCGATGTA AATGACAATC CAACCACCGA TAACACTATA 8220

55

TTTACCAGTT AATTTACTAT ATATTTGTGT TGTATATGTC CGTCCCATT TCCCAACAGT 8340
 GAATTCCATA ATGAGTAATG GCAACCCAAC AAAAATGGTG AATATTAAGA ACATAGCTAG 8400
 5 AAAGGCACCG CCGCCATAAA TCCCTGCCAT ATATGGGAAT TTCCACATGG CACCAAGACC 8460
 GATTGCAGAA CCGCACTAG CTAAAATAAA TCCAGTTGAT GACTTCCATT GTGATTGTTG 8520
 TCTTTTCATC ATTCAC 8536
 10

(2) INFORMATION FOR SEQ ID NO: 279:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4328 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 279:

GCTTTGGCCA TTTTATGTGG CGATTGAGAC AATCtGTkGT TGTCTTATTT GATGTTGTAT 60
 TTCAACTGGT AATTCTAGTT GCGATTGAAA TAATGGCAAC TTTTCCCAAT CATTAAACAA 120
 25 TAATTCAATA CCGCTATGT CTAATACTTT AGCACGTGCA TCATCAACAA GACGGCGTTC 180
 CAATTGATTT GCTTCTTCTT TAATACCTGG TGACGTACTT TCTAATATCA AATTAGATAT 240
 AGGGATGTGA CCATTAATTG CATAATATAA TGCAACACGC CCACCCATTG AATATCCAAA 300
 30 CAATGTTATT GATTTATCTT TATATTTATC TAAAATTTCGG TCTAACAACG TCGTAATATA 360
 ATCAAAATTC CACGTTTCAT CCATTGAAGA CTGATCTTCG CCATGGCCTG GTAAGTCTAT 420
 35 AGTGATGACA TGATAGTTAT CAGTAAATTT TTCGATGTGA TTATGATAAG TACGGCTGTC 480
 GCTAAGAAAT CCATGCAGAA ATACTAAAAC TTGATTGGTC TCAACGTTTG CTTCATAAAA 540
 TTTATAATGT GTCATGAATC ATTTCACTCA ATTTCTGGTA TAAAATTTGA TGCTGTTTAA 600
 40 AGTTATCTTC GCGATTCGTT ATCAATTCAT AAATCGTCGA AGTTTCAGAT AACAAATGTGG 660
 CATTTTTTAA TTCTGAAACA CTGTTAAAAC GTTTAAAATC GAATTGATAT AACTTAGCTG 720
 TATACTCGAA ATCCAATCCC GTCGGTGTGC CAAACAACCG TTCAAAATAG TCAGTTGCAC 780
 45 TTTCTTTTTG TGGTAAATAT GAAAAAATAC CGCCACCATC GTTGTTCAAT AATACAATAT 840
 TCATCTGAAT ATTATTTAAT TTTGACATTA ATAGTCCATT CATATCATGA TAAAATGATA 900
 AATCACCTAT CAATAATGTT ATTCGTTTAT GCACAGCCAT ACCCAGTGCA GTTGAAACGA 960
 50 TACCATCAAT ACCATTCGCA CCACGATTCG CATAGACATC TATATTTTTA TTCAATAACA 1020
 AGTTATCTAC ATCTCTGATA GGCATACTAT TACTAATAAA TAATGCATCT TTTTCAGATG 1080
 55

	TAATTTCTTT	ACGCCCTTTT	TTCTCTAAGC	ATTGCCATTT	TTCTAACCAA	CTTACGCGAT	1200
	TAACTGTCGT	GTCTTCCATT	AATGACCTAA	AGAAATCATT	CGCAGAAATC	TCATATGAAA	1260
5	TATCTGGCGC	TATCGGAAAG	ACATCAATCT	TATCATTGTT	TTGCACTAAT	ATTTGAAATG	1320
	CATCAGTTTT	CTTTAACCAT	TGATTTAACT	TTTTAGAAAT	CACTGGTTTC	CCAACACGAA	1380
10	TTACGAAATC	CACATTTAAG	TCTAAGCCGC	TTCTAAACAG	CAAATCATAT	GTACAGATAA	1440
	CATTCGGATG	ATCAAATTTT	CTTAAATGAC	TTAAAGGATC	AGCTAAAATA	GGCAAATCAT	1500
	ATATCGTTGA	ATACGTTAGT	ATTTGaTCAA	CTTCTTGGTG	CTGCATATCC	CCTACAATAA	1560
15	TTAAACCTTT	TTTCTTATTT	AAAATGTGTC	TTAATGCCGA	TGCATCTATA	CTTTTTTGAT	1620
	AGTGGGTAA	AATCTTCATC	TCAGAAGTTA	ACAATTCTGT	TGCATTCAAA	TCAGGTGTTA	1680
	ACGGATCTCT	AAATGGCAAG	TTAAATGAA	TTGGCCCTTT	ATGTGGTCCA	TATAAATATT	1740
20	GACTAGCAAT	TTGCATTTGA	TAGTAAATTG	CATCAATGGT	CTCTTTACTA	TCATCCGCAA	1800
	TAGGCATATC	GAAGTCATAA	CTTACATAAT	TATTAAACAT	ATTTACTTGA	TTAATCGCTT	1860
	GTGGTGCGCC	TACACTTCTT	AATTCATGCG	GACGGTCACT	TGTTAAAACG	ATTAAAGGAA	1920
25	TTCTACTAAT	TTGGCTTTCA	GCAATTGCAG	GCGTATAATT	CGCTGCTGCT	GTACCTGACG	1980
	TACATAATAT	AGCGACAGGT	CTTTCCTGCT	CTTTAATTAA	CCCAACTGCA	AAAAACGCTG	2040
30	CACCTCGCTC	ATCGGGGTGT	ATCCATGTTT	TAATATTTGG	ATGTGCTTCA	AATGCAAGTG	2100
	CAAGTGCGCT	TGAGCGTGAT	CCCGGACTGA	TAACTACTTC	CCTTACGCCG	TACGCATATA	2160
	ACTCAGATGC	AAATGTAAAA	ACTTGCTTCG	TTAAAGCTGC	TTTATGATTT	CCCATTCTAA	2220
35	TCGACTCCTA	ATGCATTCAT	CATAGGTGTG	AACTTAAGGT	TCGTTTCTGC	CAATTCACTA	2280
	TCTGGATCAG	AATCTTTAAC	AATGCCACAC	CCAGCAAATA	AAGTTGCTTG	TGCTTTCTTA	2340
	ATAAGCATCG	AACGAATTGC	AACAATAAAT	TCACAATCAT	CGTATATATC	TATATAGCCA	2400
40	ACCGGTGCAC	CATATAATCC	TCGCGTACCA	AATTCTTTCT	GCTCAATAAA	ATCCATTGCA	2460
	AATTCTTTTG	GATAGCCACC	TAAAGCAGGT	GTTGGATGTA	AATTATCAAT	TAAACTAATA	2520
	TACGAATCAT	CCTTCAGTGG	CGCCTTTATT	TCAGTGTACA	AGTGATATAA	ATGATCATTT	2580
45	TTTAGAATTT	TAGGCGTCTT	ATCATAATGT	AATTCAGTGA	TATAAGGTTT	AATATCATGT	2640
	AAAATACTGT	CAACAACAAA	TCGATGTTTC	ATTAAGTTTT	TATTATCTTT	TAAAAATGCT	2700
50	TCAACATTTT	TTGTATCTTC	GTCCTCATCT	TGTGAACGTT	TAATTGTACC	TGCTACAGCT	2760
	TTAGTCGATA	GTATTTTATT	ATTGACCTTT	ATTAATTGTT	CAGGTGTTTG	TGAAAAGAAT	2820
	ATAGAATCTT	GTGATTCTAA	CAAGAATATA	TAACTGTTTT	TTTCTTTAGA	ATATGCTTGC	2880
55							

5 ACAAATTTTT CTTCATTATT AATAGATTCT ATAGCTTCTA CTACAAGTTG ACGCCAGTCA 3000
 TCTTTATAAA TATCTTCATT TCTAGTAATT TCCCAATTT GCTCGTCCAC ATCTATGTCC 3060
 10 GATATATTGT TGAACAAATC CATTAAATCG TTCAATGCCT CAACAGTAAA ACTTTCCCTT 3120
 TTAACGTGAT AAGTTAAAAA TGTCCCATTA TTATCAGTTG AAATTAAAAAC TTCAGGTAAT 3180
 ACAAATGAT TTAGTCCAAA CTCTCGCCAT TCATCATCTG ATTTATGACT TGAAAATTGG 3240
 15 AACCTCCAA CAACTCGAAG ATGATGTTTC TCAGATTGCG GATGTATAAA TGTGATGTTA 3300
 TGTTTTAATT TTTCCAGTC TTTAAAAATA GATTGTTTAT TTTTAGAATT ATTTTGAAT 3360
 AATTGAATTG CTTTGTAGCC AAAATATGAC GTTCGATTAT CATTCAAACG CATATAAAG 3420
 CGATCTCCTG CCTCATTGTC AGTGAGATGA AATAATGTGC TCGGTCTAG TGAATGTGAT 3480
 AATTTCACTT CAACTGAAAC CCATTCTTTT GAGCTGCCAT ATATCTCTTT GACAATATCG 3540
 20 TCCTCTAATA CGCCCGTAGC CATCCATTTC ACTTCTTTCT TCGTCTTTTT TCACTCATT 3600
 TTATATTGTA TCATTTTGG ATAATTGTGT TACAAGAATT GCTTAACTT ATCTTGCAAT 3660
 TTTTCACGTC AATTGACCTT TATGCTACTT TCTATTAAAA TATCTTTGTT ATAAAAATA 3720
 25 TGATTAAAG AGGTTTGTG TTCAATGAGT AATCAATATC AGCAATATTC TACAGTTAAG 3780
 AAATATTGGC ATTTAATGCG TCCTCATACA TTAAGTCTT CCGTAGTACC CGTTTGTAGT 3840
 GGTACAGCAG CATCTAAAT ATATTTTCTT GGTAGCGAAG ATCATATTAA AATCAGCCTA 3900
 30 TTCATTGCCA TGTTACTAGC ATGCTTACTT ATTCAAGCAG CAACTAATAT GTTTAATGAA 3960
 TACTATGATT ATAAAAAAGG CCTCGATGAT CATGAATCTG TAGGCATTGG TGGTGCCATT 4020
 GTTCGCAACG GTATGAGCCC AGAGCTTGTG CTACGATTAG CCATTGCATT TTACATCTTA 4080
 35 GCAGCAATAT TAGGTTTGT TTAGCTGCT AACTCTTCAT TTTGGTTATT ACCAGTTGGA 4140
 TTAGTATGTA TGGCTGTTGG TTACCTATAT ACAGGTGGCC CTTTCCCTAT TTCATGGACG 4200
 40 CCTTCGGTG AATTATTCTC AGGCGTATTT ATGGGTATGT TTATTATCGT TATTGCATTC 4260
 TTTATTCAA CTGGCAATAT TCAAAGTTAT GTAATTTGGT TAAGTGTACC TATAGTAATC 4320
 ACTATCGG 4328

45 (2) INFORMATION FOR SEQ ID NO: 280:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1450 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

EP 0 786 519 A2

	GTTCAATACA GAAAAAATAA ATTTAGATGT TGAAGCATCC TACAATTAAT ACAGATCCAT	60
	TTCAATATAT TTAAACTAAA ATCTCGGGAT TTCTAAATTT TGAAATTTTCG AGGTTTTnAT	120
5	ATTTTTATTT AAAATAGCAC ATTTATACTT TATAATAGTA AAGATGAACA TATAAGGAGG	180
	CCAAATCATG GCAAAACATC CATTGCAACA ATTTAATCTA GAATCTAGTT TAATTGACGC	240
10	TGTGAAAGAC CTTAATTTTG AAAAACCAAC TGAAATTCAG AATCGAATTA TTCCAAGAAT	300
	ACTAAAGAGA ACAAATTTAA TTGGTCAATC TCAAACGGGT ACAGGGAAAT CTCATGCATT	360
	TTTATTACCA TTAATGCACT TAATTGATAG TGAAATAAAA GAACCACAAG CAATCGTAGT	420
15	TGCACCAACA AGAGAACTTG CACAACAACT ATACGATGCA GCGAACCATT TAAGCCAATT	480
	TAAAGCTGGT GTTTCAGTTA AAGTTTTTAT TGGTGGTACA GATATAGAGA AAGATAGACA	540
	ACGTTGTAAT GCACAACCAC AATTGATTAT AGGCACCCCT ACTAGAATTA ATGACTTAGC	600
20	TAAAACGGGA CATTTACATG TGCACCTAGC ATCATATTTA GTTATTGATG AAGCGGATCT	660
	TATGATTGAC TTAGGATTAA TTGAAGATGT AGATTACATT GCTGCAAGAT TGGAAGATAA	720
	TGCAAATATT GCGGTGTTTA GTGCTACAAT CCCACAACAG TTACAACCAT TTTTAAATAA	780
25	ATATTTAAGT CATCCAGAAT ATGTAGCTGT CGACAGTAAA AAACAAAATA AAAAGAACAT	840
	CGAATTCTAT TTAATACCTA CTAAAGGTGC AGCTAAAGTT GAAAAGACTT TAAATTTAAT	900
30	TGATATACTA AATCCATACT TATGTATTAT TTTCTGTAAT AGTAGAGATA ATGCAAATGA	960
	TTTAGCACGT TCACTAAATG AAGCTGGTAT TAAAGTTGGT ATGATTCATG GTGGCTTAAC	1020
	GCCaCGTgAA CGTAAACAAC AAATGaRACG TATACGTAAT TtaGAATTCC aATACGTTAT	1080
35	TGCCaGCGAT TTAGCATCTC GTGGTATTGA TATTGAAGGT GTTAGTCrTG TCATCaATTT	1140
	TGATGTGCCA AATGATATTG ACTTCTTTAC GCATAGAGTC GGACGAACTG GTCGTGGGAA	1200
	TTATTrAGGT GTAGCAATTA CGCTTTATAG TCCTGATGAA GAACACAATA TTTCATTAAAT	1260
40	AGAAGATCGC GGTTTTGTAT TCAATACTGT TGATATTAAA GATGGTGAGT TAAAAGAAGT	1320
	TAAAGCGCAC AATCAGCGTC AAGCAAGAAT GCGCAAAGAT GACCATTTAA CTAATCAAGT	1380
	GAAGAACAAA GTTCGAAGTA AAATTAAAAA CAAAGTTAAA CCAGGTTATA AGAAGAAATT	1440
45	TAAACAAGAA	1450

(2) INFORMATION FOR SEQ ID NO: 281:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1139 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

5 AGTCAGGTAT ATCATGCCaT yCTGAATTGG TCGATATTAA TATCAGTGGT GTTAAAGAAC 60
 GAATTGTATA CCAATAGACG CTTTATATTG TAAAATAGTA TTAAATGCaG AATAGAGAGG 120
 AGATTTAATG CGATATGACA AATTATAAAG TTGTCGTTTT AGATATGGAT GACACATTGC 180
 10 TAAATTCAGA TAATGTGATA TCAGAAGAAA CTGCAAATTA TTAAACAGCA ATTCAAGATG 240
 AAGGTTATTA TGTGTGTCTA GCATCTGGTA GACCTACTGA AGGTATGATT CCAACTGCTA 300
 GAGATTTAAA ATTACCTGAA CATCATAGCT ATATTATTAG TTATAACGGT AGTAAAACGA 360
 15 TTAACATGAC TAATGAAGAA GTAGAAGTAA GTAAATCGAT TGGTAAGCAA GATTTCGATG 420
 AAATTGTAGA TTATTGTGCA GATAGAGGCT TTTTCGTTCT TACATATCAT GATGGTCAAA 480
 TTATTTaCGA CAGCGAACAT GAGTATATGA ATATTGAAGC AGAATTAACA GGTTTACCGA 540
 20 TGAAACGTGT TGATGATATC AAAGCGTATA TTCAAGGCGA TGTACCCAAG GTCATGGGTG 600
 TAGATTATGT AGCGAATATT ACAGAAGCTA GAATTGATTT GAATGGTGTG TTCAATGATA 660
 ATGTAGATGC TACGACAAGT AAGCCATTCT TCTTAGAATT TATGGCCAAA GACGTTTCAA 720
 25 AAGGTAATGC AATTAAAGCG TTATGTCACA AATTGGGATA TTCGGTGGAT CAAGTCATTG 780
 CTTTTGGTGA TAGTATGAAT GATAAATCAA TGTTTGAAGT CGCAGGTCTA GCTATTGCTA 840
 TGGGGaATGC ATCAGATGAA CTTAAGCAAT ATGCAAATGA AGTTACGTTG GATCATAATG 900
 30 AAAATGGTAT TCCACATGCG CTCAAAAAAT TGTTATAAAT TTTAAATAA GCCTTAACAC 960
 ATGATATTTG AATAAGATAT CTTGTGGTTA AGGCTTTTTA TTTTGTGAA AATGACTTCA 1020
 35 GTTATACTAT GGAGGATTTG AAATACATAT TTTAGATTAG TAATGATATC AAACGAATAG 1080
 AGTAAATGTA TATTTttTGA ATAAATCAAG TATTAAGTAG TCACGGAAGG nAGATAAAT 1139

(2) INFORMATION FOR SEQ ID NO: 282:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2931 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

50 TCTAAAAATG CTGTGAAATT CTTTATAAAA TATCTAAAAG GAATTAATGT TGATAACATT 60
 GCTGTGATAG GAAGTAAGAC AGCGCAATAT TGTGAATCAC TTGGCATTCTG AGTTGATTTT 120
 AtGCCAAACG ACTTTTCTCA AGAAGGATTT TTAATCAAT TTAATCAAAC TAACCAAAAA 180

	AATGAAGTTG TTAAATAGA TTTATATACT TCACTGCCTA ACAAACAAAA TATACAAGAT	300
	GTAAAGAAA TGATAGAACA TCAACAAATC GATGCATTAA CATTTTCAAG TTCGTCGGCA	360
5	GTACGTTATT ATTTAATGA AGGATTTGTA CCAAAATTCA AGTCGTATTT TGCTATTGGA	420
	GAACAAACAG CACGGACCAT TAAATCATAT CAACAACCAG TAACAATTGC AGAAATTCAA	480
	ACACTCGAAT CACTAATTGA AAAGATTTTA GAAAGTAGGG GCTAAAAATG AAATTTGATA	540
10	GACATAGAAG ATTGAGATCA TCAGCGACAA TGAGAGATAT GGTAGAGAG AATCATGTAA	600
	GAAAAGAAGA TTTAATATAT CCAATTTTTG TAGTTGAAAA AGACGATGTG AAAAAAGAAA	660
	TTAAGTCATT GCCAGGTGTA TACCAAATCA GTTTGAATTT ACTTGAAAGT GAATTAAAAG	720
15	AAGCTTATGA CTTAGGCATA CGTGCCATTA TGTTTTTTCGG TGTTCCAAAC TCAAAAGATG	780
	ATATAGGTAC TGGTGCATAC ATTACGATG GTGTTATTCA ACAGGCAACA CGTATTGCTA	840
20	AAAAATGTA TGATGACTTA TTAATTGTTG CAGACACTTG TTTATGTGAA TATACTGATC	900
	ATGGTCATTG TGGCGTGATT GATGACCATA CACATGACGT TGACAATGAT AAATCATTGC	960
	CACTACTTGT TAAAACAGCA ATTTCTCAAG TGGAAGCTGG TGCTGATATT ATTGCGCCAA	1020
25	GTAATATGAT GGATGGTTTT GTTGCTGAAA TTCGTCGTGG ATTAGATGAA GCCGGCTATT	1080
	ACAATATTCC TATAATGAGT TATGGTGTCA AGTATGCATC AAGTTTCTTT GGACCTTTTA	1140
	GAGATGCAGC AGATTTCAGCG CCATCATTTG GGGATAGAAA AACGTATCAG ATGGACCCTG	1200
30	CTAACCGTTT GGAAGCACTT CGTGAATTAG AAAGTGATCT TAAAGAAGGG TCGACATGA	1260
	TGATTGTTAA ACCTGCTCTA AGTTATTTAG ATATAGTTCT AGATGTTAAA AATCATACGA	1320
	ATGTTCCAGT TGTTCATAT AATGTGAGTG GAGAATATAG TATGACTAAA GCAGCGGCAC	1380
35	AAAATGGTTG GATAGATGAA GAACGTGTCG TTATGGAACA AATGGTTTCA ATGAAACGTG	1440
	CAGGTGCTGA TATGATTATT ACGTATTTTG CAAAGGACAT TTGTCGCTAT TTAGATAAAT	1500
40	AAGGTTTTAT ATTTATGATT TTCCATAAAC TGTAGGAGGA ATTTACTTTA TGAGATATAC	1560
	GAAATCAGAA GAAGCAATGa AGGTTGCTGA AACTTTAATG CCTGGTGGTG TAAATAGTCC	1620
	AGTACGCGCA TTTAAATCAG TAGATACACC AGCAATTTTT ATGGATCACG GTAAAGGTTT	1680
45	AAAAATTTAT GATATCGATG GTAACGAGTA TATCGACTAT GTACTAAGTT GGGGACCACT	1740
	TATTTTAGGA CATAGAGACC CTCAAGTTAT TAGTCATTIA CATGAAGCAA TTGATAAAGG	1800
	TACAAGTTTT GGTGCATCAA CATTACTTGA AAATAAATG GCGCAGcTCG TTATTGACCG	1860
50	AGTACCTTCA ATAGAAAAAG TGCGTATGGT GTCATCTGGT ACAGAAGCTA CATTGGATAC	1920
	TTTAAGATTA GCACGTGGTT ATACTGGCAG AAATAAAATT GTGAAATTTG AAGGTTGCTA	1980

55

GCCGGATTCT CTGGGTGTGC CTGAAGGTAT TGCTAAAAAT ACAATTACAG TTCCATACAA 2100
 TGATTTAGAT GCACTTAAAA TCGCTTTCGA AAAATTGGa AACGATATTG CTGGTGTAAT 2160
 5 CGTAGAACCT GTTGCTGGTA ATATGGGTGT CGTACCGCCG ATTGAAGGTT TTTTACAGGG 2220
 ATTAAGAGAT ATTACGACTG AATACGGCGC ATTGCTAATT TTCGATGAAG TAATGACTGG 2280
 10 TTTTACAGAGTC GGTATCATT GTGCACAAGG TTACTTTGGT GTGACACCAG ATTTAACTTG 2340
 CTTAGGAAAA GTTATCGGTG GAGGACTACC TGTAGGTGCA TTTGGTGGTA AAAAAGAAAT 2400
 CATGGATCAT ATAGCACCAT TAGGAAATAT TTATCAAGCG GGTACGTTAT CAGGAAATCC 2460
 15 TCTTGCAATG ACAAGTGGTT ATGAAACGTT AAGCCAATTA ACGCCAGAGA CATATGAGTA 2520
 TTTTAATATG TTAGGCGATA TACTTGAAGA CGTTTAAAA CGTGTATTTG CTAAACACAA 2580
 TGTACCAATA ACTGTAAATA GAGCAGGTTT AATGATTGGT TATTTCTTAA ATGAAGGACC 2640
 20 TGTAATAAT TTTGAACAAG CGAATAAAG TGATTTGAAA TTATTTGCAG AAATGTATCG 2700
 AGAAATGGCA AAAGAAGGTG TGTTTTTACC ACCATCTCAA TTTGAAGGTA CATTCTTATC 2760
 TACGGCACAC ACGAAAGAAG ATATTGAAAA AACGATTCAA GCATTTGATA CGGCTTTAAG 2820
 25 TCGTATTGTA AAATAAATAT ACGGACAAAT TGAGAGCCTG AACTTTGTTC AGGCTCnTTT 2880
 TAAATGTATA TAAGGCATGG GCGGCGACTT GATAGTGAAA GTCCACTACT A 2931

(2) INFORMATION FOR SEQ ID NO: 283:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1421 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:

40 AATTATGAAT GCATTACCAG TATTATTACA AAAGAACAAT TAAAAATGTT TGTTTATGAT 60
 TATGATACGC ATCTCATTAA AAATGTAAaTG GTTGACAGCAG ACGTGTTAAA GGCAATGAT 120
 45 ATTCAAGGAC ATGAACCATT AATCGTTAAC CTTCAAACGA TTGATGAAAC ATTACATCGT 180
 TTACCTATGC ATAATAGAAA AGACATGATG GTTAATGGCG GTGTACTTAT GGCACATTTA 240
 AATGCCAAAA GTGGTCCGTG GTTAAAAGAT GTGCTAAGAC AAATTGAGAT AGCGATTGTA 300
 50 ACAGGTAAAG TAAGCAACGA AGAACTGAA ATTTTGAAAT GGGTGGATAA TCATGTCAAA 360
 ATATAGTCAA GATGTACTTC AATTACTCTA TAAAAATAAA CCGAATTATA TATCTGGACA 420
 AAGCATTGCG GAATCACTTA ATATTTACCG CACTGCAGTA AAAAAAGTGA TTGACCAATT 480

CCCAGATATT TGGTATCAAG GTATAATAGA CCAATATACA AAAAGTTCTG CTTTGTGTTGA 600
 TTTTAGTGAA GTATACGATT CAATAGATTC TACACAACCTT GCTGCGAAAA AGTCACTTGT 660
 5 TGGAAATCAA TCTTCATTTT TTATCTTGAG TGATGAACAA ACGAAAGGTC GTGGGCGATT 720
 TAATAGACAT TGGAGTTCTT CAAAAGGGCA AGGACTTTGG ATGTCTGTCG TGTTAAGACC 780
 10 TAACGTTGCA TTCTCAATGA TATCTAAATT TAATTTATTT ATTGCATTAG GGATAAGAGA 840
 TGCATTCAA CATTTTAGTC AAGATGAAGT CAAAGTGAAA TGGCCGAATG ATATATTTAT 900
 TGATAATGGT AAAGTGTGTG GTTTCTTAAC TGAAATGGTT GCTAATAATG ATGGTATAGA 960
 15 AGCAATAATA TGTGGTATAG GTATTAATTT GACGCAACAA CTAGAAACT TTGATGAAAG 1020
 TATTAGACAT AGAGCAACAA GTATACAATT ACATGATAAA AATAAATTAG ATAGATATCA 1080
 ATTTTGTAGAG ATATTACTTC AAGAAATTGA AAAAAGATAT AATCAATTTT TAACGTTACC 1140
 20 TTTTCTGAA ATTCGTGAAG AATATACTGC AGCTTCTAAT ATTTGGAATA GAACGTTGCT 1200
 ATTTACAGAA AATGATAAAC AGTTTAAAG ACAAGCAATT GATTAGATT ACGATGGCTA 1260
 TCTAATTGTT AGAGATGAAG CGGGTGAATC ACACCGTTTA ATTAGTGCAG ATATAGATTT 1320
 25 TTAACACTAA AGCAAGGAGA GATAGCTATG GGTATGGCAA CCTATGCCGT TGTGGATTG 1380
 GNAACAACAG GCAACCAATT AGATTTTGAC GATATCATT C A 1421

30 (2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2202 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

CCAAGTTGCC TAAAATGATT AAGCAAGGTT TATACCCTAT GCAACGATT GAACAAGAAT 60
 CTGGAGCCAT CCGACTGCCA ACGATTCTTA GAGTGANGCG TTCATTACAA TGGGGTAATG 120
 45 ATGCTTATAC AATGATTTTA GATCGTATGA ATATTGAAAC AAATGAATAA TAAATGAACG 180
 ATAAACAATG GTTATCTATC TGCCTAATA AGGTAGATAA TCATTGTTTT TTCACGAAAA 240
 AATTTACAGA GTAAAAGAAC TTAAATTTCA TATTAAGTCT TTAGAACTCG AACTTAAAA 300
 50 ATGCTATAAT CATATGTATG TTAAAAAGG AGTTTCGGAA AATGTATGAC ATTAAAAAAT 360
 GGCGCCATAT TTTTAAATTA GACCCAGCTA AACATATTTT AGATGATGAT TTAGATGCGA 420
 TTTGTATGTC TCAAACAGAT GCAATTATGA TTGGTGGAA C TGATGACGTT ACTGAAGATA 480
 55

	CAAACATCGA AAGTGTAAATG CCTGGTTTTG ATTCTTATTT TGTACCTACA GTECTGAACA	600
	GTACAGATGT TGTATTTTAC AATGGTACAT TATTAGAAGC GCTTAAAACA TATGGACATA	660
5	GTATAGATTT TGAGGAAGTA ATATTTGAAG GGTATGTCGT GTGCAATGCT GATAGCAAAG	720
	TGGCAAAACA TACCAAAGCA AATACAGATT TAACAACAGA AGATTTAGAA GCATATGCCC	780
10	AAATGGTCAA TCATATGTAT CGATTACCGG TTATGTATAT AGAGTATAGT GGCATTTATG	840
	GCGACGTATC AAAGGTTCAA GCTGTCTCAG AACATCTAAC AGAAACGCAA CTTTTTTATG	900
	GTGGCGGTAT TTCCTCAGAA CAACAAGCGA CAGAGATGGC AGCTATTGCA GATACAATTA	960
15	TCGTCGGTGA TATTATTTAT AAAGATATTA AAAAAGCTTT AAAAACAGTA AAAATAAAGG	1020
	AGTCTAGTAA ATGAATGCGT TATTAAATCA TATGAATACA GAGCAAAGTG AAGCTGTAAA	1080
	GACAACAGAA GGACCATTGT TAATTATGGC AGGTGCTGGT TCAGGGAAGA CACGTGTTTT	1140
20	AACACATAGA ATTGCTTATT TATTAGACGA AAAAGATGTC TCACCATACA ATGTTTTGGC	1200
	TATTACTTTT ACAAATAAAG CTGCAAGAGA AATGAAAGAA CGTGTTCAAA AATTAGTAGG	1260
25	TGATCAAGCA GAAGTTATTT GGATGTCAAC ATTCCACTCA ATGTGTGTTC GTATTTTACG	1320
	TCGTGATGCA GATCGAATTG GTATAGAACG CAATTTTACG ATAATTGATC CTACAGACCA	1380
	AAAATCTGTT ATTAAAGACG TCTTAAAAA TGAAATATT GATAGTAAAA AGTTTGAACC	1440
30	TCGTATGTTT ATCGGTGCGA TCAGTAATTT GAAAAATGAA CTTAAAACAC CTGCAGATGC	1500
	TCAAAAAGAA GCCACAGATT aTCACTCgca AwTGGTaGCA ACgGTTTaTA GTgGATATCA	1560
	ACGCCAATTG TCACGTAATG AAGCGTTAGA TTTTGATGAC CTTATTATGA CAACGATTAA	1620
35	CTTATTTGAG CGTGTACCAG AAGTTCTAGA ATATTATCAG AACAAATTCC AATATATTCA	1680
	TGTAGATGAG TATCAAGATA CTAATAAAGC ACAATACACA TTAGTTAAAT TATTAGCAAG	1740
	TAAGTTTAAA AACTTATGTG TTGTAGGTGA CTCAGATCAG TCAATTTATG GTTGGCGTGG	1800
40	TGCTGATATT CAAAATATCT TATCATTTGA AAAAGACTAT CCAGAAGCGA ATACAATCTT	1860
	TTTAGAGCAA AATTATCGTT CGACGAAAAC GATTTTAAAT GCGGCTAACG AAGTGrTTAA	1920
45	AAATAATTCT GAACGTAAGC CAAAAGGACT GTGGACTGCA AATACGAATG GTGAGAAAAT	1980
	TCATTACTAT GAAGCAATGA CGGAACtGAT GAAGCGGAAT TTGTAATACG AGAAATTATG	2040
	AAGCATCAAC GTAATGGTAA GAAATATCAA GATATGGCAA TTTTATATAG AACGAATGCA	2100
50	CAATCACGTG TACTTGAGGA AACATTCATG AAATCTAATA TGCCATACAC AATGGTTGGT	2160
	GGCCAAAAGT TCTATGACCG TAAAGnAATC CAAAGATTTA TT	2202

(2) INFORMATION FOR SEQ ID NO: 285:

55

(A) LENGTH: 785 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

10	AGTGGTGCAA AGATAGGCAT TGATaATACC GCTAAGCCAG CAAGATGATG GCACGATAAA	60
	ACCTAAACAG AAGAaMATAA ATAGTAATAC GATGATAAAT AATGGTCCAC TCATATGTTG	120
	AACTAAAGAT GATGAAAAGT GTAAGATTGT ATCTGAAATC ATACCTTCAT TCAACACTAA	180
15	ATTAAATACCT CGAGCTAAAC CAATAATTAA AGATACACCT ACTAAACTTG ATGCACCATT	240
	GACAAATGCA TCTACAGTTC CTTTTCTCC CAATCCAGAT TTACCTGTCC CAGCAATAAA	300
	CATTATTATA ATTGTAAATA TTAAAAATGC TGAAGCCATA ACTGGGAACC ACCAACCTTG	360
20	CGTCATAACT CCCCATACCA TAATTGAAA TGGTAGTACA AATAATGTAA GGATTATCTT	420
	CTTACGCAAA GTAAAATGGG CACTATCGTC ATCTTTTAAT ACAGACCATT GCTGTTCAAA	480
	AGCATCTTTG TCTTCATAAG AATATGACGC TTTAGGATCG TTTTAAATT TTTTACAGTA	540
25	CCAAATATAA TAACTAATAA CAAAATCGC ACCGACAATA CAAGCACCTA TTCTCCAATA	600
	CAAGCCATCC GTAAAAGTTG TACCAGCGGC ATTAGAGGCA ATTACAACCG AGAACGGGTT	660
30	AATAGTTGAA AATGTACTAC CGACAGAGCT GGcAAGGAAT ATGGCACCAA CTGGAAACGA	720
	TAGAATCGTA TCCTAACGCT AATAAATATA GGGACTAAAA TCGGATAAAA TGCTACAGCT	780
	TCTTC	785

35

(2) INFORMATION FOR SEQ ID NO: 286:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 812 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

45

45	CTAACGnGAT AAGGTTGCAA nTTTATCTGA ACATCTGATG ACTGTAATTT TGTtAATGAT	60
	AAAATATTTG TCACTAATAG ATATAAATAC TGACTTTCTT GAAAACTATG TACAAGTAAT	120
50	TGTTCCTTTT CTATGATAGA CATATCTTTA CTATGTGATA CTAAAATATC TAAATkTCCC	180
	ATAATTGTTG TTAACGGTGT ACGTATGTCA TGCgAAATTG ATCTTAAAAA ATTTGAATGT	240
	GTCAGTTGAC GTTCAGCCTG TAACATGGAT TCTCTCGTTT GTTTAAGTAA CGTCACATTT	300

55

ATCACTTGAG AACTTTGGTA ATCAATGGCT AGAATGCCTT TAATCGGAGA TGTGCCAATT 420
 GGTATCAACC ATTTATTAAT GCCTGGAAAT GTATCTGTTG TTGCACCAGC TTGTCTTTCA 480
 5 TTTTAAATTA CCCAGCTTAA TGCTTGTTCA TGCTGTTGAG TCGTATTATC GATATGGTTT 540
 TGCAATGGTA TTGTTTTAAT TACTTTTCGAT TGATTGATAA CGTATATAGT AATTGATTGT 600
 TGCAATAATT GATTAAATTTG GTATCCAGCA TTTATTAGTA AGTTTTCAAC TGTATAAGTT 660
 10 TGTTTAATCG AATCATTAAA TTGAAATAAT AAATCTGTAC GATAAAGTTG CTTTTTAGTA 720
 ATGGaGTAWT GGAATTTAAT TTGTnTTAAT AAAGCACTCG TTAAAATACT TGTnAAAATG 780
 CTAACGATAA ATGTAATAGG ATAGTCAAAG CG 812
 15

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1732 base pairs
 20 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 287:

ATnnATTATT ATTACTGCTA TTTTAAATTT TAAAAAATGC TTTTGATTAT ATTCAACaTT 60
 TTGTATAAAA TTAAATTTGC TTTTGATTAA AGCATGAAAA TTGTAATCAA ACCATAAATT 120
 30 GTCGTATGAT GTAGTTAGAA TTTTAAATG CAGGAGGTCa AGTATATGAC TGAAATaACA 180
 TTCAAAGGTG GACCAATCCA CTTAAAAGGT CAACAAATTA ATGAAGGTGA TTTTGCACCT 240
 35 GATTTTACAG TGTTAGATAA TGACTTAAAT CAAGTAACAT TAGCAGATTA TGCTGGTAAA 300
 AAGAAATTAA TTAgtGTGGT ACCATCAATT GATACAGGTG TTTGTGATCA GCAGACTCGC 360
 AAATfCAACT CTGATGCTTC TAAAGAAGAG GGGATTGTGC TTACAATTTT AGCAGACTTA 420
 40 CCATTCGCAC AAAAAAGATG GTGCGCTTCA GCAGGTTTAG ACAATGTCAT TACATTAAGT 480
 GACCACCGTG ACTTATCATT TGGTGAAAAC TATGGCGTTG TTATGGAAGA ACTTCGCTTA 540
 TTAGCTCGTG CAGTATTTGT ATTAGATGCA GATAATAAAG TTGTTTATAA AGAAATCGTT 600
 45 AGTGAAGGTA CTGATTTCCC AGATTTTGAT GCTGCTTTAG CTGCATACAA AAATATTTAA 660
 TCATTAAAGA GATAAATCTT AAAATGTATA CATCGTGTCC ATCGTTGTCA ACAGCATTAA 720
 AATAGAATTG TTTTCTATGA TTGCTAAGAC CTATGGGCAC TTTTATTGG AGAGGGACGA 780
 50 ATATGGCAGA ACAACAAACA ATTATGGAAC GCTTGTTTCA TACATTAGAT GAAAAAGCTA 840
 AAACATTAAA TAATGAAAT GGCCaAAGTT TTATTGAAAA TCTTGGGCTA GCAATGGAAC 900

55

CATTCCAATT TGCATATTTA AGTTTAATGC aGGAAGAAAA GATAcAAGCA AATCATCAAA 1020
 TTACACCAGA TTCAATTGGA TTGATACTAG GATTTTTAGT TGAGCGTTTT ATGAACAACC 1080
 5 AAGAAGAATT ACATATTGTT GATATTGCAA GTGGTGCCGG TCATTTAAGT GCTACTGTAA 1140
 AAGAAGTGTT ACCTGraAtt GcGGTTATGc ATcATTTaAT TGaAGTTGAt CCAGTTTTAT 1200
 CACGTGTTAG TGTACATTTA GCAAACCTTCT TAGAAATTCC TTTCGATGTG TATCCTCAAG 1260
 10 ATGCCATCAT GCCACTACCA TTAGAAGAAG CAGATATCGT TATTGGTGAT TTTCCAGTAG 1320
 GCTATTATCC AATTGATGAA AGAAGTAAGG AGTTTAAGCT AGGTTTTGAA GAAGGACATA 1380
 GTTATTCACA TTATTTATTA ATAGAACAAG CAATAAATGC ATTAAAAGAT GCTGGATATG 1440
 CCTTTCTAGT GGTACCAAGT AATATTTTTTA CAGGTGAACA TGTAACACAG CTTGAAAAAT 1500
 ATATTGCAAC AGAGACAGAG ATGCAAGCAT TTTTAAATTT ACCACCAACT TTATTTAAAA 1560
 20 ATGAAAAAGC GCGAAAATCT ATATTAAATTT TACAAAAGAA AAAATCGGGT GaAACAAAGC 1620
 CAGTTGAAGT ATTATTGGCA AATATTCCTg ATTTCCAAAA TTCCTTCACC AATTTCCAAG 1680
 GATTTATGGA CAGAGTTAAA ATCCAGTGGG ATGGGACCAC CAAATCGTCC TA 1732

25 (2) INFORMATION FOR SEQ ID NO: 288:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2779 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:

AAAAGAACTA GCTAAACGCA AGCAAGAAGC TATTAGTAGA ATTAAGACT TTTCAAATGA 60
 AAAATAAAT AGTATTCGAA ATAGTGAAAT TGGCACAGCT GATGAAAAAC AAGCAGCAAT 120
 40 GAATCAAATT AACGAAATTG TGCTTGAAAC AATTAGAGAT ATTAATAATG CGCATACATT 180
 ACAGCAAGTT GAGGCTGCAT TGAACAATGG TATTGCTCGA ATTCAGCAG TACAAATTGT 240
 AACATCTGAT CGTGCTAAAC AATCGTCAAG TACTGGAAAT GAATCTAATA GCCATTTAAC 300
 45 AATTGGTTAT GGAAC TGCAA ATCATCCATT TAACAGTTCTG ACTATTGGAC ATAAAAAGAA 360
 ACTTGATGAA GATGATGACA TTGATCCACT TCATATGCGT CACTTTAGTA ATAATTTCCG 420
 TAATGTTATT AAAAACGCTA TTGGTGTGGT GGGTATCTCT GGTTTACTAG CTAGTTTCTG 480
 50 GTTCTTCATT GCCAAACGTC GTCGTAAAGA AGATGAAGAG GAAGAATTAG AAATAAGAGA 540
 TAATAATAAA GATTCAATAA AAGAGACTTT AGACGATACA AAACATTTAC CACTTTTATT 600

55

	AAATAATGGC GAGTCACTCG ATAAAGTTAA ACATACGCCG TTCTTCTTAC CAAAACGTCG	720
	TCGTAAAGAA GATGAAGAAG ATGTGGAAGT TACAAATGAA AACACAGATG AAAAAGTGTT	780
5	GAAAGATAAC GAACATTAC CACTCTTATT CGCAAAACGA CGCAAAGATA AAGAGGAAGA	840
	TGTTGAAACA ACAACTAGTA TTGAATCTAA AGATGAGGAC GTTCCTTTAT TATTGGCTAA	900
	AAAGAAAAAT CAAAAGATA ACCAATCCAA AGACAAAAAG TCAGCATCAA AAAATACTTC	960
10	TAAAAAGGTA GCAGCTAAAA AGAAGAAAAA GAAAGCTAAG AAAAATAAAA AATAATTTGT	1020
	TTCTTTTGATA AATAGaGGAG CACCGATTGA CATCACATCA GTCGGTGCTC CTTTTATTTA	1080
	TTCTTTTTAA TTAATTTATA CAATGCCTGT TGAGCGTGTT GATTCGCTTC TTGTTTTGT	1140
15	TCTCTCGGTA TCCATTTAAC AAATAATAAA TCAAAATCTT TTTCAAATAT TTCTATTTGA	1200
	TCAAAATAAG GTTTGAAATT TGCGTTTTTC ACATAACCAG CTTCAATGCT ATCTGCAATT	1260
20	AGCTTTGAGT CTGTATATAA TAGTGCGTTT TGAACATTTA ATTCACGTGC ATGTTCTAGT	1320
	GCATAAATAC ATGCAGCCCA TTCTGCAGTG TGGTTATCCA TTTGCGCTAA CTCATGTGTA	1380
	TATGTATAAT GCTGCTCATC TTCTTTGATT ACAATGGCAC ATGTACTTAT GCCTGGATTT	1440
25	cCTTTCGTCG CAGCATCAAA ATTTATTTTC GCCATAATAA ACCTACTTTC TATTCAATAC	1500
	TTAGTTAAAG TTACTATTAC TGTAATACAA AATATGTTGG GTAATCCATT AAAAAACACG	1560
	CATCACTTAA ATAAGTAACA CGTGTTTAAA ATACTCGCTG ATTCAAAGAT GATTTTCTAA	1620
30	TACGTaACT GTaATATACT TCCTAAAAAA ATCATCTTCA GGCTGGGACA TAAATCAATG	1680
	TTCTATGCTC TACGATGTTA TATTGGCAGT AGTTGACTGA ACGAAAATGC GCTTGTAACA	1740
	AGCTTTTTTC AATTCTAGTC AGGGGCCCCA ACACAGAGAA TTTGAAAAG AAATTCTACA	1800
35	GGCAATGCGA GTTGGGGTGT GGGTCCCAAC ACAGAAGATG ACGAAAAGTC AGCTTACAAT	1860
	AATGTGCaAG TTTGGGATGG GCCCCAACAA AGAGAAATTG GATTCCCAAT TTCTACAGAC	1920
40	AATGCAAGTT GGGGTGGGAC GACGAAATAA ATTTTGCgAA AATATTATTT CTGTCCCACT	1980
	CCCTTAAAAC TTATTCTTTT GTGTAGTAAG TCGGTAAATA GCCTTGATCT AACTTATCAA	2040
	TCTTACCTTT ACGATAAAAT GATTTAGCAA TATATCCAAA TGGTACATTG AAAACTGTTG	2100
45	AAGCTAATTT TAATACGTAC GTTGTAATAA ATATTTCAaA TACAaMTGta CCAGGTAAAC	2160
	TTCCGATAAA TCGGATAGCT ACAAATAAAG CTGTATCAAT TATTGgGCTT AAAAATGTAC	2220
	TACCATATGg CACGGATGAA AAACGTTTTa TCMGaACTAA ATACTTTTTT AATTAGTGAA	2280
50	AAGATAAATA CATCAATATG TTGACCAATA ATATATGCGA CGATTGAGCC TAAAGCAATG	2340
	CGTGGCACAA CATCAAAGAT TGCGTGTAAT GCTTTTTGTG CCATATCTTC TGGTGCAGGA	2400

55

CAAACCTGCTC TTTTGTCAAC TCTACGCCCA TAAATATCGT TTAATATATC TGTTGCTAAA 2520
TAAATAGAAG CAAACATGAC ATTACCTAAA GTTGCTGAAA TACCAAAGAT TTCTACAGTT 2580
5 TTAATCACTT GTATGTTGGC AATGATTGTG CCAATTGCAA CCCATGCAAT TAAACCTTGT 2640
TTACCAAAAA AGCGATACAT AAGTACCATA AGCACGAACG TTGCAATAAA CGTAACTAGT 2700
CCTAAAATTT CATTATACAT ATTAAAATGT CCTCCTAAAT TTTGATCATG CGGGTGTTTA 2760
10 GAAACCGCTC AATAAATAA 2779

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

ACTGATGTGC GTTCATCAAA AACAATATAA TCAAATTCAT TTTCATCAAA TGCTTAAAA 60
25 TTATCATCTC TAGATAATGT TTGAATGGTT GCAAATAAAT-ATTTGGCATC GACATCTCTA 120
TGTTTTCTCG TCAACAATCC AAAATCACTA TCATTTTTTA TAGGTAATAC TTTTTTAAAT 180
TCTTCCTTAG CTCTATTTAA AATCCCCTCA TTATGAACAA TAAATAAAAA TTTATTAGGG 240
30 TTTACTTCTC TAACATCTAA TGCACATAAA ATCGTTTTAC CTGTACCAGT TGCAGATATT 300
ATTAACGCCT TATCTTTGGC TTTATCCCTA ATAGCTTTTA ATGACCTTAA TGCTTCTGCT 360
TGCAATAAAT TGGGTACAAT TTCCACTGAT TTTTTCACCT TATCAGCTAG CAGCATTGTA 420
35 GTTTGTTCAA CCTCCGCTAA TTTTCTAAG GAACGGTACT CAAATGATTC TTTATATGAA 480
TTAATCCATT GCTCAGTCAG TGGGGTACTC TTTTGCCATA ACAAGTCAAA TTCACTTTTT 540
40 ACACTATCAA CTAAATCGCC ATTTTTCATA GTAGACAGTA AAACATTATG CTCATAATTA 600
ACCTTTAACG CATTAGATGT TAAATTAGAG CTTCTTATTA CCATAGAACT ATAATCCTTA 660
TGCTCAAAAA TATATCCTTT GGCATGGAAT CCAGCAATAT CAGTTAATCT TACCTCTACA 720
45 TTTTTTAATT TAAGTAATTC TCCATACATT TTAGGACTAT TAAACCCTAA GTAATTAGAT 780
GTTAATATTT TCCCTTTAAC ACCCTTATTG CTTAAATCTA ATAGTTGAGC CTTTAAGCTG 840
GcTAAACCGC TTTCTGTTAT AAAAGCCACA GAAAAATAAA ACGTTTCACA TTTTGAAGT 900
50 TCATCTATAA TTGTTGAAAG AACTTTTTCA TTTTATTAT TACTAAAAG CTTCGGTGTA 960
TAATTCCTT TATGAGAAAT ATGTTTGTCT ATAAACCCTT TATGTAAAGA TGATTGAAA 1020

CGCTGGGGCC CAATTTAATT TATCAAGTTC GTTTATTGAC AACCATTCAA TACTCTTATG 1140
TTCAGTTAGA GTTGTAAC CTCTTTTAA AGTACATTG TATGTTGTTA ACCTAACAAT 1200
5 TCCAAAATCA TATTCATGTT CTGTAGTTAT AACTTTGTCT CCAACAATTA AATCACATTT 1260
CATTTCTTCT CTAATTTCTC TAATCAAAGC GTCTTTTCA GTTTCATTCT TTTCAACCTT 1320
ACnGCCAGGA AATTCCCACA TTAAAGGCAG ACTCATTTTT TCACTTCTCT GTGCACAAAG 1380
10 AATTTTGTGA TCAGAAAAAA TAATAGCTCC TACTACATTG ATTACTTTTT TCATAAGACT 1440
CACCCTTCAA TTTAAATCA TCTTAATTGT TATTCTATCA AAAATTACAA AACTATATAT 1500
AAATCAATAT TAAAAATTAA TATTTTACAT TCACATGAAC GCTCTACTCC ATGCATTTTC 1560
15 ATACACATCT ATTATATAAT ACTTGTGAAA AGTATTGTCT TGGGGCTGTG TTTTTTTACT 1620
TTTGGGGCGT ATTTCTTTAT AATTCATTAC ATAAATGTAA GGGCTTTAGT TTTTCATGTTT 1680
20 TATTAAGTCT AACTGAGATT TTGAAAGGAT GTTTAGCAAC AATGGATAAA GAATTATGGA 1740
TAGAACGAGc TAATGATAGT TTAGTTAAAC ATTTTTATGa GCAGCAATCT GACATTGAAC 1800
AGCGAGAAGG TTTTGAAAGT AAATTAACAT TTGGTACTGC GGGTATACGC GGAAAATTCG 1860
25 GTCTTGGTGA AGGTGCACTT AATAAGTTTA CTATTGAAAA ATTGGCATTG GGTTCAGCGC 1920
GTTATTTAAA TGCCCAAACA AACAGTCCAA CAATAGTCAT TCATTATGAT ATTAGACATC 1980
TTTCcAACTG AATTGCCCC 1999

(2) INFORMATION FOR SEQ ID NO: 290:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1933 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

GATGACTTTT CCCCCTCATA ATCTTCATGG TCCAGGCGTC CATTAAATGCG TCAAAGGATG 60
GCACATTTTA CCTGGAACAA ATGATTTCATA TGGTTCATAA AAATCACGCG TCGTAATATA 120
45 ATCTTCTAAA TCAAATGCAT AGAAAATCAT TGGCTTTTTA AATACTGCAT ATTCATATAT 180
TAAAGATGAA TAGTCACTAA TTAATAAATC TGTTATGAAC AGTATATCAT TAACTTCTCT 240
AAAGTCAGAA ACGTCAACAA AATATTGTTT ATGTTTGTCT GCAATATTAA GTCTATTTTT 300
50 CACAAATGGA TGCATTTTAA ATAATACAAC CGCGTTATTT TTTTCGCAAT ATCTTGCTAA 360
ACGTTCAAAA TCAATTTTGA AAAATGGGTA ATGTGCTGTA CCATGACCAC TACCTCTAAA 420

	TTGTTTGATC TGTGTCGCAT AAGCTTCATC AAATAGTACA TCAGTACGTG GAACACCTGT	540
	AGGCACTACA TTTTCTCTT TAATACCAA TGCTTCAGCG TAGAATGGAA TATCGGTTTC	600
5	AGATGATACA TAAGCTTTTG TATAGCTACG ATGATTTAAT GAATCAATAA ATGGTCCACC	660
	CTTTTACCA GTACGACTAA AGCCAACTGT TTAAAGGCA CCAACGGCAT GCCATACTTG	720
	AATAACTTCT TGAGAACGTC TAAAACGCAC TGTATAATC AATGGGTGAA AGTCATCAAC	780
10	AAAGATGTAG TCTGCCTTCC CAAGTAAATA TGGCAATCTA AACTTGTCGA TGATGCCACG	840
	TCTATCTGTA ATATTCGCTT TAAAAACAGT GTGAATATCA TACTTTTTAT CTAAATTTTG	900
	ACGTAACATT TCGTTATAGA TGTATTCAA GTTTCAGAC ATCGTTGGTC TAGAGTCTGA	960
15	TGTGAACAAC ACCGTATTCC CTTTTTCAA GTGGAAAAAT TTCGTCGTAT TAAATATCGC	1020
	TTTAAAAATA AATTGTCTTG TATTAAATGA TTGTTTGGCG AAATACTTAC GTAATCTTT	1080
20	ATATTTACGA ACGATATAAA TACTTTTAAAC TTCCGGAGTC GTTACAACAA CATCAAGGAC	1140
	AAATTCATTA ACATCGCTAG AAATTTTCAGG TGTAACAGTA TAAACCGTTT TCTTCGAAAT	1200
	GCCGCCTTTT CTAAATTCTT TTAGGTAAGT CTGCAATAAG AAATTGATTT TACCATTTTG	1260
25	TGTTTCTAAT TCGTTGTATT CTTCTTCTTG TTCTGGCTTT AGATTTTGAT ATGCATCATT	1320
	AATCACATCT GGGTTTAACT GTGCAATATA ATCAAGTTCT TGCTCATTCA CTAATAAGTA	1380
	CTTATCTTCA GGTAAGTAAT AACCATTATC TAAGATAGCT ACATTGAAAC GACAAACGAA	1440
30	TTGATTCCCA TCTATTTTGA CATCATTCGC CTTCAATTGTA CGTGTCTCAG TTAAATTTCT	1500
	TAATACAAA TTACTATCTT CTAAATCTAG GTTTTCACTA TGTCTTCAA CGAATAACTG	1560
	AACACGTTCC CAATAGATTT TATCTATATA TATCTTACTT TTAACCAACG TTAATTCATC	1620
35	CTTTTCTATT TACATAATCC ATTTTAATAC TGTTTTACCC CAAGATGTAG ACAGGTCTGC	1680
	TTCAAAGCT TCTGTAAGAT CATTAAATTGT TGCAATTTC AATTCTTGAC CTTTAAACAA	1740
40	CGCTAATTTA nCTACAATAT CTGGGTATTG AATGTATAAG TCTACAACAT CTGGGAAATC	1800
	TTTTGAACCA CTTGCACTAC TACCAATCAA CGTTAGTCCT TTTTCCAATA CTAGACGTGT	1860
	ATTAACCTTCT ACTGGGAACT CACTTACACC TAACAGTnCA ATGCTTCCTT CTGGTGAAAT	1920
45	GTAATCGATC ATT	1933

(2) INFORMATION FOR SEQ ID NO: 291:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2049 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

	nGTnCGGnCA GATATATTGG TGGTCTTTAG TAAGTGTATC AAATTCATCA GATGTCAAGG	60
5	GCATGTTATC ACCTCCTTAG GTTGATAACA aCATTATACa CGaAAGGAGC ATAAaCAAaT	120
	GAACACAaGA TCAGAAGGAT TGC GTATAGG CGTCCCACAA GTTTCTAGCA AAGCTGATGC	180
10	TTCTTCATCC TATTTAACGG AAAAGGAACG TAACTTAGGA GCGGAAATAT TAGAACTTAT	240
	TAAAAAAGT GATTACAGCT ACTTAGAAAT AAACAAAGTT TTCTATGCAT TAGATAGAGA	300
	ACTTCAATAC AGGGCGAATA ATAACAACT TTAaCATTa TCTAAAGGAG TGATAGAGAT	360
15	GCCAAAAATC ATAATACCAC CAACACCAGA AAACACATAT CGAGGCGAAG AAAAATTTGT	420
	GAAAAAGTTA TACGCAACAC CTACACAAAT CCATCAATTG TTTGGAGTAT GTAGAAGTAC	480
	AGTATACAAC TGGTTGAAAT ATTACCGTGA AGATAATTa GGTGTAGAAA ATTTATACAT	540
20	TGATTATTCA GCAACGGGAA CATTGATTAA TATTTCTAAA TTAGAAGAGT ATTTGATCAG	600
	AAAGCATAAA AAATGGTATT AGGAGGATTA TCAAATGAGC GACACATATA AAAGCTACCT	660
	ATTAGCAGTG TTGTGCTTCA CGGTCTTAGC GATTGTACTC ATGCCATTGC TGTACTTCAC	720
25	TACAGCATGG TCAATTGCAG GATTGCGAAG TATAGTGACA TTCATATTTT ATAAGGAATA	780
	CTTTTATGAA GAATAAAAAA ACTGCTACTT GTTGGAGCAA GTAACAGTGC AAGATGAGCA	840
	ATTGTCTTAA ATAATTATAT AAGGAGTTAT TAATATGACC TTACAACAAA AAATACTATC	900
30	ACATTTTGCA ACATATGACA ATTTCAATTC TGATGATGTT GTTGAACTT TTGGGATATC	960
	TAAAACACAT GCAAAATCCA CACTTTCAAA ACTTAAGAAA AAAGGAAAGA TTGCAATGGA	1020
35	AAGTTGGGGT GTCTGGCGTG TTATTGAATC GCAATTGCAT TTAAGTGTAG TCGAACGTAA	1080
	AAAAGAAATT TTAGAAGAAC AATTTGAATT GTTAGCAAGA TTAAATGAAC AAAGTGATGA	1140
	CCCTAGAGAA ATAGAAGAAC GTATCAAGTT AATGATTCTG CTAGCTAACC AATTTTAAGG	1200
40	AGGAGTTAAT CAATGGCAGT ATTAGAAGGT ATTTTGAAG AATTAAACT ATTAATAAG	1260
	AACTTACGTG TGTTAAACAC TGAACATCA ACTGTAGATT CATCAATTGT ACAAGAGAAA	1320
	GTAAAGAAG CACCAATGCC AAAAGAAGAA ACAGCTCAAC TGAATCAAT TGAAGAAGTT	1380
45	AAGGAACTT CTGCTGATTT GACTAAAGAT TATGTTTTAT CAGTAGGAAA AGAGTTCCTT	1440
	AAAAAGCAG ACACCTCTGA TAAGAAAGAA TTTAGAAATA AACTTAACGA ACTTGGTGCG	1500
	GATAAGCTAT CTACTATCAA AGAAGAGCAT TATGAAAAA TTGTTGATTT TATGAATGCG	1560
50	AGAATAAATG CATGAAGCTA GATCACTCAA ATAGAGCTCA TGCAAAGCTT AGTGCAAGTG	1620
	GAGCAAAACA ATGGCTAAAC TGTCCACCGA GTATTAAGGC AAGTGAAGGT ATTGCAGATA	1680

55

5 GTCTTAAATA TGAAGGCCTA ACACAGTTTG AGTTTAATAA AGCTTTTCAA AATTATAAGC 1800
 GAAATCAATA TTACAGTGAA GAGTTGCGCG AATATGTTGa AGAGTACGTA GCTAATGTAG 1860
 AAGAAAAGTA TAACGAGGCT TTGaGTAGAG ATGACGATGT AATAGCTTTA TTTGAAACAA 1920
 AATTGGATyT AGGTAAATAC GTCCCTGAAT CTTTTGGTAY TGGTGATGTC AtTATATTTT 1980
 10 CAGGTGGTGT ACTTGAAATT ATTGACCTTA AATACGGTAA AGGCATTGAA GTTTCAGCTA 2040
 TAGATAATC 2049

(2) INFORMATION FOR SEQ ID NO: 292:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 942 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

25 ATGATGTTTC TATATTCGTA TTAGGAAAAC CTGTTGGTAT TACAACAAAC GCCCTAAAAT 60
 TACAATACCG CTGCTCCTAT ACCAATTGCA ACAACAGTTT TAACTGAAAT ATCTTGTTTT 120
 TTCATCTTCA TTACTCCTt ACATAAAAAA TTCATTATAT TGATGGTGCT TTAGATAAAT 180
 30 GAATCGTCCA ATCATTTCCTA GTACCAATAT GATATAAATC TGAAATGAG TCTTGATTGA 240
 CTGCTACACC AATATTTTACT AGCGAGTTAA CATAACAAG AGGTTCCACC ACATTAACAT 300
 CTGCAAACGA TCGCGCAAAT TTAATAATAT TTTGATAGAC TTTCTTATCT TGATGATAAA 360
 35 TTGTTACCAC CAAATTATTA CCATGAACAA TTTCCAAGGA TTTTAAGAAT GCCAATGGAA 420
 TATTTGTCCA TAATGACCCA AATCTGATAT CTAAATATC AATGCTTCCT GTAACAGAAT 480
 CCTCATTITT TGTCACTTCT CTTATTTCTA ATGCCTCAAT ACTATCAACA TTAATTGCCT 540
 40 GACCGAGACG TTCAAACGCT ATCTTATTTG CAGCTAATCT CGCACCATTG TATGCATAAA 600
 CATCTCTACC ATGAAAAGTA TGACTTTCTT CCGAATGAGG CAATCGGCTT TTCACTTCAT 660
 CAATTTTCGAT AACTTTTTTTA ATACCTTCGT AATGTTTAAT ATGACTTAAA GAACCATTAT 720
 45 CAGGTGTAAT AATGTAATGA CCTGAATATG TTAAGCAAGC AATGTCCGCC TATCACTACC 780
 TACACCCGGG TCTACCACTG ATACAAaAc TGTGCCTTTA GGCCAGTATT TTACAGTTTG 840
 50 ATATAAACGA TATGACGCTA CCCAAATGTC ATACGGTGGT ATATCATGCG TTAAGTTTTTC 900
 AACACGTATA TCATCATTAC AGTATATGCA ACTCCATACA TT 942

(2) INFORMATION FOR SEQ ID NO: 293:

55

(A) LENGTH: 1268 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293:

10	TGTAATAAAA TTTTATGnAA CATGCTGnGA TGCTACCATG AACCTTCTTC ATTTCTTTAT	60
	GTGAGATTGT GaAATTAATC AAATAATAAT ACGGTGGATA CTTTCCTAAT TTACGATATT	120
	CCATTTCTCTG ACGATAAAAT GTTAAATAAT CATTTTTTTG AACATCCAAT ATTGAATAAT	180
15	GATCTGGATT ATACGTTTGA ATGATGACTT GACCTGCCTT TTCATGACGA CCAGCTCTAC	240
	CAGCCACTTG CGTTAATAGT TGATAAGTAC GTTCGCTCGC CCGAAAATCA GGTAAATTTA	300
	ACATTGTATC TGCATTGAGC ACACCAACTA AAGTAATATT TGGATAATCT AATCCTTTTCG	360
20	CAATCATCTG AGTACCTAGT AAAATGTCAC CGTTACCTTT TTCGAATTCA GTCAATAACT	420
	TTTCATGTGC ACCTTTCTTT GAGGTTGTAT CTACATCCAT CCTAATTATG CGCGCATCTT	480
	CAAATTCTTG TTGCAATAGT TCTTCAACTT TCTGAGTACC AGTACCTACT TGTCGAATGT	540
25	GTTCACTCTC ACAATTTGGA CATTGATTCTG GTGGCGTCTC TTGGTAACCA CAATAGTGAC	600
	ATTTTAATAA GTCTGTCGTT TTATGATACG TTAATGAAAT ATCACAGTTT GGACATTGCG	660
30	GTACATATCC ACAATCCCGA CATAACATAA ACGATGCATA ACCACGTCGA TTTAAAAATA	720
	AAACAACCTG TTCCTGTGCA TCTAATCTTA ATTGTATGGC TTCACGTAAA TCTTTTGAAA	780
	ACATTGACCG ATTACCTTCA CTCAATTCTT CACGCATGTC TACTATATCA ATTCAGGTA	840
35	AAGCTTGTTG GTTCACTCTG TTTGGTAATG ATAGCAAATG ATAAACGCCT TTTTCAGCTC	900
	GTGCATAACT TTCAAGACAT GGTGTTGCAC TTCCTAAAAT GACTGGACAG TGATGATATT	960
	CAC T TCGCCA TTGGGCAATT TCTCTAGCGT GATATCTCGG ATAATCTTCT TGT T TATATG	1020
40	TAGATTCATG TTCTTCATCA ATGATGATTA ACCCTAAATT TTTGAAAGGT GCGAACACAC	1080
	TTGACCTTGC ACCAACAATT ACTCTCGCAC GACCATCCCT AATTTTTTGC CATTATCAT	1140
45	AACGTTCCCC ATTAGATAAG CCAGAATGTA ATACAGCAAC GTCATCACCA AATCGACGTT	1200
	TGAAGCGTAA AACCATTGTC GGTGTTAGAG CGATTTCAGG AACTAACATC ATCGCCTGTT	1260
	TTCCTTGG	1268

50

(2) INFORMATION FOR SEQ ID NO: 294:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 629 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

5 TACCACCAAA TAATATATTA GCTGGCATT TAATAACATT TAnATTGTC ATGATATCAT 60
 CAATAAAATG TTGAAACTTC GTAATTTAC CTTCATAATC ATCAATTGCT GCTAATTGCG 120
 CACTCGATGC TTGCTGATCT AAATTTAAAA TATTCGACAT GCGTTGACTA TAATAAACTA 180
 10 AATGTTCTAT TAAGCCATCG TCACTCTTTT CCTTTGGTGC TGACATGACA GCGATACGTT 240
 TCAAAGGATA GTGTTGCGCC AATTTTAATG TCATTAATCC ACCTAAAGAC ACACCCGTTG 300
 CACTGATAGA TTCATAACCT TCATTGACTA AAAATTGGTA AGCTTTCTCA ACTTCTTCCC 360
 15 ACCAATCATC TACATTATAT GTCATGAAAT CTTTCAACAA TAAACCATGA CCTGGATAAT 420
 TCGGTGCATA ACAACTAAAT CCTTGGTCAT TTAACTCAGC TGCAAGATGC TTCACATCCC 480
 20 GATTTGTACC TGTAATGAA TGTAATAATA ATATCGCATG TCCATTTGTG CCTTTTAAAT 540
 ACGtGGACTC GGTGTTTTAA TTCTCATTTT TctatATAcC TCCACTatGT CTAAAGakGT 600
 TkGCTAAACG CGTTGtCGTC GATGATTAA 629

25 (2) INFORMATION FOR SEQ ID NO: 295:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2817 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:

35 TATGAAAGTA ATGAATGGTA ATATTATTAA ATTTGATGGA AAAGTAGATA TTGATAATGC 60
 AGATAATATC GGTTTTTTAA TTGAGCATCC TAAATTATAT GATAATAAAT CAGGATTGTA 120
 40 TAACTTGAAA TTATTTGCAC AAGTATTAGG TAAGGGTTT GATAAAGCAT ACACAGACAA 180
 AATTATAGAT GCATTTGGTA TGAGACCTTA TATTA AAAAG AAAGTTAAGA AATATTCAAT 240
 GGGGATGAAG CAAAAGTTAG CAATTGCACT ATCTTTATG AATAAACCTA AATTTTAAAT 300
 45 CTTGGATGAG CCTACAAATG GkATGGATCC AGATGGCTCa ATTGATGTGC TGACTACAAT 360
 TAAGTCTTTA GkAAATGaAC TTGATATGAG AATTCTaATA TCAAGTCATA AGTTAGAAGA 420
 TATTGAATTA ATTTGTGATA GAGCTGTATT TTTAAGAGAC GGnCATTTTG TTCAAGATGT 480
 50 AAACATGGAG GAAGGTGTTG CATCTGACAC AACGATAGTT ACTGTTGATC ATAAAGACTT 540
 TGATAGA ACT GAAAAATATC TTGCAGAGCA TTTCCAATTA CAAAATGTCG ACAAAGCAGA 600

55

	ATTAGATATT TATCCGAAAT ATATTGAAAC ACGTAAAAGT TCATTGCGTG ATACGTACTT	720
	CAATATAAAT CAAAGAGGTG ATAAATAATG AGAATTTTAA ATTTAGTTAA GTATGATTTT	780
5	TATAGTATAT TTAAbATCC TTAAACATAT TTAGCGaTAC TAGTCGTATC TAGTTTGATT	840
	GCAACTCAAA GTATACTTAT GGCAAATTCG ATGGATAACC CGAAACATAT TATTGTCTAT	900
	GGATCTGTAT TTGCTGCAGC AAAATGGTTA TTGTTAATAA TTGGATTAAT GTTTGTTGTT	960
10	AAGACAATTA CGCGTGATTT TTCACAAGGT ACAaTTCAAC TATATATGAG TAAAGTTAAA	1020
	ACACGCGTTG GALACATTAT TTCGAAAACA AMTCAATTa TTTTAATTTT AATATTATTT	1080
15	GCATTAATTC ATTATGTGAT TTTGATTGTT GTGCAGGCAT CTAGTAATGG AAAAAATTTG	1140
	GCGTTTTCTA AATATGTAGA TAATTTATGG TTCTTCCTAA TCTTTTTACT ATTCTTTGGC	1200
	TTGTTTTTAT TCTTAATCAC ACTTGCATCA CAAAAACAG CAATGATATT TTCATTAGGT	1260
20	GTATTTTTAG TACTCATTGT ACCGTTTATT AAACCTTTTA TTACATTTAT CCCAAGATAC	1320
	GGTGAAAAAG TTTTAGATGC TTTTGATTAT ATCCCTTTTG CTTACTTAAC TGATAAAATG	1380
	ATTAGCTCTA ACTTTGATTT TAGCAATTGG CAATGGGTAA TTTCAATTAGG TTCTATAGTG	1440
25	ATATTCTTCA TTTTGAATAT CTTATATGTC GCTAAAAAG ACATyTAATA AAAATAATTT	1500
	TGAGGTTGGG AATTTTAAAT TTTCCCAACC TCAAAGTTTG TCTTATTGTA AATTTATTTA	1560
	TTTTCTAATT TATTTAGGAT GGAATTATAA ACTGCTTTCC AAAATGAAGC GTCAGTTTTA	1620
30	TAGCGGTTTG ATATAACTAA GTGTGTTTCT TTTTCTAAAT CTGCATAGTC TGGATGATCT	1680
	TTGCTCGGTA ATTTATCAGC ACGAACATCA GTTACAAATT TTTGGACTTC ATTTGCTCTT	1740
35	GGTCCCCAAA CTGTTTCTTG TTCGAATTGA TCATTCAAGA ATACGAAGAT AGGAATTGCA	1800
	CGTGATTTAC CATTTGTAA ATATTGATCG ATCAGTTTTG TATCATCATC TCTATGGAAC	1860
	ACGCGTACTT CTAAATTTAA TGCTTCACTG ATGTGTTTTA GAATTGGGAG ATTCATCATT	1920
40	GCATCTCCAC ACCAGTCTTC AGTAATTACT AATACTTTAG AATAATTCAT CTCTTTTATT	1980
	TTTTTGATGC GTGAATCATC TTCTGGTAAC TCAATGATT GATAGATACT GAGAACGGTA	2040
	TCTTGATTTG TCTTCATTCC ATCAATGTAT TCATTTAAGG GTTGGCTATT TTTGAAATAA	2100
45	GTTTCTAAAT TTGTCATTGT AAAAACCTCC TTTAGCATTT ACAACATTAT ACCAATTTAT	2160
	AGTAATAAAA GGTAATGAAA TAAATTAATT GCAAATTCCT TGTTAATTTT TGTTAAGGAT	2220
	GAAACGGGAA GCACCTTATG CTATATTTAA ATAAGTACAA AGAAAGGGTG ACATCAGTGC	2280
50	GTATTCAAAA TCGCTGGGTT GTGTTTATAT TATTTTTAAT CTGTTCTTTT GGTGTATTAA	2340
	TTGGTTTATA TCAATATCGT CATACGAAAA CTGTGGATTT GTCTAATCTT GAAATAAACG	2400

55

TTGATCGTTT TAAATTTTAT AACAGTAAAG CTCACCCTGA TCTTACCGTT AAAGTGAGAG 2520
 AAAAGGATAA CATCGTTAAG GGGATAATAT TAGTAAGAGA TGAAAAGATA CATACTAATT 2580
 5 TTGATGGGGG AATTGGTTCG CCGATAAATA ACGCGATTGA AAATCTTGGa TTCgGATATA 2640
 AAAGrACaAA AGTTGGcAAT GrtTkCtCAT CgGTAAAGTA TATTGATAGA GATAACCATT 2700
 TAAAATTAAA CTTACTTTTAT CAAGATTTAG AAATTAAACG TATTGAATTT TTTAGTAAAT 2760
 10 AGCTTTAGGT CTTAAAGTTw TAAAAAACGA ATGAaTAATT TTATTGGGAT GAGTGAC 2817

(2) INFORMATION FOR SEQ ID NO: 296:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1607 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

TCTGTAAAA TGATTTTTCT TTTAnAAAGG CCGnAAATCA ATGTTGCGATT nTTATTTGCA 60
 25 TTATGGTCTC GATATTGGTA GAATATCAAA TGGTTAAATG AGAAAACTT GGAGGTGCTC 120
 ACATGTCATC AATCGTAGTA GTTGGGACAC AATGGGGAGA CGAAGGAAAA GGAAAAATAA 180
 CGGATTTCTT GGcAGAACAG TCAGATGTTA TCGCGCGTTT TTCAGGTGGT AATAATGCAG 240
 30 GCCATACCAT TCAATTTGGC GGAGAAACAT ATAAATTACA TTTAGTACCA TCTGGTATCT 300
 TTTACAAAGA CAAATTAGCG GTAATCGGTA ACGGAGTCGT TGTGATCCA GTTGCACTAT 360
 TGAAAGAATT AGACGGATTA AATGAACGTG GCATTCTTAC AAGTAATTTA CGTATATCTA 420
 35 ATCGTGCGCA AGTGATTTTA CCATATCACT TAGCACAAGA TGAATATGAA GAACGTTTAC 480
 gTGGgGACAA TAAGATTGGT ACAACTAAAA AAGGTATCGG TCCAGCATAT GTAGACAAAG 540
 40 TTCAACGTAT CGGTATTCGT ATGGCAGATT TACTTGAAAA AGAACATTTC GAAAGATTAT 600
 TAAAATCAAA CATTGAATAT AAACAAGCAT ATTTCAAAGG TATGTTTAAC GAAACATGTC 660
 CATCATTTGA TGATATCTTT GAAGAATATT ATGCAGCAGG TCAACGTCTA AAAGAATTTG 720
 45 TAACAGACAC ATCAAAAATC TTAGACGATG CATTGTAGC AGATGAAAAG GTACTTTTCG 780
 AAGGTGCGCA AGGTGTAATG TTAGATATCG ACCATGGTAC ATATCCATTTC GTTACATCAA 840
 GTAATCCAAT TGCAGGTAAC GTTACTGTTG GTACAGGTGT AGGTCCTACA TTCGTTTCAA 900
 50 AGGTAATTGG TGTATGTAAG GCTTATACAT CACGTGTTGG TGATGGTCCA TTCCCTACTG 960
 AATTATTCCA TGAAGATGGA CATCATATTA GAGAGGTTGG TCGTGAATAC GGTACAACAA 1020

55

TAAGTGGTAT TACAGATTTA TCTATTAAGT CAATCGATGT TTAAACAGGC CTAGACACAG 1140
 TGAAGATCTG TACAGCTTAT GAATTAGACG GTAAAGAAAT TACTGAGTAC CCAGCAAAGT 1200
 5 TAGATCAATT AAAACGTTGT AAACCAATCT TTGAAGAGTT ACCAGGTTGG ACAGAAGACG 1260
 TAACAAATGT GCGTACTTTA GAAGAATTAC CTGAAAATGC ACGTAAATAT TTAGAGCGTA 1320
 TTTCAGAATT ATGTAATGTA CAAATTTCTA TCTTCTCAGT TGGTCCAGAT AGAGAACAAA 1380
 10 CAAACCTATT AAAAGAATTG TGGTAGAACT TTATATAAGT CATAACAAT GATTATAAAT 1440
 ACATGAGCCT TCTATCTTTA TTGGTAGGAG GCTTTTGTTA TGCTTGCTTC TGTATCGATT 1500
 15 CGATTATTTA GATAAAAAAT ACTAACGTAA AGGCGATATT TGCTAGTCAT AATTTAGAAG 1560
 rTTAgATGAt AtTtAACGAA AAtTAaGATG anATACtTGA ATGGTAA 1607

(2) INFORMATION FOR SEQ ID NO: 297:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3055 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:

TTAGAAGTAA GCACTTTAAT ATCTTTACCT ACCCATGTAC CAACACCTTC TTTAGGCTTT 60
 30 GGATTTTCAG CATGGTTATT TGATTTATTC ACCTGTTTAC ATCTATTTAC TTTATTACCT 120
 TTATTCGGGT TCTCTAATAC ATCAAATTTT AATCTCGGTG AATAAAAAAG ATATATTAA 180
 AATCCAATA AAAATAATAC ACCGACGACT CTTATAATTA ACTTTTTTCAT CAATCAACCA 240
 35 CCTAAAAAGT ATTAATACTA TTGTAAAAAA CAACACATTA ATTAGCAAAT TTTCAACACT 300
 GACATAACTG TGTGTTTTCG ATAAAACAAA ATCATCTTCA GGCATAATT TAACGTCATA 360
 AAGATTTTCC CTACACTCTA TATCATAACC TATCTTTATG TTTTCAGGTT GAATTTTATT 420
 40 ATTAAGATTA AAATATGTAT AAAAAAATGG ACAGTTAAGG TATCAAATGA ATACCATCAA 480
 CTGTCCGACT ATTCTTCATC AAAAAACCTG ATAAAACAAA TTGCCTTATC AGATTAGTAT 540
 45 CATTTGTATA AGCATATTAA TGACCAAATG TTGCTTTAAT CAGTGATGTC GTTTCTCCAC 600
 CTGGATATAA TACATATAAT AATAAGTAA CTGCTACACC TGTAAATGCA GTACAGAACC 660
 ATATAACTGA AGCGAATGGC CCGAATTTGC GGTGTACATT ATATTTATCT TTAAATGCAG 720
 50 TAATAATTTG AACTAGGCCT AGAATACCAC CAATTGTTGC TAAATTAATG TGGAAAAATA 780
 AGAAAAATCGT ATAATATTTT TTAATTGATG CTGGTCCGCC AAATGCTGTA TTACCGATAA 840

55

TAACATTTTT GTGTTTATTT ATTTCCCTTT TCCAAATAAG TCTCCAACCA ATGGCAATTA 960
 AAATTGCACT AATGACAATA CATGTCGTAC TAATCGTTGG TAAAATTGGA ACGCCCATAT 1020
 5 TTTTCATCCT AACTTAATTA ATCTAGATCA AAGTAAGTAA TGAAACAATC ACAGCTAACA 1080
 CGAAAAAGAT CACTAAATAA TTTAGTGAAT ATATAAACAT TTGTGTTGCC CATTTTGTTC 1140
 GATCTGAATT TTTCTTAAAT GTTGTTAAAC CTAATGCAAT CCATCCTAAA TTTAATAAGG 1200
 10 TrGCTAACAC TaCGAATACG ACACCTAAAT TTATTAGTAA TAAAGGTACT GGCAATAAAA 1260
 TAATCAACCA GATAAACATA CTGACACGTG TACGTTTAAA GCCCTTAACT GATGGTAACA 1320
 TTGGAATATT TGCAAGTGCA TATTCATCTT TACGTTTAAAT AGCTAAGGcA TAAAAATGAA 1380
 15 TTGGtTGCCA ACmAAATACA ACTAAAAACA GCGCAATCGC TGTAAACTA ATTTGTCCTT 1440
 CAATTGCAAC CCATCCAATT AGTGGTGGTA CTGCTCCAGG AAAACTCCCA ATCACTGTGT 1500
 20 TCCATGTTGT ATGTCTTTTA GACCATATTG AGTAATAAGA CACATAACCT ACAATCCCCA 1560
 TAAGACCAAG TACGCCTGAT GGTATATTCA ATAAAAACAA ACAAATTTCT CCAACTAACA 1620
 TCATACCAA ACTTAATAGT AATAAATTTT GATCTGTAAT TCTATTATTT ACAGTTGGTC 1680
 25 TATTTTGTTC ACTAGGCATA ATACGATCAA TATCTGGTC GTAATAATTA TTTAACGCAC 1740
 ATGCGCCACC CATAATTAAA GTAGATCCAA ATAGCATTAA TAAAATTGA GGTATTGATG 1800
 ATAAGAAGGA ATGATTTGTC ATTACAACTG CTAGCCATGC GCcCGCAAAA GCCGGAATTA 1860
 30 AGTTACCTTG AACAAGTCCC ATTTTAATTA TCTGTGCAA TTCTTTGAAG TTAACCTGTC 1920
 TAATATTTTG TGACmAGTA TGCTCTTGC TCATAATCCC CCTCCTTAAA TTTGTTTATA 1980
 TAAGATTATG ATATCTTAGA TTGCATAAAA AGACTAGGTT TAATAAAATT AAATTGTGAC 2040
 35 AAATTAACGA CAAGAGAAAA TGTCAATTTT GTGACACAAA TAACATTTAA TTTATTGCTA 2100
 TAATGTATAT GTTAGAAAAT TTTAATAAGT AGAATCATGC ATCTAAAAGA GATTAATATT 2160
 40 TAAGCTTCAA ATTTGAGTAA ACGTGGATTA CATAATTATC CCAATAAAAA AATCATTACG 2220
 ATTAAGTTCT TTTTATGTCG TCCACATACA ATACTTGTA AATTAAATCA TATTCCTGTC 2280
 GTTGATCCC ATCTTTTCAT ATCCTACAAT CAGGTCTATT TATAGTATCA TCTCAAATCC 2340
 45 GGCTATTAAT TCTAATTCCT AGTGATGCGT TTTTATATGA TGGGGTGAT AAATTGTTTG 2400
 GCAAAAAGAA TTTAAATGG TTAGGTGTCG TAGCAACGTT AATGATGACA TTTGTACAAC 2460
 TTGGTGGAGC CTTAGTTACC AAAACCGGAT CAGCTGATGG TTGTGGTTCT TCTGGCCAC 2520
 50 TATGTCATGG TCGTTGATT CCAGAATTCT TTCCTATTGA TACGATTATT GAGTTAAGTC 2580
 ATAGAGCCGT TTCAGCTTTG TCTTTATTAA TGGTCTTATG GTTAGTTATC ACTGCATGGA 2640

55

TATTGCAAGC ATTAATCGGA GCTGCTGCTG TTATTTGGCA ACAAACGAT TACGTTTTAG 2760
 cATTGCACTT TGGTATATCA TTAATCAGTT TCTCATCTGT ATTTTAAATA ACATTGATTA 2820
 5 TTTTCTCTAT AGATCAAAAA TATGAAGCTG ACGAATTATA TATCAAAAAG CCATTAAGAC 2880
 GTTTAACATG GTTAATGGCA ATCATCATTT ATTGTGGTGT TTATACTGGT GCACTAGTGa 2940
 GACATGCGGA TGCAAGTTTA GCATATGGTG GTTGGCCATT GCCATTACCA CGATCTTGTA 3000
 10 CCACATTcAG GAACAAGATT GGGTTCAACT CACGCATCGT ATCAnGGTCn nTTAA 3055

(2) INFORMATION FOR SEQ ID NO: 298:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 748 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

TTCTTATTTA AAGAAGTCAT TTTTAGAAAT TGTTGAGACT TTAAAAAATG ATCCGTATAA 60
 25 AATAACACAA TCTTTTGAAA AATTAGAGCC TAAATATTTA GAGCGATATT CAAGAAGAAT 120
 TAACCATCAG CACAGGGTCG TCTATACCGT AGATGATCGA AATAAAGAAG TATTAATACT 180
 ATCGGCATGG TCACATTATG ATTAATGAAT ATTCAATATC TGAATAACTT TAATGATAAG 240
 30 TTAATTAAAG AACTAGTAT TTAAGTGTAG GGAAATAGC GACGTTAATG CGTTGTTATC 300
 TCTACACTTT TTAATTTTAT AATAGCGCAA GACTAAACAG ATTGAAATTA GTAACAATAA 360
 AAGAATAACG TATTATAATA AGGAATTTTA AATTGTGACT TTTTCGGAAT ATTAAATTTT 420
 35 AGAAATATGA GGTTTTTAAG CGGATTCCTC ACAAATTTT AAAAATATTT AAGCCTGAAA 480
 ATGATAAAGC GGTAGGGAAC GTTTTCTGA AaGTTAGTGA TACAATAGTT TTAAGTTGAA 540
 40 ATACAGGAGG ATGAATAACA TGAATCAGTC AGTCAAATTA CTTAAACATT TAACAGATGT 600
 AAACGGCATT GCTGGTTATG AAATGCAAGT TAAAGAAGCA ATGCGTaaCT ATATAGAGCC 660
 TGTCAGTGAT CaaATTATTG AAGATAACTT GGGTGGCATT TTTGGAAAGA AAAATGCTGA 720
 45 GAATGGTCAA TACTCAATTA TGGAnTTC 748

(2) INFORMATION FOR SEQ ID NO: 299:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4718 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

	ACCTCCGAAT AATTGATTCC ATTAACTTTT TTTGTGAAAA AATTAAACAG GCGAAGTCTT	60
5	CAATAAGTGA AGATGTCCAA ATTATAAAAA ATACATTCCA AAAAGAAAAG TTAGGTACAG	120
	TAATTACTAC ycTGGCGCAA GTGGTGGTGT TACGTATAAA CCAATGATGA GTAAAGAAGA	180
10	GGCGACTGAA GTTGTTAATG AGGTCATTAC TCTATTAGAA GAGAAAGAAC GTTTGTTAcC	240
	TGGCGGATAT TTATTTTTAT CAGATTTGGT AGGTAATCCA TCGCTACTAA ACAAAGTTGG	300
	TAAGTTAATT GCCAGTATTT ACATGGAAGA AAAATTAGAT GCTGTTGTTA CCATTGCGAC	360
15	AAAAGGTATT TCATTGGCAA ATGCGGTTGC TAATATTTTA AATTTACCAG TAGTAGTGAT	420
	TAGAAAAGAC AACCAAGGTGA CTGAAGGTTT TACAGTTTCA ATTAATTACG TTTCAGGATC	480
	TTCAAGAAAA ATAGAAACAA TGGTACTTTC GAAGAGAACT TTAGCAGAAA ATTCAAATGT	540
20	TTTAGTTGTC GATGATTTTA TGAGGGCTGG TGGCTCTATT AATGGTGTTA TGAATTTAAT	600
	GAATGAGTTT AAAGCCCATG TAAAAGGGGT ATCAGTACTT GTAGAATCAA AAGAAGTTAA	660
	ACAAAGATTG ATTGAAGATT ATACTTCCTT AGTGAAATTA TCTGATGTAG ATGAATATAA	720
25	TCaAGAGTTT AACGTAGAAC CTGGCAACAG TTTATCTAAG TTTTCATAAA AGGAGTTTTA	780
	GTATTATGAA AATCATTAAAC ACAACAAGAT TACCGGAAGC ACTTGGACCA TATTCGCATG	840
	CAACAGTTGT GAATGGTATG GTTTATACTT CTGGTCAGAT TCCATTGAAT ATTGATGGAC	900
30	ATATCGTAAG CGCTGATGTT CAAGCACAGA CAAAACAAGT TTTAGAAAAT TTAAAGGTTG	960
	TTTTGGAAGA AGCAGGATCT GATTTGAATT CTGTTGCGAA AGCGACCATT TTCATTAAAG	1020
	ATATGAATGA TTTCCAAAAA ATAAATGAAG TGTATGGTCA ATATTTTAAT GAACACAAGC	1080
35	CAGCGCGTAG TTGTGTAGAG GTTGCGCGTT TGCCAAAAGA TGTGAAAGTA GAAATTGAAT	1140
	TAGTAAGTAA AATTAAGGAA TTATAATTTT CGATTAATAT GTTTAATCAA GCTTCTAAAT	1200
40	AAAACAGAGA GATATATACT ATAGGGGGGC TCACTACATG AAAGTGACAG ATGTAAGACT	1260
	TAGAAAAATA CAAACAGATG GACGAATGAA AGCACTCGTT TCCATTACAT TAGATGAAGC	1320
	TTTCGTAATT CATGATTTAC GTGTAATTGA AGGAACTCT GGCTTGTTTCG TTGCAATGCC	1380
45	AAGTAAACGT ACACCAGATG GTGAATTCCG CGACATCGcg CATCCTATTA ATTCAGATAT	1440
	GAGACAAGAA ATTCAAGATG CAGTGATGAA AGTATATGAT GAAACAGATG AAGTAGTACC	1500
	AGATAAAAAC GCTACATCAG AAGATTCAGA AGAAGCTTAA TCAATTTTAT ATTTAGCGAT	1560
50	GTAATACATT TGCAATAAGT TGATTTGATA CTGTCGATAA AGCATAAAGC TTTGTGCGCA	1620
	GTTTTTTTATG TTTGTATTAA TGTTTTTTTA TTTTAAATGA AAGGCTAATA AATATATACG	1680

55

	TGaTGCTCGT	ATTTTGAAG	TAAGAAAAA	GTTGTTTTTA	AAATTACAAC	GAATTAAAA	1800
	CAATGCCTTT	TATATGTTGA	AAGAGTATTG	CaGATTAAAT	TaTAATAATG	ACGAaGgTAA	1860
5	AATTTAATGG	GGGTTAATGT	TCATGCGAAG	ACACGCGATA	ATTTTGGCAG	CAGGTAAAGG	1920
	CACAAGAATG	AAATCTAAAA	AGTATAAAGT	GCTACACGAG	GTTGCTGGGA	AACCTATGGT	1980
	CGAACATGTA	TTGGAAAGTG	TGAAAGGCTC	TGGTGTGCGAT	CAAGTTGTAA	CCATCGTAGG	2040
10	ACATGGTGCT	GAAAGTGTA	AAGGACATTT	AGGCGAGCGT	TCTTTATACA	GTTTTCAAGA	2100
	GGAACTACTC	GGTACTGCGC	ATGCaTGCAA	ATGGCGAAAT	CACACTTAGA	AGACAAGGAA	2160
15	GGTACGACAA	TCGTTGTATG	TGGTGACACA	CCGCTCATCA	CAAAGGAAAC	ATTAGTAACA	2220
	TTGATTGCGC	ATCACGAGGA	TGCTAATGCT	CAAGCAACTG	TATTATCTGC	ATCGATTCAA	2280
	CAACCATATG	GATACGGAAG	AATCGTTGGA	AATGCGTCAG	GTCGTTTAGA	ACGCATAGTT	2340
20	GAAGAGAAAG	ATGCAACGCA	AGCTGAAAAG	GATATTAATG	AAATTAGTTC	AGGTATTTTT	2400
	GCGTTTAATA	ATAAACGTT	GTTTGAAAAA	TTAACACAAG	TGAAAAATGA	TAATGCGCAA	2460
	GGTGAATATT	ACCTCCCTGa	TGTATTGTCG	TTAATTTTAA	ATGATGGCGG	CATCGTAGAA	2520
25	GTCTATCGTA	CCAATGATGT	TGAAGAAATC	ATGGGTGTAA	ATGATCGTGT	AATGCTTAGT	2580
	CAGGCTGAGA	AGGCGATGCA	ACGTCGTACG	AATCATTATC	ACATGCTAAA	TGGTGTGACA	2640
	ATCATCGATC	CTGACAGCAC	TTATATTGGT	CCAGACGTTA	CAATTGGTAG	TGATACAGTC	2700
30	ATTGAACCAG	GCGTACGAAT	TAATGGTCGT	ACAGAAATTG	GCGAAGATGT	TGTTATTGGT	2760
	CAGTACTCTG	AAATTAACAA	TAGTACGATT	GAAAATGGTG	CATGTATTCA	ACAGTCTGTT	2820
	GTTAATGATG	CTAGCGTAGG	AGCGAATACT	AAGGTCGGAC	CGTTTGCGCA	ATTGAGACCA	2880
35	GGCGCGCAAT	TAGGTGCAGA	TGTTAAGGTT	GGAAATTTTG	TAGAAATTAA	AAAAGCAGAT	2940
	CTTAAAGATG	GTGCAAGGT	TTACATTTA	AGTTATATTG	GCGATGCTGT	AATTGGCGAA	3000
40	CGTACTAATA	TTGGTTGCGG	AACGATTACA	GTAACTATG	ATGGTGAAAA	TAAATTTAAA	3060
	AcTATCGTCG	GCAAAGATTC	ATTTGTAGGT	TGCAATGTTA	ATTTAGTAGC	ACCTGTAACA	3120
	ATTGGTGATG	ATGTATTGGT	GGCAGCTGGT	TCCACAATCA	CAGATGACGT	ACCAAATGAC	3180
45	AGTTTAGCTG	TGGCAAGAGC	AAGACAAACA	ACAAAAGAAG	GATATAGGAA	ATAATCATTT	3240
	ACGTATTTAA	AATGGCTAGG	ATAAAAGGAT	AATCCTATGT	AATATTAATG	TAATCTTTAT	3300
	GATTTAATGA	TTGCGATAGT	AATGGAGTTA	CATyTTATAT	ATAATAGTAA	TTGCGTAAGT	3360
50	AAATAATTGG	AGGACTATAA	ATGTTAAATA	ATGAATATAA	GAATTCGTCA	TTAAAGATTT	3420
	TTTCATTGAA	AGGAAACGAA	GCATTAGCGC	AAGAAGTTGC	TGACCAAGTA	GGAATTGAAC	3480
55							

5 GTATTCGTGG TTGTGACGTA TTTATTATTC AACCAACATC ATATCCTGTG AATCTACATT 3600
 TAATGGAATT ATTAATTATG ATTGATGCTT GTAAACGTGC TTCTGCAGCA ACAATCAATA 3660
 10 TTGTAGTGCC ATATTATGGA TATGCAAGAC AAGATAGAAA AGCCCGTAGC CGTGAGCCAA 3720
 TCACTGCTAA ATTAGTTGCA AACTTAATCG AAACAGctGG CGCAACTCGT ATGATTGCGT 3780
 TAGACTTACA TGCACCACAA ATTCAAGGAT TCTTTGATAT TCCAATTGAC CACTTAATGG 3840
 GTGTGCCAAT TCTTGCTAAA CATTTCAAAG ATGATCCGAA TATTAACCCA GAAGAATGTG 3900
 TCGTTGTTcA CCAGACCATG GCGGsnTTAC ACGTGCACGT AAATTAGCTG ACATTTTAAA 3960
 15 AACTCCAATT GCAATTATAG ATAAACGTCTG TCCTAGACCA AATGTTGCTG AAGTGATGAA 4020
 CATTGTTGGT GAGATTGAAG GACGTACGGC AATTATTATT GACGATATTA TTGATACAGC 4080
 AGGTACAATC ACTTTAGCTG CACAAGCATT AAAAGATAAA GGTGCTAAAG AAGTATATGC 4140
 20 TTGTTGTACA CACCCGTGTT TATCAGGACC GGCTAAAGAA CGTATCGAAA ATTCTGCTAT 4200
 AAAAGAATTA ATCGTAACAA ACTCAATTCA TTTAGATGAA GATCGCAAAC CATCTAACAC 4260
 TAAAGAATTA TCTGTGCTG GTTTAATCGC ACAAGCTATC ATTCGTGTAT ACGAAAGAGA 4320
 25 ATCAGTTAGC GTATTATTTG ACTAATATTT AAAAGGCGTT TGACGAACAT ATTCCAAACG 4380
 TGTATAATAG TTTTCGTTCTG GATTATACGA ATAAATAAAC ACTTGCAAGC AACGATGATG 4440
 TTGATGGGTA AGTGAGGTGC TCGTTTTGAG CAAAAATGAA AGGTGGAAAT GAGAATGGCT 4500
 30 TCATTAAAGT CAATCATCCG TCAAGGTAAA CAAACACGTT CAGATCTTAA ACAATTAAGA 4560
 AAATCTGGTA AAGTACCAGC AGTAGTATAC GGTTACGGTA CTAAAAACGT GTCAGTTAAA 4620
 GTTGATGAAG TAGAATTCAT CAAAGTTATC CGTGAAGTAG GTCGTAACGG TGTTATCGAA 4680
 35 TTAGGCGTTG GTTCTAAAAAC TATCAAAGTT ATGGTTGC 4718

(2) INFORMATION FOR SEQ ID NO: 300:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3181 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

50 AAAATGCATA TAAATACATA TTAAGGAGGA TTTTATGAAT TTTCTTAAAC CTGCAAAGCA 60
 TATTAAGCCT TTGCCAGAAA ATCAGATAGA TGATACCTAT AAACGATTAC GTCTCCAAGT 120
 ATTTCTTGGT ATTTTCATCG GTTACGCTGG GTACTATTTA TTACGTAAAA ACTTTTCGTT 180

55

	TGCTGTTTCC ATCGCATATG GATTTAGTAA GTTCTTTATG GGTACTGTAA gTGATCGGAG	300
	CAATGCTCGG ATATTCTTAG TTCTTGGATT AGCACTCACT GCTATCGTCA ATTTGTTAAT	360
5	GGGATTTGTA CCGTTCTTTA CATCAGGTAT CGGTATTATG TTTGTCCTAT TATTCTTAAA	420
	TGGATGGTTT CAAGGTATGG GCTGGCCACC TTCAGGCCGT GTTCTCGTTC ACTGGTTTAG	480
	TGTAAGTGAA CGCGGAAGTA AGACTGCCCT TTGGAACGTT GCGCATAATG TTGGTGGAGG	540
10	TATTATGGCA CCTATTGCTG CTTGGGGTAT TACAACAACT GCATTTATCA ACTTTGGTTA	600
	TTTAAAAGGT TTCGAAGGTG TATTCAITTA CCCTGCACTC TTAGCACTTA TCATTGCCGC	660
15	AATTTCATAC GTATTGATTA GAGACACACC TCAATCTCAA GGTTTACCTC CAATCGAAAT	720
	TTATAAAAAT GACTTTGCTA CAAGCGATAA GAAAACATTA GAAACAGAAT TAACTACAAA	780
	AGAAATTTTA TTTAAATATG TACTGAACAA TAAATGGGTA TGGGCAATTG CCTTTGCAAA	840
20	TATATTTGTT TATTTCTGTC GTTATGGTGT ACTTGATTGG GCGCCAGTCT ACTTAAGTGA	900
	AGAAAAACAT TTCGACTTAA AAGCATCAGG TTGGGCATAC TTCTTATACG AATGGGCTGG	960
	AATTCCTGGT ACATTATTAT GTGGTTACAT TTCTGATAAA TTATTCAAAG GTCGTCGTGG	1020
25	ACCTGCAGGT TTCTTCTTTA TGTTAGGTGT CACAGTATTT GTATTAATTT ATTGGTTAAA	1080
	TCCTCCAGGC AATGCTTGGT TAGACAATGT CTCATTAATT GCCATTGGIT TCTTAATATA	1140
	TGGACCAGTT ATGTTAATTG GTTTACAAGC ATTAGATTAT GTACCTAAAA AAGCAGCTGG	1200
30	CACAgcAGCT GGATTAACAG GATTATTTGG TTATCTGTTT GGTGCTGTAA TGGCCAACAT	1260
	CGTCTTAGGT GCTGTAGTTG ATAAATTCGG ATGGGATGTC GGTTTTATTT TATTAAACAGC	1320
	AATTAGTGTG TTTGCAATGT TGAGCTTTAT CCTCACTTGG AATAAAGTAG GACAAGAAAC	1380
35	CGTTCATCAT TAAATGATAA AAAATAAAGT CATATGGTTA TCTTATCGAA AGATGATATA	1440
	TTCTCTCTT ATAAGTTCAA CCATATGACT TTTTATTAGT ATTCAAAAAA ATATTTACAT	1500
40	TGCCACTTTT GTGTTTGCCC TGCTGTTTTA TTCAATTGAT TACACCACTT AGGATAAACT	1560
	CTAAAAGCCA TTTTCCCTTG ATACTTAGAT GAAGCTAAGA TACCTTTTTT TACCAATAAT	1620
	TCCCTAGGAA ATAGAAAGTA GCCATTTAAT TCATCATCGA TAACAGCAAC AATCAGGTAA	1680
45	TCAGCAAACCT CTTCAATTTT GTATGGCCGA TTATAATTAT CtTCGTCTTT TGTCCAACAT	1740
	GTCACGAAAT ATCCCGATTT AGTCGGTGTT TTCTTAGCTA ACCTACTTTG ATATGTTTCC	1800
	TCTTTAAAGC TAAATGTTAA TGCyTCGTAA TCTTGATTAT ATTTTCTTC AGTTAAATCT	1860
50	TTAACTTCTG ATTCTTCACT AAAAATATTT TTCAACAGTA TTTTAGATTT ACACATATGC	1920
	CAATCTCACT TTATTTTTTC ATAATCGTAT CATATATTTA TTTTTTTCGA AAAATACACT	1980

TGTATCAACT GTCAAACGAT CACGTAATAA ATAGACGATT AACATTGcTG CTAAAGCACA 2100
 AAGTGACTCA GCAATTAATA ATGACCAAAT GACACCTGTT AGTCCAAACA AAGCATTTCAT 2160
 5 AATAAATAAT ACTGGAATGA TAATTGCACC TTGTAAAATG GCCATAATTG TAGCACCACG 2220
 ACCTTGCCCA GTCGCTTGAA GCATACCAGT AAACAAGAAA CCTATACCAT TTAATAATAA 2280
 TGATGCCATT GTTACTTTCA AAATAAATGt CGCCATCTCA ACAATGGCTT GATCAGTAGT 2340
 10 AAATAGTCCG ACCATATGAT GTCCAATTGT AAATACAGCA CTCATACATA CAACAAAGAT 2400
 AACGCCGATA GACATGATAA CTGCTTTGAT AACGTCTTTC ATACGGCCTT TATTTGCCAT 2460
 AAAGTTATAT GCAATTAGTG GTACAACACC TTCACATAAT CCCATGATAA TAAGTTCTGG 2520
 15 aAATTGCACA AGTCTAAATG AGATACCATA ACTTGCaATC GCGAAGTTTC CATAATGTGC 2580
 TAAAAATAAA TTAAAACTA ATCCTGTGAA TCCCATTAG ATACTCATT AAAATGCAGG 2640
 20 AATACCGATT TTAAAGATTT CAGAAAGCAT TTCTTTATTA GGTTCGCAA GTTTAATATT 2700
 AACTGACACA ACGTCACTAT TTTTCATAAA ATAAATGATA AAGAACAGAG CAGCAGCAAC 2760
 ATTACTGATT GCAGTACCCA AAGCTGcACC AACACGTTT aAATCAAAAC CAAAAATTAA 2820
 25 AATTGGATCT AAAATAATAT TTAAGCCTAC ACTAGCTAAC ATACCAATCA TAGAAACCAT 2880
 TGGTGCCCCA ATTGCACGTG CAAATTGTTC TAATATGAAG AACAAAATTA CAAAAGGTGC 2940
 ACTTAAAAAC ATTACTTTCA AATAATTACT TGTTAAAGCT AACGTTTCAC CTCTCGCCCC 3000
 30 TAAAATTGCT GCGATTTGAT CACTGAATGG TAAAGTAACT AAAATCACGA TAAGTCCTAG 3060
 TGCAATACCA CCATAAATAG AGAACTACT TACAAATTTA CTCTTAmTAT AGTCTTTGCG 3120
 ACCTAATAAA CGTGAAATAT AAGTTCCTGC ACCAACGCCA AATAAATTAC CTAACCCCAT 3180
 35 T 3181

(2) INFORMATION FOR SEQ ID NO: 301:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4029 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:

TCATCACCTT CAAATAAATA GGCATGCGAT AATTTATTTG AATGATATGC ATTGTCAT 60
 50 TGTGCTGTT CATCCATTCA TTACAACTC CCTTTTGCTT TTATATAAAA AGGTTGCCAA 120
 AGAGCAACTA ATTACTGCAA CAATAGCGAC CATGATTGTG GATTATTTTT AGAACCACAA 180

55

	TTATTAATAC ACTATTTTAA GTATCTCTAA TTGATGATTG TATTATAGAA TTAAAATTGA	300
	TGGAATGCAT CAACTGGCAT AACAAATACA GTAGCACC GC CAACTTCAAC TTCAACTGGA	360
5	TATGGAATGT ACGAATCCGC ACTGCCTCCC ATAGGTGTAA TAGGTGAAAC CAACTGTTCT	420
	CTATTACCAC ACGTTTGATT AATCACAGAC AATATTTTCAT CTACACGGTC ATCATTGACA	480
	CCACATAAGA ATGTTGTATT ACCCGCTCTT AAAAACCAC CTGTTGTTGC CAATTTTGTT	540
10	GCTCTAAAGT TATTTTAAAC AAGTTGATCT GCAAGTTCCT GACTATCTTG ATCTTGTAGC	600
	ATCGCTATAA TCATTTTCAT TTTATAACAC CTCTTCTAAC AATTATATCA TATCTTTTCT	660
15	AAATATTGGA TGATAGTTTG ATACGTGTCT TCAACAACAT TTTCAAGAGG TTGATCTGCA	720
	TTAACGCTTT TGAACCGTTG TGATTCATTA TGAATGATTT CTTGGTAACC TTCAATTACT	780
	TTTTCGTGAA CTTTAAATCT TCTTGATCTA ATCTATTTTG ATCTCTTGAA TTTTAAATAA	840
20	TACGTTCCGC ACCTACTTCA GCACTAACAT TTAAATAAAT CGTCAAGTCT GGATATAATC	900
	CATTTATTGC AAATTCGTTT AATGCTCTTA CTTCTTCAAC GCCAATCCCT CTAGCATAAC	960
	CTTGATAAGC TAATGAAC TAATGATATAGC GATCACACAA CACAACCTTA CCTTCTTTTA	1020
25	AAGCTGGTAT GACCTTTAAT ACAAGATGTT CTCTTCTAGA TGCAGCAAAT AACATTGCTT	1080
	CaGtCTAAT GTCCATATCA TTGCCCTTCT AATACAATTT TACGTATTTT TTCACCAGTA	1140
	GGAACACCAC CTGTTCTCT AGTCATAATG ACATCATAAT CTTTTACTAA TCTATGGTAA	1200
30	ACTTCATTAA TTACAGTTGT TTTTCCAGAG CCTTCTGGGC CCTCAAAAGT TATAAAAGCT	1260
	GACATTTTAT TCATCCTCAA CTAAAATTTT ATTATTTTTA ATTCCTTCAA CTATCATTC	1320
	AGTTTCCAGA TATTCATTTA CCAATTCTAT CATATTTTCA GTTATTGTTT CTCCTTTGAA	1380
35	AATAATAGGA ATcCCTGGcG GATACGGGAC AATATGTCGC GCCAGAACTT TACCTTTgCC	1440
	TTTTCAAATC ACACCAAGTA ACATATTCAA AGCGTTTTGG TTTATAGTTA CCTTCAGTyG	1500
40	TTAAAAGTTG TGTTTGTTTA ACTTTAGAAA CTGATTTTTT CGGTAAAATC ATATCTTCAA	1560
	TTTTACGTAA AAGCGAATCA AATAAATACG TATCATCATG ATGCCATAAC GGCAATATCG	1620
	CTAATGCTTG ATAGTCGTCC GCTAATTCTA AATAGATATG TGCATTCATA AACCAATTTT	1680
45	GAATATCATG ACCTGTAAAA CCTTCATATT TTATCAGCAA CTTCAACGGA TCATCAACTT	1740
	GAAGCATTTT AAAACCCTTC TTCTCCAAAC ATTCGATTAA TTGCGCTCTC TTATCAAAAA	1800
	ACACGGTACT ATCATATGTT TTATAGAACT CGGCAGCTGA CTCTAAACTA GCCATAATCA	1860
50	AATACGAAGG ACTAGATGTT TGGAAGTAGC TTAGATATTC TATAATAGTT TCTCTATAAG	1920
	GTGCATTTTT ATGAATATAA AGTACCGAGC CCATCGTTAA AGCTGGTAAC GTTTTATGAA	1980

55

	CAAAGTGCGC GCCGTGTGCT TCGTCAATGA GTACAGGAAT ATTTAATTGG TGCAAAGATT	2100
	TGATAACCTC TTCTACATTA AATGTTTCAC CGTAATAGTT AGGATAAGTC AACACAGCAA	2160
5	GTTTGTGACC GTCATTATTC AAACGGCTTA AATTAACTTT ATTATAATGA TTCGTTAACG	2220
	GACTTTGATG CGTTTCAATA AAATGCCCTT CTTGTTGGCT AATATCGAGC GCATGTAACA	2280
	CAGATTTATG TACATTTCTT GCCATTAAGA TATCGCCTTT TTTCTGTGAA AAAGACTGGA	2340
10	TGACAGATAA TATTCCTGAA GTGGTGCCAT TCACTAAGAA ATAAGCATCA TAATCTGAAT	2400
	GTTTCTCCAC CTGCTTCATA CTTTCCAAAA TGACTTCTTC AGGATGATGT AAATCAnCTn	2460
	AATCCAGGTA TTTCAGTTTT ATCCATTGTC ATTGATAATT GAGATAAATG ACCGATAGTC	2520
15	ATATTTTTAT GACCCGGAAC ATGCAAAGAA ATCGCTTCTT CTTGATTTAA ACTTTCTAAT	2580
	TTATTTAAAA TAGGTTGCTT CATGATATAC GCTTCCTTTA TTTACACTGT TTTGGAATTA	2640
20	GTTACTTTCA AAAGTATTAA TTATATAGTA ACACCTCTTT GACAAAAGTT AGTGTTACTT	2700
	ATGCAATAGC TTGTCTATTG TATAATAATT AATTTCTTTT TTGTACTTCG ATTTAAAAGA	2760
	TATTAGACAT AAAATCTAAA AACAGCAGTA AGATGATTTA TGATTAAAAA CTATCTTACT	2820
25	GCTGTTCACT TTTTATAATA CTTCTGAATG TCTTCACTTA TACTTCTAGT CACAGATTTA	2880
	AATAATCAAA AGTGCACATT ATTAATAATAT CAATTTCACA CTCAATGCCG CTCATCGCAT	2940
	TCATTTCTTG TCTAGCAACG TTCTACTCTA GCGGAACGTA AGTTAGCTAC CATCCTCGCT	3000
30	AAGAACCTTT CTTGACTTGT GACAATCGCT TGCTTCTTTC CTCTCCTTCG GCTCTCGCTT	3060
	ACTCATTTAG CTCTACTAAA CTCGTTGCGC TCTTTTCTCG TTTTCGTCAGA TTCAAACGTT	3120
	TTCACTTCGC CAAGCCATTT TTCTTTGTGT TTACTTTTTA TTTTGACGTT TTAGACATAA	3180
35	AAAAAGAGAC cTCACGGTCT CAACTTGCCT GGCAACGTTT TACTCTAGCG GAACGTAAGT	3240
	TGGCTACCAT CGTCGCTAAA GACCTTTCTT GACTTGTGAC AATCGCTTGC TTCTTTCTTC	3300
	TCCTTCGGCT CTCGCTTACT CATTTAGCTC TACTAACTC GTTGCCTCT TTTCTCGTTT	3360
40	CGTCAGATTC AAACGTTTTT ACTTCGCCAA GCCATTTTTC TTTGTGTTTA CTTTTTATTT	3420
	TGACGTTTTA GrCATAAAAA AAAGAGACcT TGCGGTCTCA ATGCGGCTCA TCGCATCCAT	3480
45	TTTTTGCTG GCAACGTTCT ACTCTAGCGG AACGTAAGTT GGCTACCATC GwCGCTAAgA	3540
	acCTTTCTTG ACTTGTGACA ATCGCTTGCT TCTTTCTCT yCTTCGGCTC TCGCTTACTC	3600
	ATTTAGCTCT ACTAAACTCG TTGCGCTCTT TTCTCGTTTC GTCAGATTCA AACGTTTTCA	3660
50	CTTCGCCAAG CCATTTTTCT TTGTGTTTGC TTTTATTTT GACGTTTTAG ACATAAAAAA	3720
	AAGAGACCTT GCGGTCTCAA TGCGGCTCAT CGCATCCATT TTTTGCCTGG CAACGTTCTA	3780

55

CATGGGAACA GGTGTGACCT CCTTGCTATA GTCACCAGAC ATATGAATGT AATTTATACA 3900
 TTCAAAACTA GATAGTAAGT gAAAAGTGGG TTTTGCTTCG CAAAACATTT ATTTTGGATT 3960
 5 AAGTCTTCGA TCGATTAGTG ATTCGTGCAG CTCCACATGT GCACCATGCT TGCCACCTCG 4020
 GAACCTATT 4029

(2)/ INFORMATION FOR SEQ ID NO: 302:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7159 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:

20 GATGCAATAG TTGAGTAGTT ATAAGTAGCT ACATACAACC GCTCAAATAT AGGTTCAAGA 60
 ACATTTTTTAA TGAAGAAAGC GATTTTCCCT CAGAGTGCAA ACGCTTGA CT CAATAACGAA 120
 ATGTCACCTA TGTATGGCGT GACTTATTAT ACAGGAGGTG CAAAGTATGT TTGCTATTAT 180
 25 TGAAACAGGT GGAAACAAA TCAAAGTAGA AGAAGGTCAA GAAATCTTCG TTGAAAAATT 240
 AGACGTAAAC GAAGGAGATA CTTTTACATT TGATAAAGTA TTATTTGTAG GTGGAGATTCT 300
 AGTTAAAGTT GGAGCGCCAA CAGTTGAAGG TGCAACAGTT ACTGCTACTG TTAATAAACA 360
 30 AGGTCGCGGT AAAAAAATCA CTGTATTCAC ATACAAACGT CGTAAAAATT CAAAACGTAA 420
 AAAAGGCCAT CGTCAACCAT AACTAAATT AACAATCGAT AAAATCAACG CGTAATTATT 480
 ATGATTACTG TTGATATTAC AGTTAATGAT GAAGGCAAAG TAACAGACGT TATTATGGAT 540
 35 GGCCATGCTG aCCATGGTGa ATATGGTCAT GATATCGTTT GTGCTGGAGC TTCaGCTGTA 600
 TTGTTTGGTA GTGTTAATGC GATTATAGGA TTGACATCTG AGAGACCAGA TATCAATTAT 660
 GACGACAATG GTGGTCATTT TCATATAAGA AGCGTTGATA CAAACAACGA TGAAGCGCAA 720
 40 CTAATCTTTC AAACAATGCT TGTGTCTTTA CAACTATTG AAGAAGAATA TAATGAGAAT 780
 ATTAGATTAA ATTATAAGTG AGGTGCATTC CGATGTTAAA ATTAACTTA CAATCTTTCG 840
 45 CATCTAAAAA AGGGGTAAGT TCTACAAAAA ACGGACGTGA CTCTGAATCA AAACGCTTAG 900
 GTGCTAAACG TGCTGACGGT CAATTCGTAA CAGGTGGTTC AATTTTATAT CGCCAACGTG 960
 GTACTAAAAT TTACCCTGGT GAAAATGTAG GTCGTGGTGG CGATGATACA TTATTCGCTA 1020
 50 AAATCGACGG CGTTGTAAAA TTCGAACGTA AAGGTCGCGA CAAAAACAA GTTTCTGTAT 1080
 ATGCAGTAGC TGAATAATTT TGTCTAGTTA ACACCAGAAG TGAATCTTCT GGTGTTTCTT 1140

	AGACGTTATA CTAAATGTGC AcTGTATAAG AGCCCCTAAT CACTAAACTA TAAGGGGGAC	1260
	AAAGGAATAC AGTTGCAGCG TTTAAAGAAT AACTGTACC ACAATTGGTG CTGAGAAATA	1320
5	TAAGTATTTT AAAGCAAAGA TTTATAAAAG TAACTGCATA AGAGCCCCTA ATTATTTACA	1380
	ATATATAAGG GGCTCTAATA TGCTATAATT ATTGGGAAAA TGAAAATTAT ATGTAAAAGA	1440
10	GGTGAGATAT ATGTTTGTCTG ATCAAGTCAA AATATCTCTT AAAGCCGGTG ATGGTGGTAA	1500
	TGGTATTACC GCATACAGAA GAGAAAAATA TGTACCATTI GGTGGACCAG CTGGCGGTGA	1560
	CGGTGGTAAA GGTGCTTCAG TCGTATTTGA AGTGGATGAA GGTTTAAGAA CGTTATTAGA	1620
15	TTTTAGATAT CAACGTCATT TTAAAGCAAG CAAAGGTGAA AATGGCCAAA GTAGTAATAT	1680
	GCATGGTAAA AATGCGGAAG ATTTAGTATT AAAAGTTCCA CCTGGTACAA TTATTAAAAA	1740
	TGTTGAAACA GACGAAGTGT TAGCAGATCT TGTGAAGAT GGTCAAAGAG CTGTAGTAGC	1800
20	GAAGGGCGGT CGAGGTGGCC GAGGTAATTC ACGTTTTGCA ACACCTAGAA ACCCTGCACC	1860
	TGACTTCAGT GAAAAAGGTG AACCAGGTGA GGAATTAGAT GTATCTTTAG AATTGAAATT	1920
	ATTAGCTGAT GTAGGATTAG TAGGTTTCCC TAGTGTGGGT AAATCGACTT TATTATCTAT	1980
25	CGTTTCAAAA GCTAAGCCTA AAATTGGGGC ATATCATTTT ACAACGATTA AACCAAATCT	2040
	AGGTGTTGTT TCAACGCCTG ATCAACGTAG TTTTGTATG GCAGATTTAC CAGGTTTAAT	2100
	TGAAGGTGCA TCTGATGGCG TTGGATTAGG ACATCAATTT TTAAGACATG TAGAGAGAAC	2160
30	AAAAGTTATT GTTCACATGA TTGATATGAG CGGTTCTGAA GGTAGAGAAC CTATTGAAGA	2220
	TTATAAAGTC ATTAATCAAG AATTAGCTGC GTACGAGCAA CGTTTAGAAG ATAGACCTCA	2280
	AATCGTAGTA GCTAACAAGA TGGATTTACC TGAATCACAA GATAATTTAA ACTTGTTTAA	2340
35	AGAAGAAATT GGCGAAGATG TGCCAGTTAT TCCAGTTTCA ACAATAACGC GTGATAATAT	2400
	TGATCAATTA TTATATGCAA TAGCAGATAA ATTAGAAGAA TATAAAGATG TTGACTTCAC	2460
40	AGTTGAAGAA GAGGAGTCAG TTGGCATTAA CCGAGTATTA TATAAACATA CACCGTCACA	2520
	AGATAAATTT ACAATTTCAA GAGATGATGA TGGTGCTTAT GTGGTAAGTG GTAATGCTAT	2580
	TGAAAGAATG TTTAAAATGA CTGACTTTAA CAGTGATCCA GCAGTACGTC GATTTGCTCG	2640
45	TCAAATGCGT TCGATGGGTA TTGATGATGC GCTTAGAGAA CGTGGTTGTA AAAATGGTGA	2700
	TATCGTTAGA ATTCTTGGCG GAGAATTTGA ATTCGTTGAA TAGGAGCGAA ACATGATGGA	2760
	CAATAAAGAT TATAAAAAGT TTTATTTAAT TAGAGAAGAT GTCTTGCCTG AATCCGTGGT	2820
50	TAAAACATTG AAGATTAAAG ATGCCTTAAA AAGTGATCCG ACATTGTCCA TTTATGATGC	2880
	CGTTAAACAG TTTGATCTAT CTAGAAGTGC TTTTATAAA TATAGAGAAA CGATATTTCC	2940

55

	TGTTGGTATG TTGGCACGTG TACTAGATGT TATATCAAAG TTAGAACTAT CTGTATTAAAC	3060
	GATTCATCAA AGTATTCCAA TGGAAGAAAA AGCAACAATA ACATTATCAC TGAATGCTAA	3120
5	ATCTAAAGAA ACTTCAGTAG AAGATGTTAT TGGCGCTTTG AGAAATTTAG ATTATGTATC	3180
	AAAAGTAGAA TTAATTAGTA TGAGTATGTA AGGATGTGCC TATAATGTAC GCGTATGTCA	3240
10	AAGGTAAGTT AACACATTTA TATCCTACAC ACGTAGTTGT TGAAACTGCT GGTGTTGGTT	3300
	ATGAAATTCA AACACCAAAT TCTTATCGTT TTCAAAGCA TCTAGATCAT GAAGTTTTAA	3360
	TTCATACATC TTTAATTGTT CGTGAAGATG CACAATTATT GTATGGATTT AGTAGTGAAG	3420
15	AAGAGAAAGA TATGTTCTTG AGTTTAAATTA AAGTTACTGG TATTGGTCCG AAATCAGCTT	3480
	TAGCTATTTT AGCGACAAGT ACGCCTAATG AAGTAAAACG TGCCATTGAA AATGAAAATG	3540
	ATACGTATTT AACTAAATTC CCAGGAATTG GTAAGAAAAC GGCAAGACAG ATTGTCTTAG	3600
20	ATTTAAAGG TAAAGTGAAG ATTACTGAAG AAGATAGCGA TTCATTATTA CAAGTAGACG	3660
	CTACTTCGAC GGTGCAAGAT CAATTCGTGC AAGAAGCAAT GTTAGCGTTA GAAGCATTAG	3720
	GTTATTCTAA ACGAGAGCTT GCAAAAGTTG AGAAAACGTT AAATAAAAAT AAATATGACT	3780
25	CAGTTGATGA AGCTGTTAAG GCAGGTCTTC AATTAGTTGT ATCTTAATTT TAAATAGATT	3840
	AATAGGGGAA GTGTTGTCAT GAATGAGCGT ATGTTGATC AATCAATGCA TAGTGAAGAA	3900
	ACTGATTTTCG AATTGTCGCT TAGACCTACG AGATTACGAC AATATATTGG TCAAAATTCA	3960
30	ATAAAAAGTA ATTTAGAAGT ATTTATTAAA GCGGCTAAAC TTCGTCATGA ACCATTAGAT	4020
	CATGTATTGC TTTTGGCCCC CCCTGGATTA GGTAAGACAA CATTATCTAA TATCATTGCC	4080
	AATGAAATGG AAGTTAATAT ACGTACAGTA TCAGGGCCTT CATTAGAAAG ACCTGGTGAT	4140
35	TTGGCTGCAA TTTTATCAGG ACTTCAACCT GGAGATGTTT TGTTTATTGA TGAAATACAC	4200
	AGACTGAGTA GTGTTGTTGA AGAAGTGTTA TACCCTGCAA TGGAAGATTT CTTTTTAGAT	4260
40	ATTATCATTG GTAAAGGCGA TGAGGCTAGA AGTATCCGTA TCGACTTACC TCCATTCACT	4320
	TTGGTAGGTG CAACAACGCG AGCTGGCAGC TTAACAGGTC CACTAAGGGA TCGATTTGGT	4380
	GTGCACTTAA GATTAGAATA TTATAACGAA TCAGATTTAA AAGAAATCAT TATTAGAACA	4440
45	GCTGAGGTTT TAGGCACAGG TATTGATGAA GAAAGTGCCA TTGAACTTGC TAAACGTTCT	4500
	AGAGGGACTC CAAGAGTAGC AAATCGACTA TTGAAGCGGG TAAGAGACTT CCAGCAAGTG	4560
	AATGAAGATG AACAAATATA CATTGAAACA ACGAAGCACG CATTAGGTTT ACTTCAAGTT	4620
50	GATCAACACG GACTAGATTA CATTGATCAT AAAATGATGA ACTGTATTAT TAAGCAGTAT	4680
	AATGGCGGAC CTGTTGGTTT AGATACGATT GCCGTAACAA TTGGTGAAGA ACGTATTACA	4740

55

	GGCAGAAAAG CAACACCATT AGCTTATGAA CATTTTGCAA AGTCGAATGA GGAGAGAGAA	4860
	TAACGTGAAT ATTGAAGAAT TTGACTATGA CTTACCAGAA TCATTAAATTG CTCAAACGCC	4920
5	TTTAAAAGAT CGTGATCATA GTCGTTTATT AGTCATGGAT AGAGAAACTG GTGAAATGAA	4980
	ACATTTACAT TTCAAAGATA TCATTGAGTA TTTTAGACCT GGTGATACAT TAGTGCTTAA	5040
10	CGATACGCCA GTAATGCCAG CTAGACTTTT TGGTTTAAAA GAAGAAACTG GTGCAAAAGT	5100
	TGAAATGTTA ATGTLAACTC AAATTGAAGG TAATGATTGG GAAGTCTTAC TGAAACCAGC	5160
	TAAGCGTATT AAAGTTGGTA ATAAATTGAA TTTTGGTAAT GGCAAAATTA TAGCTGAATG	5220
15	CATAAAAGAA ATGGATCAAG GTGGACGCAT CATGCGTTTA CATTATGAAG GTATTTTACA	5280
	AGAAAGATTA GATGAATTAG GGGAAATGCC ACTGCCACCA TACATCAAAG AACGTTTAGA	5340
	TGATCCAGAT CGTTATCAAA CAGTTTACGC TAAAGAAAGT GGTTCCAGCG CAGCACCAAC	5400
20	AGCAGGATTA CATTTTACTG ATGAGTTATT AATTGAAATT AAAAATAAAG GTGTTAATAT	5460
	CGCATTGTGTT ACATTACATG TTGGGTTAGG TACGTTTAGA CCGGTGAGCG TAGACGATGT	5520
	GAATGACCAC GAAATGCATA GTGAATATTA TCAAATGACm CAAGAAACAG CTGATTTATT	5580
25	AAATGATACT AAGcCAAAGG ACATCGCATT ATATCAGTTG GTACAAC TTC AACACGTACA	5640
	CTTGAAACAA TTCGACGCCA TCATGATAAA TTTGTTGAAA CGAGTGGCTG GACTAATATA	5700
	TTTATTTATC CAGGATTGTA TTTTAAAGCA ATTGATGGCC AGATTACTAA TTTTCATTTA	5760
30	CCAAATCAA CATTAGTTAT GCTAGTATCA GCGTTTAGTA GTCGTGAAAA TGTTCTGAAT	5820
	GCTTATAAAA CGGCAGTAAA TTTAGAATAT AGATTCTTTA GTTTTGCGCA TGCAATGTTA	5880
	ATTATATAAA AAGAATGTGA GGATTTTGAA TATGCCTGCA GTAACATACG AACACATTAA	5940
35	AACTTGTAAT CAATCAGGTG CCGGTTTAGG TATCGTGCAC ACACCACAG GTTCATTGTA	6000
	AACATCTATG TTTATGCCAG TTGGTACTAA AGCAACCGTT AAAACAATGA GTCCAGAAGA	6060
40	GTTAAGACAA ATTGAAGCAA AAATCATTTT GGGCAACACA TATCATTTGT GGTTACAACC	6120
	CGGAAATGAT ATTATCAAAC ACGCTGGGG ATTACATAAA TTCATGAATT GGGATGGTCC	6180
	GATTCTTACA GATTCAGGCG GTTCCAAGT GTTTAGTTTA AGTAATTTAC GTAAAATTAC	6240
45	AGAAGAAGGC GTGGAATTTA GACATCATAC TAATGGGTCT AAATTATTTT TGAGTCCTGA	6300
	GAAATCAATG CAAATTCAAA ATGATTTAGG ATCTGATAIT ATGATGGCAT TTGATGAATG	6360
	TCCACCGATG CCTGCTGAAT ATGATTATGT AAAAAATCT ATTGAACGTA CAACACGTTG	6420
50	GGCGAAAAGA TGTCTAGATG CACACCAAAG ACCTGAAGAT CAAGCATTGT TCGGCATTAT	6480
	ACAAGGTGGC GAATATGAAG ATTTAAGAGA ACAAAGTGCA AAGGATTTAG TAGAATTAGA	6540

55

AATGGTTGAA CATACAGAGC AGTTTATGCC TAAAGATAAA CCAAGATATT TAATGGGTGT 6660
 AGGATCTCCa GATGCGTTAA TCGAATGTAG TATTCGCGGC ATGGATATGT TTGATTGTGT 6720
 5 CTTACCGACA CGTATTGCCA GAAATGGTAC TTGTATGACA TCGCAAGGTC GTTTAGTTAT 6780
 TAAAAATGCA AAATTTGCAG ATGATTTAAG ACCGTTAGAT GAGAATTGTG ACTGTTATAC 6840
 ATGTCAAAAC TATTCAAGAG CGTATATACG TCATTTAATC AAGGCAGAGG AAACTTTGG 6900
 10 TATTCGTCTT ACTACTATTC ATAATTTACA TTTTCGCTA AAATTAATGG AAGATATAAG 6960
 ACAAGCCATT CGAGAAGATC GTCTTTTAGA TTTCAAAGAA GAATTCCTCG AGCAATATGG 7020
 15 ATTAAATGTT GAGAACCCAA AAAACTTTTA AGCAAGAGGA GCGTATAAAA TGCAATTTTC 7080
 ATTACTAATA TATATAGTCG TAATTTTTGC GGTATGTAT TTCTTGATGA TCAGACCACA 7140
 ACAAAAAC TGCGAAACA 7159

20 (2) INFORMATION FOR SEQ ID NO: 303:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3159 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:

TCCATTTATA GAAATTTCTA AAGCAGAnAA TAAGATAGAA GATATCGGCC AAGGTGCAGA 60
 AATCATCAAA AGAACACAAG ACATTACTAG CAAACGATTA GCTATAACTC AAAACATTCA 120
 35 ATTTGATTTT GTAAAAGATA AAAAATATAA CAAAGATGCC CTAGTTGTTA AGATGCAAGG 180
 CTTCATTAGC TCTAGaCAA CATATTCAGA CTTAAAAAA TATCCATATA TTAAAAGAAT 240
 GATATGGCCA TTTCAATATA ATATCAGTTT GAaAACGAAA GACTCTAATG TTGATTTAAT 300
 40 TAATTATCTT CCTAAAAATA rAATTGATTC AGCAGATGTT AGTCAGAAAT TAGGCTATAA 360
 TATCGGCGGA AACTTCCAAT CAGCGCCATC AATCGGAGGC AGTGGCTCAT TCAACTACTC 420
 TAAAACAATT AGTTATAATC AAAAAAATA TGTACTGAA GTAGAAAGTC AGAACTCTAA 480
 45 AGGTGTTAAA TGGGGAGTGA AAGCAAATTC ATTGTTTACA CCGAATGGTC AAGTATCTGC 540
 ATATGATCAA TACTTATTTG CACAAGACCC AACTGGTCCA GCAGCAGGAG ACTATTTTCGT 600
 50 CCCAGATAAT CAACTACCTC CTTTAATTCA AAGTGGCTTT AATCCATCAT TTATTACAAC 660
 ATTGTCACAC GAAAGAGGTA AAGGTGATAA AAGCGAGTTT GAAATcACTT ACGGCAGAAA 720
 CATGGATGCT ACATATGCTT ACGTGACAAG ACATCGTTTA GCCGTTGATA GAAAACATGA 780

55

	AAAAATTAAA AGCATCACAC CTAAGTAAAC AGTTCaATCA TCTTAAAAAA TCCTGGGACA	900
	CTTCATACTT GTCTCAGGAT TTTTmACAA ATTGAATCAG CCTCATAACA TTAAATTATT	960
5	TTATCGTACA TTAAATTTAA TAATAACAAC TGATTTTAT AAGAATAAAG TATCGAmCCA	1020
	TAGTAGATAC ACAAATAATA CAAATGAAAC AATTTAACCTT GAAAGCTTaa ATAAATATTA	1080
	TCAAGTTAAT AAACAATTAA TTTTtagATG GATTcATCAa AAATCGTAAA AAAGCACAAT	1140
10	TTGTATTTTA CAAACATTAA TTAAAAAGA AAGCAAGACA TTCGTGCAAT CGGTTACCTT	1200
	AAATTGTTTA CAACTGTCAA CAATACCAAG GTTTTATTAA CTATATTTCT CACAAAATTA	1260
15	GCTTTTAGCA TTCCAAACAA AAAAGGTAA ATCGAACGGA ATTATGGCAT TTTTAACTTA	1320
	ATTGTAAAAA AAGTTGATAA TGGTCAATTG TTAATGAACA GTTAATTATA ATAACGCCCA	1380
	AAATATATTA TTATTTAATT AAGTTAAATA AAATTATAGA AAGAAAGTGA AACTTATGCT	1440
20	TAAAAATAAA ATATTAAC TAACCTTATC TGTGAGCTTA CTTGCCCTC TTGCCAATCC	1500
	GTTATTAGAA AATGctAAAG CTGCTAACGA TACTGAAGAC ATCGGTAAAG GAAGCGATAT	1560
	AGAAATTATC AAAAGGACAG AAGATAAAC AAGTAATAAA TGGGGCGTGA CTCAAAATAT	1620
25	TCAATTTGAT TTTGTAAAGG ATAAAAATA TAACAAAGAT GCTTTGATAT TAAAGATGCA	1680
	AGGATTCATT AGCTCTAGAA CAACATATTA CAACTATAAA AAACTAATC ATGTTAAAGC	1740
	TATGCGATGG CCATTCCAAT aTaATATTGG TTTAAAAACA AATGATAAAT ATGTTTCTTT	1800
30	AATTAATTAT TTACCTAAAA ATAAATTGA ATCTACAAAC GTGAGTCAGA CATTAGGATA	1860
	CAATATCGGT GGTAATTTCC AATCAGCCCC ATCACTCGGT GGTAATGGAT CATTTAAC TA	1920
35	TTCTAAATCG ATTAGCTATA CACAACAAAA TTATGTAAGT GAAGTAGAAC AACAAAATC	1980
	AAAAAGTGTT TTATGGGGCG TCAAAGCGAA TTCATTGCGC ACTGAATCAG GTCAAAAATC	2040
	AGCCTTTGAT AGCGATTTAT TTGTAGGCTA CAAACCTCAT AGTAAAGATC CTAGAGATTA	2100
40	TTTCGTTCCA GACAGTGAGT TACCACCTCT TGTACAAAGT GGATTTAACC CTTCAATTAT	2160
	CGCCACAGTA TCTCATGAAA AAGGTTCAAG CGATACAAGC GAATTTGAAA TTACTTACGG	2220
	AAGAAACATG GATGTCACTC ATGCCATTAA AAGATCAACG CATTATGGCA ACAGTTATTT	2280
45	AGACGGACAT AGAGTCCATA ATGCATTTGT AAATAGAAAC TATACTGTGA AATACGAGGT	2340
	CAATTGGAAG ACTCATGAAA TCAAGGTGAA AGGACAGAAT TGATATGAAA ATGAATAAAT	2400
	TAGTCAAATC ATCCGTTGCT ACATCTATGG CATTATTATT ACTTTCTGGT ACTGCTAATG	2460
50	CTGAAGGTAA AATAACACCA GTCAGCGTAA AAAAGTCTGA TGACAAAGTT ACTTTATACA	2520
	AAACAACAGC CACAGCAGAT TCTGATAAAT TTAAAATTC ACAGATTTTA ACATTTAATT	2580
55		

ACTCAGGCTT TGTGnAACCT AATCCTAATG ACTATGACTT TTCAAAATTA TATTGGGGAG 2700
 CTAAATACAA TGTATCTATA AGCTCACAAT CTAATGATTC AGTAAACGTC GTTGATTATG 2760
 5 CACCAAAAAA TCAAAATGAA GAGTTTCAAG TTCAAAATAC TTTAGGCTAT ACATTTGGTG 2820
 GTGACATTAG TATCTCTAAT GGTtTATCTG GTGGACTTAA TGGAAATACA GCTTTTTCTG 2880
 AAACAATTAA TTATAAACAA GAAAGTTACA GAACAACATT AAGTCGCAAC ACAAATTATA 2940
 10 AAAATGTTGG CTGGGGAGTT GAAGCACATA AAATTATGAA TAATGGTTGG GGACCTTATG 3000
 GAAGAGATAG CTTCCACCCA ACATATGGTA ATGAACTCTT CTTAGCTGGC AGACAAAGCA 3060
 15 GTGCATACGC TGGCCAAAAC TTCATAGCGC AACACCAAAT GCCATTATTA TCTAGAAGTA 3120
 ACTTCAATCC AGAATTTTTA AGCGTACTAT CACACAGAC 3159

(2) INFORMATION FOR SEQ ID NO: 304:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3821 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:

30 GCAAAACTT TTCTCCAAT ATGTCAGACT ATAATGGCAT GAACAAATT GATATGATGA 60
 ATCAAAATTA AGTTGATACG ATGTTGCATG GTTATCACGC TGGATTTTTA TTTCATTAT 120
 TGATTACAGT TGTTAGTTTC TTTTGTTTCAT TTATGTTACA AGGTAAGAAA AAAGAAGTTG 180
 35 ATTCTCGTCA ATAAATATAA GTTGCTAGCT ATATAAAGCT TTTTAGCAA AGTTCAACAT 240
 ATTGACTTAT CCGGCATTTC AGATTAAAAT ATTTTATTC CGATTAGAAT AATAAGAATA 300
 AGGAGATATA TTCTATGAAA AGACTTTTAT TTGTGATGAT AGCTTTCGTT TTCATATTGG 360
 40 CTGCATGCGG AAACAATTCG TCGAAAGACA AGGAAGCTAG TAAAGATAGC AAGACAATTA 420
 ATGTTGGGAC TGAGGGGACT TATGCACCAT TTAGTTTCCA CGATAAAGAT GGTAAATTAA 480
 CTGGTTACGA TATTGATGTT ATTAAAGCAG TGGCTAAAGA AGAAGGTTTA AACTTAAAT 540
 45 TTAATGAAAC TTCTTGGGAT TCTATGTTTG CAGGTTTAGA CGCAGGGCGT TTTGATGTAA 600
 TCGCGAACCA AGTAGGTATT AATCCTGATA GAGAAAAGAA ATATAAATTT TCTAAGCCTT 660
 50 ACACATTCTC AAGTGCTGTT TTAGTTATTC GTGAAAATGA AAAAGATATT AAAGATTTTG 720
 ATGATGTTAA AGGTAAGAAG TTAGCACAAA CATTACATC TAATTATGGT AAATTAGCTA 780
 AGGATAAAGG TGCTGATATT ACAAAGTTG ATGGCTTTAA CCAATCAATG GATTTATTAT 840

	AAAAACCTAA TGcTAAGATC AAAGCAATCA AAGGTAATGC TGAACAAAGT AGATCTGCAT	960
	TTGCATTTTC TAAAAAAGCA GATGATGAAA CAGTTCAAAA ATTCAATGAT GGCTTGAAAA	1020
5	AAATCGAGGA AAACGGTGAA TTAGCTAAAA TAGSTAAGAA ATGGTTTGGT CAAGATGTTT	1080
	CTAAATCTAA ATAGCGAACA ACAACATGCG CTAGATGCTG CAAAACAAGC TTTCGGACCT	1140
10	ATGCTAGAAG GTTTGGTCAA ATATTCAATT CCTATTACAT TAGTTACATT TGTTTTAGGA	1200
	TTGATTATTG CATTATTTAC AGCATTAAATG CGAATTTCAA CGAGTAAAAT TTTAAGAAGT	1260
	ATTTACAGTG TCTATGTATC TATTATTCTGA GGAACACCAA TGATAGTACA ACTATTTATC	1320
15	ATATTTTATG GTATTCCAGA ATTAGGTAGA TTATTAACAA ATGACGCTGA CAACCAATGG	1380
	ACATTGGCAC CTGTAAGTGGC TGCTATTATT GGTATTATCAT TAAATGTAGG TCGGTATGCT	1440
	TCGGAAATTA TTCGTGGCGG TATTATTTCT ATACCGAAAG GACAAACAGA AGCTGCaTAT	1500
20	TcCaTCGGTA TGACGTATGG TCAAACGATA CAACGTATCA TTTTACCGCA GGCAATTCTGA	1560
	GTGTCGATTC CTGCACTAGG TAATACATTT TTAAGTTTAA TCAAAGATAC ATCATTATTA	1620
	GGATTTATTT TAGTGGCTGA AATGTTTGA AAGCTCAAG AAGTTGCGTC TACAACATAT	1680
25	GAATATTTAA CAATTTATGT GTTAGTTGCG CTAATGTACT GGGTGGTATG CTTTATTATT	1740
	TCAATTATCC AAGGTATCTA TGAATCTTAT ATTGAAAGAG GGTATCGCTC ATGATTCAAT	1800
30	TGAACAATAT CCATAAATCA TTTAATGATG TTGAAGTCAT CAAAGGTATT GATTTATCTG	1860
	TTGAACAAGG TGAGGTTGTA ACCTTAATCG GTCGATCTGG TTCAGGTAAA ACAACATTGT	1920
	TACGTATGAT TAATGCATTA GAAATTCCAA CTGAAGGTAC AGTTTATGTT AACGGCAAAA	1980
35	CATATACATC TAAAGATAAA AAATCACAAA TAGAAGTTCTG TAAACAGTCT GGTATGGTAT	2040
	TTCAAAGTTA TAACCTTTTT CCGCATAAGA CGGCATTAGA AAATGTAATG GAAGGTCTTA	2100
	TCAÇAGTTAA AAAGTTGAAA AAGGATGAGG CACGTGGGAA ATCACTTGAG TTACTTGAGA	2160
40	AAGTTGGTTT AACACATGTC AAAGATCAAC GTCCACATGC ATTATCAGGT GGTCAACAAC	2220
	AACGTGTTGC TATTGTCAAG AGCACTAGCA ATGAACCCTA AAGTGATGTT GTTTGATGAA	2280
	CCAACATCTG CACTTGATCC TGAACCTGTG AATGATGTTT TAAAGGTTAT TAAAGATTG	2340
45	GCTAATGAAG GCATGACAAT GGTCAATGTG ACACATGAAA TCGGTTTTGC TAAAGAAGTA	2400
	TCTAATAACA TTGTATTTAT TcmTGAAGGC ATGATCGGAG AACAAAGGGC TCCAGAAGAG	2460
50	ATGTTCAATC GTCCGAAAAC AGAAGAATTA AGACGTTTCT TAAATGTTAT AAATGAAGAA	2520
	TAATCAAATA GAACCACGTA TCATGTTTTA GTATGGCGAT GAAGCCATAT ACATGATGCG	2580
55	TGGTTCTTTG TTATGTTGTC ATAATCTTGG AGCGATATTT TAACGACGTT TATGATTTAA	2640

TTCTACATGT GCGTTAAAAC CTTTTTTGAA TTGTTGGACG CCATAGTCTT CTGATGACTC 2760
 TGAAAAGTCA CCGGTAATAC CATAAAAATT ATAGCGATCA ATATGATGCG CTTTAGCAAA 2820
 5 CTTAATCATT TCCCacTGCA AATGGTAGGC ACCCATATAA GCATTATATT TAGGGTTTGA 2880
 ACCACTAGAT AAGTAATAAA CTTCATGCTC ATTGTAGATA AATAAAGCAG AAGCTAAGTT 2940
 10 TAAGACTGCA CCATCTTGTT CAATTTGTTT TATTGTATTG TCGATTTTAC GCTTATTGCT 3000
 ATTTAGCTGT TGTTCTAGCT GTGTGCGTTT CGTTTTATTT TTCTTTGAAT TAGGACTTTC 3060
 TTCCAATGCT TCTTCAACAC CTGAGAGTTC AGCTGTTAAT TGTGTGTGCT TTAATTGTAA 3120
 15 CGTTTTTAAA TACTCGTTTA AATCAATATA CGCCAACTTT AACATGGCGT GGTCATCGTA 3180
 TAACTTTTGC ATTTCTTCAA AGTATGGTAA CTCACGGAAT TTGAAACCGT GCTTTTCCTC 3240
 AGCCATATGG AATAAGTCGA AAAAAGTTTG CGTTTCATCA ATCGTTAAcG TTTTaGTTTT 3300
 20 GACACCAaTa TCaTATGTtT tTtTAATATT ACGTCTCGtT TGATAATCCA TTTCTTTTAA 3360
 AAGTTGGTCT TCAGTCTTAT CTTTTAAATC TAACACTGAC AGCCAACGGA TTTGGCTCAT 3420
 TGAATCATAA CCTACAGGGA AACCTTGGTG TTTATAACCT AATTTATCCA TTGTTCTaAC 3480
 25 AAATGCTCGG TTATCATAAG ATTTAACAAT TTCACCGTCT GCATTGCGTA AATTTTCAAT 3540
 TAAATATGGA TCTACAAGGA CATATAACA ATTGTGTTTC TTAAATATG ACGTTAATGC 3600
 30 TTTAAAGAAA AATGCTACTA ATGATTGATT TGTATAATCC ATCACTGGCC CGCGATGTGT 3660
 ATAAAAATAT TTGAAAATT TAAGTGTGCG TGCTTCTGTC AATAAGCATC CTGCAATCAC 3720
 TTGACCATTG TCATCTTTAA CCCCTACrAG ATGCACATCG CCTTTTAAAT CAACTCTATG 3780
 35 ATTGTAATGA ATAGCTGATT GTGTGTAATG TGAAAAATGC T 3821

(2) INFORMATION FOR SEQ ID NO: 305:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1422 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:

GCGACACCAT TGTTTGcAGC AACTGTTAAT TTACCAGTTG ATGGTTTCGA CGGTGTTGTT 60
 50 GGTTTTGATG GTGTAGTAGG GGTAGTTGTA GATTGCGTAC CCCATGGCGC CACTTTACCC 120
 ATTTTTATTA AATATTTTTC ATTAATTAAG TCATATAATT GATCATAACT ATAATTATGA 180
 CTTCTTAAAT ATCCATGTGG ATcGGCATGG TCAGTACCAC CTAAATATTT ACTTACAGCG 240

TATTGTAATT GTGTAGCTGC ATAGTCAGCA TAGTTATTCA TTGAACGTGC AAATGAAGCA 360
 TAGTCGTGTG TGTGTACGAT TTCAACATTG ATGAATCTAG GGTTACCGAC TGCACCGACA 420
 5 CCCCCAAGATA AGTAATCCGT TGGTGCTGTT TCGATTATAC GATCCCCATC AACAAATGCA 480
 TGTACGAATG CGTTTTGATA GTTATTTTTC ATATAACTAA TTTCACCAIT TATCGTCGAA 540
 CGATCATTAG CTGTATCATG AACTACGATA CCTTCAGGAC GACCTACGCC GTTACGGTAT 600
 10 GCGTATTTAG GGAAGTAAGA TGTATAATCT TCTTCAATTT TAGGTGCTTT TAAGTTATTT 660
 TTACAAATGT AATCGTTAAT TGAAGAGTTT ACTTGTGGTT TATATTTTGG CAAACTCGTT 720
 TTTGGTGTG CAGCAACTGA TCTTGGTTGT GCTGAAGCGC TAAAAGTAGT TACTTTAGGT 780
 15 GTCGCTTCAG TTTTAGCTTT AGGTGCTGAT GTAGTTGCAG CTTTAGGTGC TCGGGTTTTA 840
 TATTGCGTYT CAAGAGCTGC AGGTTTAGCA GCTGATTAA TTAATTCTGG ATTAATTTGA 900
 20 TTTTCTGAAT TATCATCTTC ATCATCAACT AAACATAAC CAGCATTGT AACATTAGTG 960
 TTAGTTTTAG GTGCTGTAGT GCTTGTGAC TTTGCAACAG GCTGCGTATT ATTTGTAGTC 1020
 GCTGATTGAT TAGCAGGAGT GTCACCATTT ACTTGTGCAG TATCAACTTT TTGACTTACT 1080
 25 TGAGCATTGC CTGTTTTGTT ATTTGCTGTT TTTGGTTGGA CAATAGCAGG GTCTTGATAT 1140
 ACTTGAGTGC CAGAAATGTT TTGCGTTGGA TTTTTTACCT CAGCTTTTGC TTGTTTCAGTA 1200
 GTTGCTTTAA CTTTATTACT ATCTAAAACG TTTTATTAG TAGTTTGATC TTGTGTCGTC 1260
 30 TCAGCTGCTT GAACTTGATG TGCAGTGACT GCTGAACCTA CAAGCGTTAA TGCAACCATT 1320
 GAGGTAGTTT GTAATTGAAT TTTTTCGCCA TTCTATTTAT TACTCCyAAC ATTTATTAAT 1380
 TATtACTAAC ATTATAGTAC CTGTrTTATA TACCTGTGCG TA 1422
 35

(2) INFORMATION FOR SEQ ID NO: 306:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6076 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:

ACGCGTTGGT TATTTCAATC ATAATATTAC TCTGCAAATA CACAGCCCTG TAACCGAGTA 60
 AATGGATAGA GATTTGAACA AATGAAAACA ATCAACTAAT GGAAAGGATA AAATATTATG 120
 50 CACAACAATA ATGAATTCAA CAAAAAAGT AAAGATTTTA TAGGTAGCGA TAAACGGATG 180
 GCTTTAGTAA AGGGTTATGT TAACGAGTAT AAATTAACAA CTGTTTTAAG AGCTTTAAAT 240

	GTAGATAATC GAATTATTCC TAAAAATATA ACTCAAACA CAATCTTCAA ATTGAGTAAT	360
	TTAACCTTAG AGGTTAACCT CTATGAAAGA AACACAATTT ATCACGGAGA CTTTGCTATT	420
5	TACTATCCTG TAGAATCTGC TCTAATAATA GAAACAGATA CTCCTAGGCT AATTAATCAT	480
	ATAAATAACA ACCCAGTCCA AAAAATATTT ATAATAACAA CAAATGACTG GAGCTTTAAT	540
10	ACACTTGaAT TAGAAAAAAT AGTAGACGAA ACTATAATTT ATGATTTAAA ACAAGAAGAT	600
	CCTAAACAAT ATGAAATTTT ATATAAAAAAT AAACACGGAA AATTGCCTTA TTGATGAATT	660
	AAAACGCAAA CTTAGAACAC CTTACCCCTT TTCCGTCCTT ATTGTAATTT TCCAACACT	720
15	CAAATTAAAA ACCACCCGTG CAAACAGGTG GTTTTATTAT ATATAATTAT TCCCACCTCAA	780
	TTGTGCTTGG TGGTTTTGAT GTAATGTCAT AGACTACGCG GTTGACGTGA TCTACTTCGT	840
	TTACGATACG ACTAGAAATC TTTTGTAAGA CTTCCCAATC GATGCGTGCG AAGTCACTTG	900
20	TCATACCGTC GATAGATGTT ACTGCACGAA TACCTACTGT GTGATCATAC GTACGGTAGT	960
	CTCCCATAAC ACCTACTGAT TGAATGTTTG GTAACACTGT GAAGTATTGC CAAATTTCTC	1020
	TTTCAAGACC TTCTTCTCTA ATCACTTGGC GTAAAATCGC GTCTGATTCT CTAACGATT	1080
25	CTAGTTTATC TTCAGTAATT TCTCCAAGTA CACGAATACC AAGACCAGGT CCTGGGAATG	1140
	GTTGTCTCCA TACTAAATGT TCTGGAATAC CTAACTCAAT ACCTAATTTA CGTACTTCAT	1200
30	CTTTAAACAA TGTATTGATT GGTTCGATTA ATTGGAATTC CATGTCTTCT GGTAATCCAC	1260
	CAACATTGTG GTGTGATTG ATTGTTTGTG CTGTTTTAGT ACCTGATTCTG ATGACGTCTG	1320
	TATATAGTGT TCCTTGCGCA AGGAAGTCTA CACCTTTCAG TTTTGATGCT TCATCATCAA	1380
35	ATACGTATAC AAATTCATTA CCAATGATTT TACGTTTTTG TTCAGGATCT GAAACACCTT	1440
	TTAATTTATT CATAAAGCGA TCTTTCGCAT TAACACGAAT AATATTCATG TTGAAACCTT	1500
	CACCGAATTG CTCCATAACC ATGTCGCCTT CACCTTTACG AAGLAAGCCA TGGTCTACAA	1560
40	AGATACATGT TAGTTGATCA CCTATTGCTT TATGCAATAG TACAGCTACA ACAGATGAAT	1620
	CTACGCCGCC ACTCATCGCA CATAATACAC GACGGTCTCC TACGCGTTGA CGAATCTTTT	1680
	CAATTTGAT TTCGATAAAG TTTTCCATTG TCCATTGACC TCTACAATCA CAAACACGAC	1740
45	GGACAAAATT ATTTAATAAA TCATTACCAT ATTCTGTATG ACGTACTTCT GGATGGAATT	1800
	GAACACCATA AATGCGACGT TTCTTATCTT CGATTGCTGC ATAGTCTGTG CTTGGGCTAT	1860
	CAGCGATAAC TTCAAAGCCT TCTGGAATTT CAATAACTTT ATCAGAATGA CTCATCCAAA	1920
50	CAGTTTGTTT TGCTGGTAAG CCAGCGAATA ACTCATCTGA CTTGCGATTA ATGATTGCTT	1980
	TACCGTATTC ACGTTCATTG GCACGTTCAA CTTTACCACC TAATAATTTA GTAGTTAATT	2040
55		

	TAAATGAACC TTCTTCATAA ACTGAATTTG GACCACCTGA TAAGATAATA CCTTTTGGAT	2160
	TCATTTTCTT AATTTCTTCA ATTGAAATTT CATGATCGTG TAATTCACCTA TAAACGCCCA	2220
5	TTTCACGAAT TCGGCGTGTA ATTAATTGGT TGTATTGGCT ACCAAAGTCT AAGACAAGGA	2280
	TTAACTCTTG TTCTTTTGCC ATTTCCATAT TTGTCGTTCT CCTTTATCTT AATTAGAATG	2340
10	AGTAGTTCGG TGATTCTTTC GTAATTTGAA TATTATGTGG ATGGCTTTCT GCTAAACCAG	2400
	CAGGACCCAT ACGTGTAAT TGTGCTTCTT CGCGTAATTC TCTTAAATCG TGTGAACCAG	2460
	TATAACCCAT ACCAGCACGC ACACCGCCCA TTAATTGGTA AATTGTATCT TGTAACGCAC	2520
15	CTTTATAAGC CGTACGTCCT TCGATACCTT CAGGAACAAA TTTCTTAGGC GCTTTGTCCT	2580
	CTTGGAAGTA ACGGTCGTTT GAACCTTTTT CCATCGCACC TAAAGAGCCC ATACCACGGT	2640
	ATACTTTATA TTGTCTACCT TGGAAAATTT CTGTTGCGCC TGGGCTTTCT TCAGTACCTG	2700
20	CTAATAAGCT ACCTAACATA ACCGCATGTC CACCAGCAGC TAATGCTTTA ATGATATCTC	2760
	CTGAGAATTT AATACCACCA TCAGCAATGA TAGCTTTACC ATGTTTGCGT GCTTCAGTtG	2820
	CACAATCATA AATTGCTGTA ATTTGTGGTA CACCAACACC TGCTACAACA CGCGTCGTAC	2880
25	AAATTGAACC TGGGCCAATA CCAACTTTAA CAATATCTGC ACCCGTTCA AATAAATCTT	2940
	TTGTTGCTTC TGCAGTTGCT ACGTTACCTG CTACTAATGT GATTTCTGGG TAAGTCTTCT	3000
	TAATATGTTT CACTTGATCG ATAACACCTT TAGAGTGACC ATGTGCTGTA TCGATAACTA	3060
30	AGACATCCAC ACCTGCTTCG ACTAATTTTT GAGCACGAAT ATCAGTATCT TTTGAAATAC	3120
	CAATTGCTGC GGCTACAAGT AGACGACCAT GTTCATCTTT TGCTGCATTA GGAATTCCA	3180
35	TAACTTTTTC AATATCTTTA ATAGTAATAA GACCTTCTAG ACGTCCGTCT TTAACATAATG	3240
	GTAACTTTTC AATCTTATGT TTTTGGAGAA TTTTTCTGCT TTCTTCAAGT GTTGTATTCA	3300
	CTGGAGCTGT AATTAAATtT TCTTGCGTCA TTACATCTAC AATTTTAATC GAGAAGTCTT	3360
40	CAATAAAACG TAAGTCACGG TTTGTTAAAA TACCTACTAA GTTGCGATCT TCTTTATTAT	3420
	CAACAATTGG TACACCTGAA ATACGGTATT TACCCATTAA TGCTTCTGCT TCATAAACGC	3480
	TTTCTTCTGG CGTTAAGAAA AATGGGTTTG AAATGACACC ATTTTCTGAG CGTTTTACTT	3540
45	TTTGAACCTC GTCCGCTTGT TCTTCAACGC CCATATTTTT ATGAATAACA CCTAAACCAC	3600
	CTTGACGAGC CATAGCAATC GCCATTTTATG ATTCAAGTTAC AGTATCCATA CCAGCAGAAA	3660
	TAAGTGAAT ATTTAATTTA ACTTTGTCTG ATAATTGTAC GCTTAAATCA ACGTCTTTTCG	3720
50	GTAAATATATC AGATTGTGCT GGAATTAATA ACACATCATC AAACGTTAAT GATTCTTTTG	3780
	CAAATTTACT TTCCACATT AAAAACAGCC TCCATTTTTT AAATTAATTA GTTATATTAT	3840

55

	GCAGAGATTG CGCCTAAAAC AATTCCGTTT TGAGTCAACC ATGCAAATTG TTCACCTAAA	3960
	CCTTTAAATG CTTGTGGTAC AGCGCTTATA CCAGTACCTA AtCCTACTGA TACAGCGATA	4020
5	ATTAATAAAT TGTTTTGATT TTTAAAATCG ATATGTCCTA ATATACTAAC ACCATATGCC	4080
	ATTACCATGC CAAACATAGC TATCATCGCA CCGCCTAACA CAGGTAGCGG TATGATATTT	4140
10	GCTAATGCGC CAAGCTTAGG TATACAACCA CATATAAGTA ATAACACGAC CATGCCGTAT	4200
	ATAACATTGT TTTTCTTAGC GCCGGATAAA GAAACAAGTC CTACATTTTG CGAATAGGCT	4260
	GTATACGGAA ATGAATTGAA TATAGAACCT AACACTATCG CTAGACCTTC CGCAGTATAA	4320
15	CCTTTACGAA AATCTTTTCT TTCTAACTTC TTACCGGTAA TTTCACTTAA CGCATGATAG	4380
	ACACCTGTCG ACTCAATTAA ACTAACGATA GCTACAATAA AGAACACTAA CGTCGATGTC	4440
	ACATCAAAGC TAAATCCAGA GAATCTAAAC GGCCTGGGA TGCCTAACCA ACCGGCATGA	4500
20	TTGACTTGAT TAATATCGAC CATCCCAAGT AAGCCAGCAC CTATCGTTCC TAAAACGAGT	4560
	CCAATTAATA TGGCAATACT CTTAATAAAT CCAGTTGTGA ATCTTTGTAA AAGAAGAATA	4620
	ATGATTAATG TCATTAAACC TAACAAAATG TTCTTAACAT CTCCATAGTC CTTTGCACCT	4680
25	TGACCTCCAG CTAAGTAATT CATTGCTACT GGCATTAAAT TGATACCAAT GATAGTAACA	4740
	ACACTACCCG TTACTACTGG TGGGAAGAAT TTTACAAGAT GTGAAAAGAA AGGCGCGATG	4800
30	ATAATACTA ATATCCCTGA TAAAAATAGC GAACCATAAA GTACATCTAT TCCTTTCGTT	4860
	TGACCAATTA AAATCATGGG CGCAACAGCC GTGAATGTAC ATCCAAGAAC GATTGGTAAT	4920
	CCTGTTCCCTG TTACTTTATT GGCTTGTAAG AATGTGGCAA CCCCACACAT AAATATATCT	4980
35	ACTGTACTA AGTAAGCGAT TTGTTCAAGT GTAAACTTCA AACTTGTACC AACAAATGATT	5040
	GGAACCTAAGA TAGCACCTGC GTACATAGCT AAAAGATGTT GAACACTTAG GATTAAATTT	5100
	TTCAATTATC TTCTCCACCC AATGTCACCT TGTTCCTTC TAGTGAAGCA ACCTTGCAGA	5160
40	GAGAAGAAAC TGTAAACCT GCTTCTTCTA AACGTTGATG CCCATTTTGG AAACCTCTTT	5220
	CAACAACAAT ACCAATACCA GCTGTCTTAG CATTGCTTG CTGTGCGATA TCGTATAATC	5280
	CTAATGAAGC ATCACCATT TCTAAAAAGT CATCGATGAT AAGTACAGTA TCTTCTTCTG	5340
45	ATAAAAACTC TTTTGAAACA ATGACCGTAC TTGTTTTATT TTTAGTAAAT GAATGAATAG	5400
	ATGTTTCATA ATAACCATCC GTCAAAGTGC TAGGTTTTGC TTTTTCGCA AATAAACATG	5460
50	GCACATCAA ATGCAGTGCA GCCATGATTG CAGGTGCGAT ACCGGAAGCT TCAATGGTTA	5520
	AGATTTTAGT AATCCCTTTA TCTTTAAATT GCTCGTAAAA AGTGCGACCA ACTTCATTCA	5580
55	TTAACTTTGC ATCAATTTGA TGATTTAAAA ATCCATCGAC TTTTAAATC TTCTCATCAA	5640

ATTTGTGTGA AACATTTTGC TCTTAAATTG GTGCTAGATA CAAAAAATC CCCAAACTAA 5760
 ATAATAGTTT CAGGGTTTAT GACTGAACGa ACATGCATAA CGAATTTGTC ATGCAATCAA 5820
 5 TGTAAGAGAA GTTTCATCAA ATAAACTGTG ACCATCATAT AAAATGATAT AAATCACCCA 5880
 CCATGGTTAC AATTTAATGG CTGAAGCTAC TCCTAGTATT GTGTTGTTAC TCATAGTCAT 5940
 10 GTCGTTCAAG GCAACATGGT AGAACTTCT AAAGCCATAT TCTTTAGATT ATATGAGTTT 6000
 ATGTAAATTA TTTAACGATA ATAGCAAATT TTCGGCATT TTTCAATAAC TGCTTAGGTA 6060
 ATCTTTTAAT AGTTTT 6076

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6136 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

25 ACCGCCGATT GATAGCTTTA CTGCTGCGAT AAAACTGCCT TCAACAATTG GTGCATCAAC 60
 TTTTAACACA CGATGATTAC CATCATACAT TTCAATTGCC ATATCTACAT TCATTTCTGA 120
 30 AGATCCAATA TCGTAAAAAC ACAATGCATC ATCCTCTAAT TTAGTCAAAA CTCTTGGAT 180
 GATATCAAAT GAAGTTCCAA TTGAACCATC TGGTAATCCC CCGATTGGTA TAATATCAAC 240
 GTCACCTGCC ATTTGCTTTA ACAAAGATT TGTACCACTT GcAATTTCTT TACTGTGGCT 300
 35 AACAAGTATA ATTTTAGGCA TTATCTTCAT CTCCAATCAA AGCGTTTAAA ATATAAACCA 360
 TACTTTGAGC ACCTGGATCA ATATAACCTT TTGATTCTTC TCCAAAATAT GCAGCTCTAC 420
 CTTTCGTTGC TACCATATCT TTTGTATTAT CTGCTAATTG CTGTAAATCA TTGAATGTTA 480
 40 AAGTTTCACC ATTTTAAAGC TTCTCTGCTG CTCGCGCTAC TACATCATAC ATTGTCTTTT 540
 CATTTAAAGT AACTTTACCA CGTGATTCAA CCGCTTCGGC AAATGCCTGA ATTAGTGTA 600
 TGAAATCTTG ATTATCCATA TCATCTTTGG TGAATGCAGA CATTTTAACA AAGCTAAAGC 660
 45 CATACAGTGG TCCTGATGCA CCCCCAACAT TTGACATCAA TGCCATACCA GTTGATTGTA 720
 ACAATGATTG CATTGAGCTA TCATCAAGTT TGTCTTTAAG ACTACTAAAC CCACGAACCA 780
 50 TGTTAACCCC GTGGTCACCA TCACCAATTG CTCGATCTAA TTCAGTTAAT TCAGATTCAT 840
 GTTTTTTAAA CGTTTCTTCT AAATTTAATA AACGTGCTTT CATATCATTC ACTTTCATTT 900
 GTGCCACCTC ATAAATGTAT ATTTATTCAT ATTCACCTCT TATTTAAAGT ATTGACTTGT 960

	ACCTTGCA TG TCTAAAGATG TCAATAATC ACCAACAAC CATTAGCAA CATTAAACCGT	1080
	TCTTGCA GCT AAATTTTGTT GAATATATTT AGTTACGATA TTTAATTCAG ATAACGGCGT	1140
5	ACCACCCATA CCATTTACCA TTAATATGAC ATCATTGGCA GTAACCTCTT TATACAATTC	1200
	GTCTAACAAT GTTCCAACAA TATGATCAAT ATCCTTTACT TCTTCCCTAT GAATACCTTT	1260
10	TTCACCATGT ATACCAATAC CGATTTCCAT TTTGTCGTCT TCAATATCAA AGCCATATTT	1320
	TCCAGTAGTT GGAACAAGCG GTGGCTCAAT TGCCATACCA ATACTTTTAA TTTCAGGTAA	1380
	CAACGCTTCT ACACGCGATT TTATCTCTGT TAATGAATAA CCTTTTTCAG CAAGATAACC	1440
15	GGCAAGCTTA TGAACAAACA CTGTTCTGTC AACACCACGA CGTTGTACTT CGTTTGTAC	1500
	AGCAATGTCTG TCACGAACAA TAACAGTTTG AACATTATA CCTTCCATTT CTGCAAGCTC	1560
	TTGTGCCATT TCGAAATTCA TCACGTCACC TGCATAGTTT TTTACAAC TAATAACACC	1620
20	ATCACCAGTA TCTACTGCTT TAATAGCTTC TAATATTTTA TCAGGTGTAG GTGATGTAA	1680
	TACTTCGCCA CAAACCGCTG CATCTAGCAT ACCTTCTGCA ACAAACCGG CATGCGCAGG	1740
	TTCATGTCCG CTTCCACCTC CAGAGACTAT TGCTACACCA TGTCTTTCT TAGCTTTTTT	1800
25	TACAATAACT GTATTAGCAA TCAGATCTAA CTCTGGGTGC GCAATTAACA ATCCTTCAAG	1860
	CATATCAGTT AAAAATGTTT CTTTTTATT GATTAACTTT TTCATCATGT TGTACCTCCT	1920
30	TGGTATTATC AGTCATAGTA TAAACAGAA TAAATGAATG CGCTATCATA AAGAATTAAC	1980
	CTGATACCGT TATCAAATAT ATTTTGTATA AGATCCTCTT GATAACTTAA TGGTTCATTA	2040
	TTGAAAAAAT AATAGTCTT GAGTGGCTCA TTAAATAAAA AATGCTATGA AAAGCCTTAT	2100
35	AATAAAGTGC CTTTCATAGC aATAAGTTGT GTCCATTGAC ACTATACATT TTTCGTTTTG	2160
	TACATTAAAT ATAAGAAATA CGGTGCACCA ATAATTGCTA CGACAATACC TGCTGGAACC	2220
	CCACTGGTT GTAATACAAT TTTGCCAATT GTATCAGCTA TAACAAGTAA ACATGCCCCCT	2280
40	ACTAAAATGG CAATTGGTAA AAACAAC TGG	2340
	TGCGGACCCA TTAATCCGAT AACGAAATT GAACCTGCTA cTGCTACAGC AGCAGATGAT	2400
	AACATCACTG CGATAAAGAA TAATATTAAA CGTTCCTGTC TTAACCTTAC ACCTAGACCT	2460
45	CGTGCAATAT TATCACCCGT ATGAATAATA TTTAGTGTAT TCGATTTAAA TAGTAAGTAA	2520
	GGAATAATAA TCAACACCCA CGGTAAAAAT GCAATGACAA ATGGCCATTC GTCACCCCAA	2580
50	ATATTACCTG CAAACCAAGC AGCGATGAAA TCAGATTGCT TATCATCAAA TTTTGACATA	2640
	ATTGTAATTG AGCCACCATA TAATGCTGTT TGTAACCTA CACCTATTAA TACCATACTC	2700
55	GCAGGTGTAA CACCTTCATT TTTATTAAAA CTGaAAATAA AAATAATCaA TGCAGTGgTG	2760

	CCAATTGCAA TAAATAATGC GATTGCAAAT CCGCCACtGC GTTAATACCT AATATACCTG	2880
	G TTCAGCTAT TGGATTTTTT GTGACACTTT GCACTATTGC ACCACTAATA CTAAGCtGCG	2940
5	CCAGCCAAAA TAGTAATCAT CATCCGAGGT AACCTGAAAT CTAATAAGAT TAACTCATCA	3000
	ATGGCATCAC CTTGTCCAAT TAAAGTTTGT AAAAATCTTT CAACAGGTAT GTTGTATTCA	3060
10	CCTGAGGTAA TACTCCAAGT ACAACCTAGA AATAGTAGAA TGCTAAAAAC AGCCAGTGCT	3120
	ATCAATTGTC TGC GTTTATT ATTTGAACTA ATCATATTGA GCGTCCTCCT TTTTAACTA	3180
	AATATAAAAA GTAAGGAACA CCGATAAATG AAATGATTGC ACCAACAGGC GCTTCTCCTA	3240
15	AATATCGTGC TATCACATCG GCAACAAGCA CGAGTATCCC ACCTAACAAG GCTGTTAATG	3300
	G TAGAATTTT AGCATAATCA GTTCCAATTA AAAATCTTGC TATATGAGGT ACCATCAAAC	3360
	CTACAAATGC AACTTGTCCA GCGATAGCAA CTGCAATACC TGCTAGAATC ATAGCAATAA	3420
20	TTAAACATAT GCCTCTGATC ATTGTTACAT TTTGACCTAA ACCTTTAGCT AATGATTCAC	3480
	CAAGATTTAA AATGGTAAGT TGT TTTACTAA TTGTTAATAT AATGAATAAC GCAATACCAA	3540
	TTAATGGAAT TGCCCACTTA AGGTGTGACC ATGTTGTGCC TGAAACGCCT CCAGCAGTCC	3600
25	AAAATGTTAC TGTTTGATTT AGTCTAAAAG CTAATGCAAT ACCTTGACTT AGCGCTGTTA	3660
	ACATAGCACT TACTGCTGCA CCCGCTAAAA TAATACGCAT CGGATTAAAT CCATCACGTC	3720
	TAGATCGGCC TATCATTAAT ACAATAGCAC CTCCTAGAAT AGCACCTAAA AATCCAGCAA	3780
30	ACATCAATAT TAAAAATGAA GTGTTTGTA AAACTGCATA TGTTAATGCT AAAGCAAATG	3840
	AAGCACCTGA ATTTAAACCT ATGAGCGCCG GATCAGCAAG ACCATTACGA GTAACACCTT	3900
35	GTATAATCGC ACCAGAACT GCAAGCGCCA TACCTACAAT TACTGCTGCT ATATTTCTGG	3960
	GAATCCTAAT CTCATTGATG ATGTTTGTCT GTTGATTGCT AGGATTATAA TTAAAAATAG	4020
	CCTCTATAAT TG TAGAGGCT TGAATTTTGG CGTCACCTAT TAATGTAGAA ATAAATAGTG	4080
40	TGATTAGTAG TATCATACTT AAACCTATAA TATAGGATAA AAAC TTCAAT GCGTTGGGT	4140
	TCTCTCTATT TGT CATGTTA ATTGTCCTTT TTATCATATT AACTTACTTA ATTAAGAATA	4200
	AGCTCTGCGA CATAAGTCAT AAGTTACCAG TAAAGGTTTT CCAGTTTTAG GATCTTTACT	4260
45	TAAAACAACA TCAATATTAA AAAC TTTTTC TAATATTTCC TGTGTTAATA CGTCTTCTGT	4320
	TGAACCTGTA GCGATGATAT CCCCTTCTTT CATCGCAATA AGATGATCTG AGAAACGAAT	4380
50	CGCTTGGTTG ATATCATGAA GAACCATGAC AATTGTACAA CCTTGTTCTT GATTTAGCTT	4440
	CTGAACTAAT TCTAGTATTT CTAATTGATG ACAGATATCT AAATATGTTG TTGGTTCGTC	4500
55	TAAAAAGATA ATATCAGTTC TTTGTGCTAA TGCCATTGCA ATCCAAACAC GTTGTCTTTG	4560

TGCCCAATCA ATTCTTTCT TATCCTCAGC AgTTAATCTA CCAAATCCTT TTTGATGTGG 4680
 AAAACGACCA TATGAACTA ATTCCCCAAC AGTTAAGCCA TCTGCTACTT CaGGTGaTTG 4740
 5 aGGTAAaTG GcAtTTTTt TGcAATCyCy TTCGTAGAt GTGtATGAAT ATTTTCACcA 4800
 TCTAAAAATA CTTCGCCTTC TTAACTGCC AATAAACGTG ACAATGCCTT TAGCAAAGTA 4860
 GATTTCCCGC AGCCGTTAGG ACCAATGATT GACGTCCTT TGCCATCTGG TATTTCAACA 4920
 10 TCTAATTAT TTATAATCGT GTTATCCCCG TAACCAATTT TAACTTGTTG TCCATGCAAA 4980
 CGATTCAATA TTTCCCTACT TTCAATAAAA TTCTTTCTGT TTATAAAAAA TAATTTCTAT 5040
 15 TTTTAAATTA TCAATTTTCA AAGACATCCC AATTGATAAT GATTATCATG AACATCATT 5100
 TAACATTTT CAATCTTATT GACTAACATT ACTTTTAAA TTGGATAGCT CGATTTGTCA 5160
 TGTCTGTAT ATTACTTTTA TAAAATAAAA AACGCCACA GATAAGTCTT CATAGTTCAA 5220
 20 AAACCTGTCC GTGGACTTCT ATTTAAGTAT GTGTGCTCAT ACCATTTATT TATTCATCTG 5280
 CAAGAAAGCC ATTACCATAG ACATCTCTTA CATCATGAAT TACGAGGAAT GCATCTTTAT 5340
 CGATTTGTTT AATTAATCGC TTTGCTTTTG AAACCTGTGT TTTAGAAATA ACAACGTATA 5400
 25 AGACATCTTT TTCTTCACGC GTATAATAGC CATGTCCGTT TAAAATGGTT AAACCTCTTC 5460
 CAATTTGCTC GTCTATTGCT TTGGCAAGTT TGTGGGATT AGTTGAAATA ATCGTCATAG 5520
 CTTTTTAGT GTTTAAACCT TCTATGACAT ATTCCATCAC TTTTGTTCTT ATATAAAGTG 5580
 30 ATATTACTGT TACTAATACT TTATCAAGTG GAATAACTGT AAGTGAAATT GCAACAACGA 5640
 TCATATCGAA GAAAAGCAAA GCATATGGCG TGCTTACATC GAGGTATTTT GTTGCAATTC 5700
 TCGCCAAAT TGTGTACCT GCTGTTGTAC CGCCTGCAAG GATAATTACT CCGATTCTTA 5760
 35 GTCCAACGCT TACACCACCA AAAATGGCAT TCACAATGCT GTTTCAGTT TCTACTGCC 5820
 ATGATTCTGT TAACTCAA AATATTGAAA TAAGAATTGT TACAAGAATA GTTAAGTACA 5880
 40 TACTTCTCTT ACTCAAAAAt TTATAACCTA TGGCAATCAA TACTGCGTTG ACCAAGAAGT 5940
 TAGTGATGGC TGGTGAAATA TGAAACGCAT AATATAAAAT AATTGCTAAA CCTGTAACCC 6000
 CGCCTTCACC TAAGTTACCA GAAATATAA ATGCATTAC ACCTGCAGCA AAGATAAATG 6060
 45 AACCTAAGAC AACTAGTATT AAATCTTTAA CCGTTTATT CACGAAACCA TCCCCTTTAT 6120
 ATATTTATTA GACTAT 6136

(2) INFORMATION FOR SEQ ID NO: 308:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2576 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:

5	GATATCGTAs	CTACTGAAAA	GTCATCACCA	CCATGGAATG	ATTTCTTTAA	ACGTTTTAGC	60
	TTCTATGCAA	TTGCAATTCA	ATACTTTGTT	GTACAATTTA	TCATTACATT	ATTCTTAATT	120
10	TGGTTACCGA	CGTATTTAAC	AGAAGTATTC	CACGTTAACT	TTAAAGAAAT	GAGCATTAGT	180
	TCATTACCTT	GGTTATTAAT	GTTCTTCTTA	ATCTTATCAG	CAGGTGCAAT	TTCTGACCGT	240
	GTATTAGGAT	TAGGTCGTTT	AAAATTCGTA	GCTAGAGGTG	TAATTGCAAT	TGCAGGATTT	300
15	ATTGTGTTTG	CAGTTTCAAT	TATCTTTGCT	GTACGCACAG	GAAATTTATA	TGTAAGTATT	360
	TTCTGGTTAT	CACTAGGTCT	TGGTGGTATC	GGTATTTCAA	TGGGTATGAG	TTGGGCTGCA	420
	GCAACTGACT	TAGGACGTAA	CTTCTCTGGT	ACAGTATCAG	GGTGGATGAA	CTTATGGGGT	480
20	AATATAGGTG	CATTAATCAG	TCCGCTATTA	GCAGGTCTAT	TCGTAGAACA	TTTGGGTTGG	540
	ACAATGACAT	TCCAATTGTT	AATCGTTCCA	GCAGTAATCG	CTGTGATTAT	GTGGTTCTAT	600
	GTGAAACCAG	ATCAACCTTT	AATTGTTAGT	GATGATAAAG	CAATAGAAAA	ATAATTTAAA	660
25	CAAGCAGTAA	GCTTTCACAT	AGTTGGGGCT	TATTGCTTTT	TTTGC GTTGA	AATTGAAACT	720
	TTTTAAACA	GATATGGTTT	AAGATGAAAA	TGAAGTTATT	GAAATGATAT	ATGTAAAGAA	780
30	ATAAGGTTTT	AAAACATTAG	TCAGGTAACG	CTTGTA AAAAG	TACATATAAA	TTTTAACTAG	840
	CGCAAAGGTG	GGCGACCAAA	GtTcaACGAT	GTTAAATAAC	aTTAGTAATT	AATTTTAATT	900
	GGACTTTAAA	AGTTTTTAAA	TTTAGATAAT	TGAGCATAAG	GTGTTATAAT	GACATATGTT	960
35	GCGTAATTAA	AATTTATAGC	AACAAATTCA	TTTTAACTAT	GCTAATAAAA	AGATTATGGA	1020
	AATATTTTGA	CAAGGAAAGG	AGAAGTCGAA	ATGACATCTT	TTTGACATCA	CTCATAAAAA	1080
	TCAATCGACT	TAACCTAGAC	TTTTATAAAG	GTGTAAGACA	GGGACTGTTA	ATGATTATTC	1140
40	CTGCAATAAT	CGGTTACTTA	TGTGGTAATT	TCCAATTTGG	ATTATTAGTT	GCAACCGGAA	1200
	CACTAGCCCA	TATTTATGTT	TTTAAAGGTC	CGTCGCGATC	TAAGCTGCGA	ACTGTAATAA	1260
	TTTGTAATTT	AGCGTTTGCA	ATATGTATGA	TGCTTGGTAC	GCTAACAGCC	AAAACGCCAC	1320
45	TCGTTTTTGG	AATGACATTA	TTAATTGTTA	CGGTTATACC	ATTTTATATA	TTTACTGCCT	1380
	TAAAAATAGC	TGGACCGTCA	TCGACATTCT	TCATTGTGAC	ATTCAGTCTA	CCCATTAACT	1440
50	TACCTATAGC	TCCCGAAGAA	GCATTATATA	GAGGCTTTGC	GATTTTAGTA	GGCGGTATAC	1500
	TTGCCACTAT	GATGGTGTTA	ATCACGATCG	TATTTTCTAA	AAACAAAGCT	GAAGAACAAG	1560
	CAATTCAAAA	TGATTTTAAA	CTCATATCTA	AGTTGTTACA	CACTTATAAT	GATAAATCTG	1620

55

TCACTTCTAC TTCAAGTAAC GATAAATTAA GTAGACGTTT CCAAAAATTA TTATTATTAC 1740
 ACACATCTGC CCAAGGGATT TATTCTGAAC TGTTAGAGTT GAACGCTAAA CAAATTCGAC 1800
 5 CATTGCCAGA TGAGTTAATT GAAATGATGG ATCATATCAT TGCACAAC TA GATAATAGTG 1860
 AGGAAAATGT AAGATATTGG CGAAAAGAAG TGACAGTAAC AGAGGAATTT CAAAATTTAT 1920
 TCAACCATAT ATTGAAAATT GATGAAATGG TGCATGCAAA TGAAGCGCGT ATTGCGTATG 1980
 10 AAGCAGACAT GCGAAAACCT TTATATAGTA AACGCATTTA TCaAAATTTA ACaTTAGACT 2040
 CtATkGTTTT TAGAAATACA TTGAGATATA CAGCGATTAT GATGATAGCG ATATTTATTG 2100
 CGTTAATGTT TGATTTTGAA AAAGCATACT GGATACCGTT ATCTGCACAT ACAATATTAC 2160
 15 TAGGAACATC AACTATACAT GCAATCGAGA GAGGTATGGC ACGAGGTTTA GGTACTATTT 2220
 TAGGTGTGTT AGTACTTTCA GTCATATTGT TGTTTTCAAT ACCAACACCT GTTGCGAGTAA 2280
 20 TTTTAATGGG CATTGCAGCA TTGTTTACTG AAGCATTGGT GGGAGCAAAT TATGCGATTG 2340
 CAGTAGTTTT TATTACAATA CAAGTTATTT TAATGAACGG ATTAGCATCA CAGAATTTAA 2400
 CAATTAACAT TCGGTTTCCA AGAGTTATTG ACGTTGCAAT GGGTATTGTG ATTGCAATCA 2460
 25 TAGGTTTATT TGTCTTGGG CAACGTACCG CATCCGCATT GCTTCCTAAT GTAATGGCTG 2520
 AAGTTGTTTC TAAAGAAGCA ACGCTCTTTC ATTATTTATT TTCTGAAAAT CAATAT 2576

(2) INFORMATION FOR SEQ ID NO: 309:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 668 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:

40 CAAGCACATT ACGATGGTCT AATACTTTC TAATAATTTT TTCTTGTCCT GGTCGAAACG 60
 TTTCATACCC AAAGTAATGC GATAATGTTT GTTGATCAT AAATTGACCC CTTATTGTTG 120
 TTCTTTAATT TCTTCTAACT CACTCCATCT TGTGATGTCT AAATCATATT GAATTTCAAG 180
 45 TTGTTCTTTT TCTTCGTTTA ATTCTTTAAT TTTCCATAA TCTGCACCTG CCTCAATCAT 240
 GAGCACATCA ATTTCTTCCA TTCTTACTTC CGCTTGTTCT ATGCGTTTCA TCAATTGTTT 300
 ATATTCTAAT TTTTCTTTAT ATGATAAACC ATTTTCTTA CGTACAGTTG TAGAAGATTT 360
 50 AGATTGTTGC TTCAATGTGG ATTTATTTTT ATCTAATGAT TTTTATAAC TTTTATAATC 420
 TTCAAAAGTT CCGATAATCT TTTCCATCTG ACCATCATGA ATAAACCAAT ATGACTGTGC 480

55

AATATAATCT TCAAGTATTG TTAAAGTCTC AGTATCTAAA TCATTTGTCG GTTCATCTAA 600
 CAACAGAACA TTTGGCTGGT GTACGAGTAG ACGTAATAAA TACAAACGCT TTTGCTCTCC 660
 5 ACCAGATA 668

(2) INFORMATION FOR SEQ ID NO: 310:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12173 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:

CATCAGTTTA TTTTGAAAGG CAATGCGATC ATTTTCATGTA TTTATGTTGT TTGAAACATC 60
 20 GTTAGATAAC AATAGTGATA TTGCACATT TAAGCTGAAG ATGGAAAAAT ATTCGATAAT 120
 CAAATAAACA ATGAATTTTA GAAGGTACAA TGACGTTTCA TAATTTAAAT ATAGCTGAAT 180
 GTGTTGGTGA GTGATGTTCA CTATAGATT ATATTAATAT ACAAAGACAA AGGTTGTTAA 240
 25 TTTTTATTAA GCGTTAGGTT GAATGTATGA GAATTTTAGA TTTATAATAG AAGATAGAAA 300
 CGAAAATTTT TCTTAAAGC AGTAATGTTG ACTCAAATA AGCTATAATA ATGACACTTA 360
 TTTAATTGAT TAACATTTGC TAATAAATAT CAATATAGAA TATAACTTTC CAATAATGAC 420
 30 TGAGAAAATC GAAATGTCAG TCTCGAATCA TATAATTAGA AAATTGATTA TTTTCTGTCA 480
 ATTTAGGGTT GAACTATACA TATGATATTG TTAGAATATT TTTTAACATT ATATTTTATT 540
 35 GCTTTAAAGT GGAATATACT TGAAATAATT AGTAGAGGTG AGTAAGGATG AGTAATAAAT 600
 TAGAATCATA CAGAAGTGAG ATTGTATCAC TGAATCATCA AATTTTAGAC TTATTATCTA 660
 AACGTGGTGA ACTAGCACAA AAAATTGGGG AAGAAAAATT AAAACAAGGT ACACGTATTT 720
 40 ATGATCCACA ACGTGAAAAA GAAATGCTTA ACGACTTAAT CGATAGTAAC AAAGGACCAT 780
 TCAACGATAA TACTATTAAG CAATTATTTA AAGAAATTTT CAAAGCCTCT ACAGATTTAC 840
 AAAAATCTGA AAATGAAAAA CATTTATATG TATCACGTAA GTTGAAACCT GAAGATACGA 900
 45 TTGTAACATT TGATAATGGG GGCATTATAG GAGACGGCAA TAAATCATT GTATTTGGGC 960
 CATGTTCACT TGAATCATT GAACAAGTTG AAGCTGTTGC TAAAACTTA CATGCTAAAG 1020
 GTGAAAAATT TATTCGTGGC GGTGCATTTA AACCACGTAC ATCACCATAT GATTTCCAAG 1080
 50 GCCTAGGTGT TGAAGGACTT AAAATACTTA AACAGATTAA AGATAAATAT GATTTAAATG 1140
 TTGTCAGCGA AATCGTAAAT CCAAATGATT TTGAAGTGGC TGATGAGTAT TTAGACGTAT 1200

	AAAAGCCTAT TCTATTAAAA CGTGGTTTAT CTGCTACAAT CGAAGAGTTT GTTTATGCAG	1320
	CTGAATACAT TGCTTCACAA GGTAATCAAA ACATTATTTT ATGTGAACGT GGAATCCGAA	1380
5	CTTATGAAAA GCGACACGT AACACTTTAG ATATTTTCAGC AGTACCAATT TTTAAACAAG	1440
	GTACACACTT ACCAGTCATG GTAGATGTTA CGCATAGTAC AGGTCGTAAA GATATCATGT	1500
	TACCAACTGC GAARsAgCAT TAGCAGTTGG TGCTGATGGA GTTATGGCTG AGGTGCATCC	1560
10	AGATCCATCT GTTGCACTTA GTGATGCGGG TCAACAAATG GATTTAGATG AATTCCAAGC	1620
	ATTTTATGAT GAATTAAAGC CTTTAGCTGA TTTATATAAC GCTAAAAAGT TAAAATAATA	1680
	TTCCAAGGAA ACTATAGACT ACTTAACTAA TATGTCATGT TGAAGTAGAA TATTATCTTT	1740
15	GAATCGACAA TTTTAACTT ACAGCCATTC TAAGAGTATA TTACTTTTAG AGTGGCTATT	1800
	ATTTTTTGTA TAGAAATAAA GGTATACTGC ACTTAACGAT TGTTATAATA CTTCGACACT	1860
20	TGTTCAATTT CACAATTATT AAAGATTATG ACTGATAGCA GTAATTAAAA TTATAACTAT	1920
	GAATTATCTG TAAAATATAA TAGATTCACA CATTTGTTGC TGAAATGTGA ACATTTTTC	1980
	ACAAATGCAA TTGATATTTG AAAAGGCTTT CTCAAAACAT TACAATTAAA AATGAAAAAA	2040
25	GTTTATATAA AATTAAAATA TATCGTTCGT TATCATTTAG CGTTTGTTTT TATTTCAAGC	2100
	TTTTCGCTAA ATTTTCCAA ACAAAAATAT GTTACTGTAA ATTAAAATAT GGTAACTAT	2160
	GAAAATGAAA TGAAAACATG TTATTATAAT GAATAAACG TTTACAAGGA GGAAATTATG	2220
30	ACAGTTACTA TATATGATGT AGCAAGAGAA GCGCGTGTCT CTATGGCCAC AGTGTCGCGT	2280
	GTTGTTAATG GGAACCAAAA TGTTAAAGCA GAACTAAAA ATAAAGTTAA CGAAGTCATT	2340
	AAGCGTTTGA ATTATCGTCC AAATGCTGTT GCTAGAGGTT TAGCTAGTAA AAAGACAACA	2400
35	ACAGTAGGTG TGATCAITCC AGATATATCT AATATCTATT ATTCACAACT TGCTCGTGGA	2460
	CTTGAAGATA TTGCAACAAT GTATAAATAT CACTCAATTA TTTCAAATTC AGATAACGAT	2520
40	CCTGAAAAGG AAAAAGAAAT TTTTAATAAC TTATTAAGTA AACAGGTTGA TGGTATTATT	2580
	TTCCTTGGTG GTACAATTAC TGAAGAAATG AAAGAATTGA TAAATCAATC ATCTGTACCT	2640
	GTAGTAGTAT CAGGAACAAA TGGAAGGAT GCACATATAG CATCAGTTAA TATTGATTTT	2700
45	ACTGAAGCTG CGAAAGAAAT TACGGGAGAA TTAATTGAAA AAGGCGCTAA ATCATTTGCT	2760
	TTAGTAGGTG GAGAACATTC TAAAAAGCT CAAGAAGATG TTTTAGAAGG TTAACTGAA	2820
	GTGTTAAATA AAAATGGCCT TCAATTAGGT GATACATTGA ATTGTTCTGG TGCTGAAAGT	2880
50	TATAAAGAAG GCGTAAAAGC TTTTGCCAAA ATGAAAGGCA ATTTGCCAGA TGCCATTTTA	2940
	TGTATCAGCG ACGAAGAAGC AATTGGTATT ATGCATAGTG CAATGGATGC TGGTATTAAA	3000

55

	CCACAAC TTT CTAGTGT TAT TCAACC ATTA TATGAT ATCG GTGCAG TAGG GATGCG CTTA	3120
	TTAACAAA AT ATATGA ACGA TGAAAAG ATA GAAGAAC CAA ATGTAG TTTT ACCTCAC AGA	3180
5	ATTGAAT ACC GAGGAA CTAC AAAATA AATT CACAAA ATTA GGCATT CATC TAACGAC CCA	3240
	AATTAT ATGG GTGTTG GAAG AATGCCT TTT ATTTAT CTTT TAAAAT CGTT GCAGATT AGG	3300
10	TTACTT ATTG ACGAGT AGAT TCGTAC CAAC TCGCTAT ATG TAAAGCT AAT TTTTAT TTTT	3360
	TTTCACT AAT TTCTTT TGTG CGGGGG ACAT AGGTATA ATC ATTTAA ACGA TCTTCCC ATC	3420
	TTTTAG GTTAA TAATTC AGAT GAATAAT GTT TCCATTT ATT AATCC ATTCT AACGGT AAA T	3480
15	AACCACT TTG AATTGG TTGA TCAATT AAAC TTAAGA ATAC ATGACT CCAT GCACGT GGT A	3540
	CGACTCT CCA AATATT GTAG CCTCCG CCAC CAAACATA AT TACCTTT CCA TTCGTATA AG	3600
	AATCAG CTAA ATATTTT ACA AAATAT GGAA TTTCAT ATAA TGAATGT AAC GTACAATT TA	3660
20	GATGAG TTAG TGGATC ACGA TAATGT ATAT CGACACC ATT TACGCT TAGA ATAATAT CAG	3720
	GTTTAA AACT CTTTAC GACA GGCTCA ACTG TTAATTT AAA ACACTCCA AA AATGAT GCAT	3780
	CTTCTGT ATA CGGTTCA AGT GGGACATT TA CAGTGT GTCC ATAGCC GATA TCTTCAC CGC	3840
25	GCTCAG TATA GTGACC AGAG CCTGGG AAAA GAAATTT TCC GGTTTC ATGG ATAGAATA AG	3900
	TAGTAAC ATG GTTATCG GCA TAGAACT CC ATTGTGT ACC ATCTCC ATGA TGTGCAT CGG	3960
30	TATCTAT GAT TAAAAC GCGT TGATTGT ATT CTTTAG CTAA GTATTGT GCG GTAATTG CAA	4020
	TATCATT GTA TATACAAAA CCACTTG CTC GACCAG GTT AGCGTG ATGC AAACCACC AC	4080
	CTAAGTG ACA ACCATTT AAT ACTTTGC CTG ACATAATA AG ATCTGCT AAA GTTAAAG CGC	4140
35	CTCCAACA AT TGTGGC ACTA TGGCGGT GCA TATGCT TAAA TTGACC ATT C TCTTCAT CAT	4200
	TTAATCC ATA TTTCTT AGCC TCATCTT CAC TGATAAT GCC ATGTGA AGCA TGCTTA ATAG	4260
	CTTCGAC GTA ATCATATTTA TGAATTA ACA TTAATTC GTC ATCTGTT GCA ATTCTAG GTT	4320
40	GTACTAT TTG TTCTGG AGAC AATAAA TTG CATTCA AAAAG TAGCTCT GTT GTTAATTT TA	4380
	AACGCATT TG ATTGAAG GGA TGTGTG TCAT GAAATCG ATA TTGTAATA AC TTATCT GAAT	4440
	AAACAT ATGC AGTTTTT GAT GAATGTT GTT GCATATA ATC CCTCCG ATAT TCCAAAA ATT	4500
45	AAAAGAAAA CCGATT CATA TAACGA ATAT CATCAA ACGC TTGTTG CTGT TCTAATGT AA	4560
	TGTTTTT GCC AATTC TTGCC ATTAAACA AT TAGCTGG ATG ACTTGTT ATT TCTGGAT CAT	4620
50	CTGTAG CGAA TATTTCA AGT CCACCAG TTG CCATTA ACCG CTGCATTA AT TTTTAT AGT	4680
	CAAATAC ATC TAACTTT GAA TTTTTT AAAT CCCAAT GCCA GTAATATT CT GTAGTTATA A	4740
55	CGATATA ATT CTCGAATT CT GGTGTAG AAA GGCTAAG TTG TATCAG CTTT TCTGCAAG TT	4800

	TACCTGTAGA CCAGCGCTCA ATTTTCATCAG GATAGTGGAA AGTGACATAA CCCACAATGA	4920
	GTTGATCTTG GCGAGCAACA TAAATTCTAC CTTCTGGTAA TGTTGTAATT TCTAACAAAG	4980
5	CTTTATACTG ATCTTCAGCA TCTCTAAATG CGGTTAAATG CGCATCGAAA GTAAGCGCTT	5040
	TCAAATCTtC GTGTGTTAAA GGACCTTCAA TAACAAATTG CTTGTCATGA ATGTAATAAT	5100
10	CTTCGGATTG ATACGTCTTT AAATGATTCA TATTTTCAAC TCCTCAATCG ACGTCGTGAT	5160
	TGTATTAAAT TCATTATATA GAAaATTTAC AATAATTAAT ACTAGAAAAA GGATAAAAGT	5220
	AAAAATTTTG AATAATTAGA AATGTTATGT ATAATATTGA GAAAGAAAGC GTTTTCACAT	5280
15	AACAAAGGGG GAGTTTCAAA TGAAAGTCGA AGTTTATAAA GGAGCGCAAG GTAAACATAA	5340
	CCTTAAAGAT TATGAAGAAA CATATAATAC TTTTGATTGG AAAGACGTAG AACAAGCATT	5400
	TTCTTGAGT GAAACTGGAA AAATGAACAT GGCATATGAA TGCATAGATC GCCATGTAGA	5460
20	TCAAGGATTA GGGGATAAAA TAGCGTTAAA TTACAAAGAT GAGCACAGAA AAGAATCGTA	5520
	TACTTATAAA GATATGCAAC GGTTATCTAA TAAAGCAGCG AATGTTTTGT CTGAACATGC	5580
	AGAAGTTGAC AAAGGTGACA GAGTATTTAT ATTTATGTCG CGTACACCTG AACTATATTT	5640
25	TGCGTTGTTA GGTGTTTTAA AAATTGGTGC AATTGTTGGG CCGTTATTG AAGCATTAT	5700
	GGAAAAGGCA GTTGCGGATA GATTAGAGAA CAGTGAAGCT AAAGTGTTAA TTACTAATAA	5760
30	GGCATGTGTA CCTCGAGTAC CTGTAGATAA ATTACCAAAC TTGAAAAAAA TTGTTGTCGT	5820
	AGATGAGGAT GTAGAAGACA ATTACATAGA CTTCAATTAGT TTGATGGAAA CTGCTAGCGA	5880
	TGAATTTGAC ATTGAATGGT TAAAGTCGGA TGATGGTTTG ATTTTACATT ATACATCAGG	5940
35	TTCTACTGGG caACCTAAAG GtGTATTGCA TGTTCAACAA GCAATGTTAG TGCACTATAT	6000
	TTCTGGAAAA TATGTATTAG ATTTACAAGA AGATGATGTT TATTGGTGTA CAGCAGATCC	6060
	AGGTGGGGT aCAGGAACAT CTTATGGTAT TTTTGCACCA TGGTTAAATG GCGCTACAAA	6120
40	TTGTATAGCT GGTGGTCGCT TTTGCCAGA ACAGTGGTAT AGTATGATTG AAGATTTTAA	6180
	AGTGACGATT TGGTATACGG CACCAACAGC TTTAAGAATG TTAATGAGTG CTGGTGACGA	6240
	TATTGTTGAG AAATATGACT TGTCATCGTT ACGTTGATT CTATCAGTAG GTGAGCCTTT	6300
45	AAATCCTGAA GTTATAAAAT GGGCGAAAAA AGTATACGGT TTAACGGTGT TAGATACTTG	6360
	GTGGATGACA GAAACAGGTG GACATATGAT TGTTAACTAT CCAACGATGG ACGTcaAGCT	6420
	TGGCTCAATG GGCAACCAT TACCTGGTAT TCAAGCTGCA ATTATCGATG ATGCAGGGAA	6480
50	TGAATTACCA CCAAATCGAA TGGGCAACCT TGCTATAAAA AAAGGCTGGC CATCAATGAT	6540
	GTATCGTATC TGGAAGAATC CAGAAAAATA TAAATCATAT TTTATTGGAG ACTGGTATGT	6600

55

	TGATGTAATT ATGACAGCTG GTGAACGAGT TGGACCATT T GAGGTTGAGT CTAAATTGGT	6720
	TGAACACGAA GCAGTTGCCG AAGCAGGAAT TATTGGTAAA CCTGATCCGG TTCGCGGTGA	6780
5	AATAATTAAG GCGTTTGTG CACTGAGAAA AGGATATGAA CCAACAGACG AATTAAAAGA	6840
	AGAAATTCGT ATATTTGTTA AAGAAGGTTT GTCGGCACAT GCAGCACCAC GTGAAATCGA	6900
10	ATTTAAAGAT AAATTACCTA AAACACGGTC AGGTAAAATT ATGAGACGTG TATTAAAAGC	6960
	TTGGGAATTA AATTTAGATG CTGGGGATT T AAGTACAATG GAATAATGAC ATGAATGTTA	7020
	TTGAAGATT TTTTCGAAGA ATAAAGGGTG ACAACATATT TCATGTCAAT GTTTAAATAA	7080
15	TCGTTTACTT TACGATAAGC AATATAAAGA ACTGTTAAC TGTGTCATAT CATTTCTAG	7140
	AAAGCATTG AAAATGATGA CATAACAATA ATGGCATATC TTTATATTGC TTTTATTTT	7200
	TAATATGATC TTTGGAAGAT GATTATTTTA AATAATAGAA AAATATAGTT ATCAATAGTA	7260
20	TCAAGCGCTA AAAGTTGTAT AATACAAAAC TTTAATAAGT GAATTTATTG CAAAATGAA	7320
	AGCGCTAACG CGATTTAGTC GACAAGTTTT TAACAGTTCTG TTATTATATG AATGTAAGTA	7380
	AAAATTTCTT AGCTACAACT TACATATTAT AAATGCATAA ATTAACAAA AAGGGGCGAA	7440
25	AAAAGTTGAC TCATTTATCA GATTTAGATA TTGCGAATCA ATCAACACTA CAACCAATTA	7500
	AGGATATTGC TGCATCAGTA GGTATTTCTAG AGGATGCATT AGAACCTTAT GGTCAATACA	7560
30	AAGCTAAAAT CGACATTAAT AAAATTACGC CAAGAGAAAA CAAAGGGAAA GTTGTTTTAG	7620
	TAAGTGCAT GAGCCCAACA CCAGCTGGTG AAGGTAAATC AACGGTTACA GTTGGTTTAG	7680
	CTGATGCATT CCATGAGTTA AATAAAAACG TTATGGTTGC ATTAAGAGAG CCTGCTTAG	7740
35	GACCAACATT TGGTATCAAA GGTGGTGCGA CTGGTGGTGG TTATGCGCAA GTCTTACCTA	7800
	TGGAAGATAT CAACCTACAT TTCAACGGAG ATTTCCATGC GATTACAACT GCAAATAATG	7860
	CATTGTCTGC GTTTATCGAT AATCATATTC ACCAAGGTAA CGAATTAGGA ATCGATCAAA	7920
40	GACGTATTGA GTGGAAACGT GTATTAGATA TGAATGATCG TGCACCTAGA CATGTAAACG	7980
	TTGGGTTAGG TGGACCTACA AATGGTGTAC CACGTGAAGA TGGCTTTAAT ATTACAGTAG	8040
	CGTCTGAAAT TATGGCGATT TTATGTTTAA GTAGAAGTAT TAAAGACTTA AAAGATAAAA	8100
45	TTAGTCGTAT TACTATTGGT TACACTAGAG ATCGCAAGCC AGTTACAGTT GCAGATTTAA	8160
	AAGTGAAGG TGCACCTGCA ATGATTTTAA AAGATGCAAT AAAACCAAAC TTAGTACAAT	8220
50	CAATTGAAGG GACACCTGCA TTAGTTCATG GTGGACCATT TGCGAATATC GCACACGGTT	8280
	GTAAGTCAAT TTTAGCAACT GAAACAGCAC GTGATTTAGC TGATATCGTT GTAACGGAAG	8340
	CTGGATTTGG TTCAGACTTA GCGCTGAAA AATTCATGGA CATTAAAGCG CGTGAAGCAG	8400

55

	GTGTAGCGAA AGATAATTTA AAAGAAGAAA ATGTAGAAGC AGTAAAAGCA GGAATTGTTA	8520
	ATTTAGAGCG TCATGTTAAT AATATTAAAA AATTCGGTGT AGAACCGGTT GTTGCAATTA	8580
5	ATGCATTTAT ACATGATACC GATGCAGAAG TAGAATATGT AAAATCTTGG GCTAAAGAAA	8640
	ATAACGTACG AATTGCCTTA ACTGAAGTTT GGGAAAAAGG TGGTAAAGGT GGC GTTGACT	8700
10	TAGCAAATGA AGTATTAGAA GTCATTGATC AACCTAATTC ATTTAAACCT TTATATGAAT	8760
	TAGAATTACC ATTAGAGCAA AAGATTGAAA AGATTGTGAC TGAAATCTAT GGCGGTTCAA	8820
	AAGTAACGTT TAGCAGTAAA GCGCAAAAAC AATTAAAACA ATTTAAAGAA AATGGTTGGG	8880
15	ATAATTACCC AGTATGTATG GCGAAAACAC AATATTCATT CTCAGATGAT CAAACGTTGT	8940
	TAGGTGCACC ATCAGGATTT GAAATTACAA TTCGTGAATT AGAAGCGAAA ACAGGTGCAG	9000
	GATTTATCGT AGCGTTGACA GGTGCAATCA TGA CTATGCC TGGTTTACCT AAAAAACCAG	9060
20	CAGCATTAAA CATGGATGTT ACTGATGATG GTCATGCAAT TGGGTTATTC TAATAATCA	9120
	TGTCAATTGT TTAATAAAGA TAAGTAAATA GTTTAATAGA CCGGACTGTT GGAGATGCAT	9180
	TATTTACGCA GTTCGGTTTT TTGCTGTGCT AAAAATAGAT TCAATTTGGC GAATCTAACG	9240
25	ACAATGTTTG AAGGTGGTTA ATTAATGTAT ATGAAGATAA AAAGTGGGCT TGAAGAATAG	9300
	GAAAGCGATG CAATGAATAT TCCATATTAA AAAAAATTAA TAAATAGGT TGCAATATTT	9360
	AATTGGGATG CGCTACAATT AACACTAATA ATTGATATTG ATAATTATTA TCAATTAAAT	9420
30	ATAATCTTAT AGGAGTTGTT AACACATGA ACAACATCA CCCAAAATTA AGGTCTTTCT	9480
	ATTCTATTAG AAAATCAACT CTAGGCGTTG CATCGGTCAT TGTCAGTACA CTATTTTAA	9540
35	TTACTTCTCA ACATCAAGCA CAAGCAGCAG AAAATACAAA TACTTCAGAT AAAATCTCGG	9600
	AAAATCAAAA TAATAATGCA ACTACAACTC AGCCACCTAA GGATACAAAT CAAACACAAC	9660
	CTGCTACGCA ACCAGCAAAC ACTGCGAAAA ACTATCCTGC AGCGGATGAA TCACTTAAAG	9720
40	ATGCAATTAA AGATCCTGCA TTAGAAAATA AAGAACATGA TATAGGTCCA AGAGAACAAG	9780
	TCAATTTCCA GTTATTAGAT AAAACAATG AAACGCAGTA CTATCACTTT TTCAGCATCA	9840
	AAGATCCAGC AGATGTGTAT TACACTAAAA AGAAAGCAGA AGTTGAATTA GACATCAATA	9900
45	CTGCTTCAAC ATGGAAGAAG TTTGAAGTCT ATGAAAACAA TCAAAAATTG CCAGTGAGAC	9960
	TTGTATCATA TAGTCCTGTA CCAGAAGACC ATGCCTATAT TCGATTCCCA GTTTCAGATG	10020
	GCACACAAGA ATTGAAAATT GTTCTTCGA CTCAAATTGA TGATGGAGAA GAAACAAATT	10080
50	ATGATTATAC TAAATTAGTA TTTGCTAAAC CTATTTATAA CGATCCTTCA CTTGTAAAT	10140
	CAGATACAAA TGATGCAGTA GTAACGAATG ATCAATCAAG TTCAGTCGCA AGTAATCAAA	10200

55

	AGGCAACGAC CAATATGAGT CAACCTGCAC AACCAAAATC GTCAACGAAT GCAGATCAAG	10320
	CGTCAAGCCA ACCAGCTCAT GAAACAAATT CTAATGGTAA TACTAACGAT AAAACGAATG	10380
5	AGTCAAGTAA TCAGTCGGAT GTTAATCAAC AGTATCCACC AGCAGATGAA TCACTACAAG	10440
	ATGCAATTAA AAACCCGGCT ATCATCGATA AaGAACATAC AGCTGATAAT TGGCGACCAA	10500
	TTGATTTTCA AATGAAAAAT GATAAAGGTG AAAGACAGTT CTATCATTAT GCTAGTACTG	10560
10	TTGAACCAGC AACTGTCATT TTTACAAAAA CAGGACCAAT AATTGAATTA GGTTTAAAGA	10620
	CAGCTTCAAC ATGGAAGAAA TTTGAAGTTT ATGAAGGTGA CAAAAAGTTA CCAGTCGAAT	10680
	TAGTATCATA TGATTCTGAT AAAGATTATG CCTATATTCTG TTTCCAGTA TCTAATGGTA	10740
15	CGAGAGAAGT TAAAATTGTG TCATCTATTG AATATGGTGA GAACATCCAT GAAGACTATG	10800
	ATTATACGCT AATGGTCTTT GCACAGCCTA TTAATAATA CCCAGACGAC TATGTGGATG	10860
20	AAGAAACATA CAATTTACAA AAATTATTAG CTCCGATCA CAAAGCTAAA ACGTTAGAAA	10920
	GACAAGTTTA TGAATTAGAA AAATTACAAG AGAAATTGCC AGAAAAATAT AAGGCGGAAT	10980
	ATAAAAAGAA ATTAGATCAA ACTAGAGTAG AGTTAGCTGA TCAAGTTAAA TCAGCAGTGA	11040
25	CGGAATTTGA AAATGTTACA CCTACAAATG ATCAATTAAC AGATTTACAA GAAGCGCATT	11100
	TTGTTGTTTT TGAAAGTGAA GAAAATAGTG AGTCAGTTAT GGACGGCTTT GTTGAACATC	11160
	CATTCTATAC AGCAACTTTA AATGGTCAAA AATATGTAGT GATGAAAACA AAGGATGACA	11220
30	GTTACTGGAA AGATTTAATT GTAGAAGGTA AACGTGTCAC TACTGTTTCT AAAGATCCTA	11280
	AAAATAATTC TAGAACGCTG ATTTTCCCAT ATATACCTGA CAAAGCAGTT TACAATGCGA	11340
	TTGTTAAAGT CGTTGTGGCA AACATTGGTT ATGAAGGTCA ATATCATGTC AGAATTATAA	11400
35	ATCAGGATAT CAATACAAA GATGATGATA CATCACAAA TAACACGAGT GAACCGCTAA	11460
	ATGTACAAAC AGGACAAGAA GGTAAGGTTG CTGATACAGA TGTAGCTGAA AATAGCAGCA	11520
40	CTGCAACAAA TCCTAAAGAT GCGTCTGATA AAGCAGATGT GATAGAACCA GAGTCTGACG	11580
	TGGTTAAAGA TGCTGATAAT AATATTGATA AAGATGTGCA ACATGATGTT GATCATTAT	11640
	CCGATATGTC GGATAATAAT CACTTCGATA AATATGATTT AAAAGAAATG GATACTCAAA	11700
45	TTGCCAAAGA TACTGATAGA AATGTGGATA AAGATGCCGA TAATAGCGTT GGTATGTCAT	11760
	CTAATGTGCGA TACTGATAAA GACTCTAATA AAAATAAAGA CAAAGTCATA CAGCTGAATC	11820
	ATATTGCCGA TAAAAATAAT CATACTGGAA AAGCAGCAAA GCTTGACGTA GTGAAACAAA	11880
50	ATTATAATAA TACAGACAAA GTTACTGACA AAAAAACAAC TGAACATCTG CCGAGTGATA	11940
	TTCATAAAAC TGTAAGATAA ACAGTGAAAA CAAAAGAAAA AGCCGGCACA CCATCGAAAG	12000

55

CATGGTGGGG CTTATATGCG TTATTAGGTA TGTTAGCTTT ATTCATTCCT AAATTCAGAA 12120

AAGAATCTAA ATAATTACT AAATATAGCA TATGTATGAT TAACTTTGTA GAC 12173

5

(2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1316 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

CAACATTAAT ATTGATATTA AATCTTCCTG GATAACGTGC TTGTTGAGT GATAAGTATG 60

CACGCACTTG ACTTAACTCT TTATCTAAAG TAATCGTATG TTGCTTAGAG CCTTGTAAT 120

20

TCGCTCTGAA AAAATAACTC AATTCTAATA GTAACCTCG TGCCTTTTCG CTATTTATTC 180

TAACTAAAGC TGAGATCGTG TTAATTGAAT TGAAGAAAA ATGTGGACTC ACTTGTGCCT 240

GTAATGACTT AATCTCAGCA TCTTTCAATA ACTTACTTTG CGTTTCGGCT TCACCAAGTT 300

25

CAATTTGGCT ACTAAAAATA TTTGCCAATC CTTCTGCAAG TTGACGTTCC ACAAAGTTA 360

AATCATTAGG GTTTGTAAAA TACATCTTCA ATGTACCGAC GATAGAACCA TGCATCTCAA 420

30

GTGGTATCAC GATAGCTGCT CTAAGCGGGC AATTCGGATG ACTACAACCA ATCTCTTCTT 480

TAGTATGAAC TTCTTTCAAC TTTCTGATT TCAATACATC TTTAGACAGA CTTGTTAATA 540

TTTCATTGTG TGGTATGTGA TGATCACTAC CTGCACCTAC ATGCGATAAG ATTTCAATTT 600

35

TGCTTGTAAT TGCTACGGCA GATACTTTCA TTAAATTTT AATAATCATC GCAATTTGCT 660

GTGCCGATTC TCTATTCAAT CCTTCTTTAA AATACGGCAA TGTCTGGTTC ATCAATTGCA 720

GTAATCATG TGTGTAACA GCCTTCATT GCTCCTCTG CTTAATGTT GAAATGATAA 780

40

TAGACATAAA AATCGCCGTA CCAACGCTAT TAACAATAAT CATTGGTAGT GCAATTAATG 840

ATATGAGGTC AACCGCATAT GCTTTGTCGT GGGAAAATGT TAAAATGCTC AACATTTGAA 900

TCATTTCCAT AACAATTCCA ATCATGGCAC TTTTCGCAAT ACTCGGTAA CGCTTGCCTC 960

45

TTTGAGCTTG TAAGCCAAAA TAACCAGCAA TTATACCAAT AAATATAGAT GAGATAAGAT 1020

AAACTTGTGC ATCCGCCCCA CCCATATACA CTCTGAAAAT ACCTGAAATA ACGCCAACAA 1080

ATAGACCTAC AAAAGGGCCA CCAACTAATC CTGCGACACC TATCGTTAAT ACACGTGTGT 1140

50

TAGCTAAAGA TACATCATCA TCTAAACGGA AGTACACACT TCCTGACAAA CTATGTTGAT 1200

GATCGATGAC GATACAGTT AAATTAGACA TTAAGGCAA CAACTGAAA ATAATACATA 1260

55

(2) INFORMATION FOR SEQ ID NO: 312:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7972 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

	TATAAATATT ATTTTATTAT CGTTTATCGC TATTTGTGTT ACATAATCAA AACCATAAAT	60
15	TCTTACTCAT TCAGATTAC CCAATATTTT TACTTTTATA ATGTAATGCG TTTTATCCAA	120
	GTTATTTTTT AAAAATAAAT ATTGAATTnG GGGCTGnTTT CATGTCATTA AGAGATGAAG	180
	CATTGGAAT GCACAAACGT AATCAAGGTA AATTAGAAGT TAAACCAAAT GTAAAAGTTA	240
20	CTAATAAAGA GGAATTAAGT TTAGCATACT CACCTGGCGT TGCTGAACCG TGTAAGATA	300
	TTTATGAAGA TAAAAGAAAA GTATATGATT ACACAATTAA AGGAAATACA GTTGCAGTTA	360
	TTACTGATGG AACAGCGGTA TTAGGTTTAG GTAACATTGG ACCTGAAGCA AGTATTCCTG	420
25	TAATGGAAGG TAAAGCAGTA TTATTCAAAA GCTTCGCTGG TATCAATGGG GTGCCTATTG	480
	CGTTAAATAC AACTGATACC GAAGAAaTCA TTA AACAGT TAAGTTGTTA GAACCTAATT	540
30	ATGGTGGTAT TAATTTAGAG GATATTCGG CACCACGTtG TTTTGAAATT GAAGAACGAT	600
	TGAAAAAAGA AACTAATATT CCGGTATTCC ATGACGATCA ACATGGTACA GCAATTGTAA	660
	CATTGGCAGG TTTGGTAAAT GCATTGAGAG TTGTTAACAA AGATATTGCT AAAATAAAAG	720
35	TTGTACTAAA TGGTGCTGGT GCAGCAGGAA TAGCCATTGT TAAATTACTA TACGCGTATG	780
	GTGTAAGAAA TATGTTATG TGTGACTCAA GAGGCGCAAT TTTTGAAGGA CGTTCATATG	840
	GTATGAATCC TACGAAAGAT GTTGTAGCAA AATGGACAAA TAAAGATAAG ATTGAAGGGT	900
40	CTTTAGAAGA AGTCGTAAAA GACGCAGATG TATTTATCGG GGTTCCTGTA GCTAATGCGC	960
	TGTCACAAGA TATGGTTAAG AGTATGGCAG ATAATCCAAT TATATTTGCA ATGGCTAATC	1020
	CAAATCCTGA AATAATACCT GATGATGCCA AAGCGCAGG TGCACGAGTT GTTGGTACAG	1080
45	GACGTTCAGA CTATCCTAAC CAAATTAATA ATGTATTAGC TTTCCCTGGT ATTTTTAGAG	1140
	GTGCATTAGA GGTGAAGCT ACACATATAA ATGAAGAAAT GAAAAAGGCA GCTGTAGAAG	1200
50	CGATTGCTGA TTTAATCGAT AGTTCTGAAT TAAATGAAGA CTA CTGTATC CCAGGACCGT	1260
	TTGATAAACG TGTAGCGCCA TCAGTTGCTC GTAATGTTGC TAAAGCGGCA ATGGAATCTG	1320
	GAGTAGCTAG GATTGAAGTT GATCCGCAAG ATGTGTATGA TAAAACAATG AAACCTTACAG	1380

	ATTAAATGA TGAAATGAA AGTTTATGAT AAACATTCAA CAGTCAAACG AATATAAATC	1500
	AAATAAATTT AAACCCGTTT TTAAGTGGTC AAGTTCAGTT TAAGGCTCTA AATGGTTAGA	1560
5	ACAGAGGTTA TTTGGAGGTT TTCCTATGTT TAAAGATTTT TTTAATCGAA CAAAGAAAAA	1620
	GAAATATCTT ACAGTACAAG ACTCTAAAAA TAATGATGTG CCTGCAGGTA TTATGACTAA	1680
10	GTGTCCAAAG TGTAAGAAAA TTATGTACAC AAAAGAATTA GCTGAAAATT TAAATGTGTG	1740
	CTTTAATTGT GATCATCATA TTGCTTTAAC TGCCTATAAA CGTATAGAAG CAATTTCTGA	1800
	TGAAGGATCA TTTACAGAAT TCGATAAGGG AATGACCTCT GCGAATCCAT TAGATTTTCC	1860
15	aAGTTATTTA GAAAAAATTG AAAAGGACCA ACAAAGACA GGTCTTAAAG AAGCAGTTGT	1920
	GACTGGTACA GCACAACTAG ATGGTATGAA ATTTGGCGTT GCTGTCTGAG ATTCACGTTT	1980
	TAGAATGGGA AGTATGGGAT CGGTTATCGG TGAAAAGATA TGTCGCATCA TTGATTACTG	2040
20	CACTGAGAAC CGTTTACCAT TTATCTTTTT CTCTGCAAGT GGTGGTGCAC GTATGCAAGA	2100
	AGGTATTATT TCCTTGATGC AAATGGGTAA AACCAGTGTA TCTTTAAAC GTCATTCTGA	2160
	CGCTGGACTA TTATATATAT CATATTTAAC ACATCCAAC ACTGGTGGTG TATCTGCAAG	2220
25	TTTTGCATCA GTTGGTGATA TAAATTTAAG TGAGCCAAAA GCGTTGATAG GTTTTGCAGG	2280
	TCGTGAGTT ATTGAACAGA CAATAAACGA AAAATTGCCA GATGATTTCC AACTGCAGA	2340
30	ATTTTTATTA GAGCATGGAC AATTGGATAA AGTTGTACAT CGTAATGATA TGCCTCAAAC	2400
	ATTGTCTGAA ATTCTAAAAA TCCATCAAGA GGTGACTAAA TAATGTTAGA TTTTGAAAAA	2460
	CCACTTTTTG AAATTCGAAA TAAAATTGAA TCTTTAAAG AATCTCAAGA TAAAAATGAT	2520
35	GTGGATTTAC AAGAAGAAAT TGACATGCTT GAAGCGTcAT TGGAACGAGA AACTAAAAAA	2580
	ATATATACAA ATCTAAAACC ATGGGATCGT GTGCAAATTG CGCGTTTGCA AGAAAGACCT	2640
	ACGA ^Δ CCCTAG ATTaTATTCC ATATATCTTT GATTCGTTTA TGGAAGTACA TGGTGATCGT	2700
40	AATTTTAGAG ATGATCCAGC AATGATTGGT GGTATTGGCT TTTTAAATGG TCGTGCTGTT	2760
	ACAGTTaATTG GACAACAACG TGGAAAAGAT ACAAAGATA ATATTTATCG AAATTTTGGT	2820
	ATGGCGCATC CAGAAGGTTA TCGAAAAGCA TTACGTTTAA TGAAACAAGC TGAAAAATTC	2880
45	AATCGTCCTA TCTTTACATT TATAGATACA AAAGGTGCAT ATCCTGGTAA AGCTGCTGAA	2940
	GAACGTGGAC AAAGTGAATC TATCGCAACA AATTTGATTG AGATGGCTTC ATTAAAAGTA	3000
50	CCAGTTATTG CGATTGTCAT TGGTGAAGGT GGCAGTGGAG GTGCTCTAGG TATTGGTATT	3060
	GCCAATAAAG TATTGATGTT AGAGAATAGT ACTTACTCTG TTATATCTCC TGAAGGTGCA	3120
	GCGGCATTAT TATGGAAAGA CAGTAATTTG GCTAAAATTG CAGCTGAAAC AATGAAAATT	3180

55

GGTGCACATA AAGATATTGA ACAGCAAGCT TTAGCTATTA AATCAGCGTT TGTTCACAG 3300
 TTAGATTCAC TTGAGTCATT ATCACGTGAT GAAATTGCTA ATGATCGCTT TGAAAAATTC 3360
 5 AGAAATATCG GTTCTTATAT AGAATAATCA ACTTGAGCAT TTTTATGTTA AATCGATACT 3420
 GGGTTTTACC ATAAATTGAA GTACATTAAA ACAATAATTT AATATTTAGA TACTGAATTT 3480
 TAACTAAGAT TAGTAGTCAA AATTGTGGCT ACTAATCTTT TTTAATTAA GTTAAAATAA 3540
 10 AATTCAATAT TTA AACGTT TACATCAATT CAATACATTA GTTTTGATGG AATGACATAT 3600
 CAATTTGTGG TAATTTAGAG TTAAAGATAA ATCAGTTATA GAAAGGTATG TCGTCATGAA 3660
 15 GAAAATTGCA GTTTTAACTA GTGGTGGAGA TTCACCTGGA ATGAATGCTG CCGTAAGAGC 3720
 AGTTGTTTCGT ACAGCAATTT ACAATGAAAT TGAAGTTTAT GGTGTGTATC ATGGTTACCA 3780
 AGGATTGTTA AATGATGATA TTCATAAACT TGAATTAGGA TCAGTTGGGG ATACGATTCA 3840
 20 GCGTGGAGGT ACATTCTTGT ATTACAGCAAG ATGTCCAGAG TTTAAGGAGC AAGAAGTACG 3900
 TAAAGTTGCA ATCGAAAAC TACGTAAAAG AGGGATTGAG GGCCTTGTA TATTGGTGG 3960
 TGACGGTAGT TATCGCGGTG CACAACGCAT CAGTGAGGAA TGTAAAGAAA TTCAACTAT 4020
 25 CCGTATTCCT GGTACGATTG ACAATGATAT CAATGGTACT GATTTTACAA TTGGATTTGA 4080
 CACAGCATT AATACGATTA TTGGCTTAGT CGACAAAATT AGAGATACTG CGTCAAGTCA 4140
 CGCACGAACA TTTATCATTG AAGCAATGGG CCGTGATTGT GGAGATCTAG CATTATGGGC 4200
 30 TGGATTATCA GTTGGTGCTG AGACAATTGT AGTTCCAGAA GTGAAAACAG ATATTAAAGA 4260
 AATAGCTGAT AAAATTGAAC AAGGTATTAA ACGTGGTAAG AAACACTCAA TCGTTCTTGT 4320
 AGCAGAAGGT TGTATGACTG CGCAAGATTG TCAAAAAGAA TTATCACAAT ACATCAATGT 4380
 35 TGATAATAGA GTGTCTGTGT TAGGTCACGT TCAACGTGGT GGTAGCCCAA CAGGTGCGGA 4440
 TAGAGTTTGA GCATCACGTT TAGGTGGATA TGCGGTAGAC TTATTAATGC AAGGTGAAAC 4500
 40 AGCTAAGGGT GTTGGGAATTA AGACAATAA AATTGTAGCA ACATCTTTTG ATGAAATTTT 4560
 TGATGGTAAA GATCATAAAT TTGATTATAG TCTATATGAA CTTGCTAACA AGTTATCTAT 4620
 ATAAGATTTC AGGAGGAATT ATAAAATGAG AAAAATAA ATTGTATGTA CAATTGGACC 4680
 45 AGCTTCAGAA TCAGAAGAAA TGATTGAGAA ATTAATCAAT GCTGGTATGA ACGTTGCACG 4740
 ATTAACTTT TCACATGGTA GTCATGAAGA GCATAAAGGT AGAATTGATA CAATTCGTAA 4800
 AGTAGCTAAA AGATTAGACA AAATTGAGC AATTTTATTA GATACAAAAG GTCCAGAAAT 4860
 50 TCGTACGCAT AATATGAAAG ACGGTATCAT TGAACCTGAA CGTGGAACG AAGTTATTGT 4920
 TAGCATGAAT GAAGTTGAAG GAACACCTGA AAAGTTCTCA GTAACATATG AAAACTTAAT 4980

55

	TAAAGATATT GACCATGCTA AAAAAGAAGT TAAATGTGAT ATTTTAACT CTGGTGAGCT	5100
	TAAAAACAAA AAAGGTGTTA ACTTACCTGG CGTAAGAGTA AGTTTACCTG GTATTACAGA	5160
5	AAAAGATGCT GAAGATATCC GTTTCGGTAT TAAAGAAAAT GTTGACTTCA TTGCAGCAAG	5220
	TTTCGTACGT CGTCCTAGTG ATGTTTTAGA AATTCGTGAA ATTTTAGAAG AACAAAAAGC	5280
10	TAACATTTC A GTATTCCCTA AAATTGAAAA CCAAGAAGGT ATTGATAATA TTGCGGAAAT	5340
	TCTTGAAGTG TCTGATGGTT TAATGGTTGC ACGTGGTGAC ATGGGTGTTG AAATTCCACC	5400
	TGAAAAAGTA CCAATGGTTC AAAAAGATTT AATCAGACAA TGTAACAAAT TAGGTAAACC	5460
15	AGTTATTACA GCTACACAAA TGTTAGATTC TATGCAACGT AACCCACGTG CTACACGTGC	5520
	AGAAGCTAGT GACGTTGCCA ACGCAATCTA TGATGGTACA GATGCAGTAA TGTTATCTGG	5580
	TGAAACTGCT GCTGGTTTAT ATCCTGAAGA AGCTGTTAAA ACAATGAGAA ATATTGCTGT	5640
20	ATCAGCTGAA GCAGCCCAAG ATTACAAAAA GTTATTGTCA GATCGTACTA AATTAGTTGA	5700
	AAC TTCATTA GTGAATGCTA TCGGTATTTT GGTGTCACAT ACAGCTTTAA ACTTAAATGT	5760
	TAAAGCAATT GTAGCTGCTA CTGAAAGTGG TTCAACGGCA CGTACTATCT CCAAATATCG	5820
25	TCCACATTCA GACATTATG CGGTGACTCC AAGTGAAGAA ACTGCACGTC AATGTTCAAT	5880
	TGTTTGGGGA GTTCAACCTG TAGTTAAAAA AGGACGTAAG AGTACAGATG CATTGTTAAA	5940
	CAATGCAGTT GCAACAGCTG TTGAACTGG TAGAGTATCT AATGGTGATT TAATCATTAT	6000
30	TACTGCTGGT GTACCAACTG GTGAACTGG AACTACTAAT ATGATGAAAA TCCACCTAGT	6060
	TGGTGACGAA ATTGCTAATG GTCAAGGTAT TGGACGTGGA TCAGTTGTTG GTACTACGTT	6120
	AGTTGCTGAA ACTGTTAAAG ATTTAGAAGG TAAAGATTTA TCTGACAAAG TTATCGTTAC	6180
35	TAACTCAATC GATGAAACGT TTGTACCTTA TGTAAGAAAA GCTTTAGGCT TAATTACAGA	6240
	AGAAATGGT ATTACATCAC CAAGTGCAAT TGTTGGTTTA GAAAAAGGTA TTCCAACAGT	6300
40	TGTAGGTGTA GAAAAAGCTG TTAAAAACAT AAGCAATAAC ATGTTAGTTA CGATTGATGC	6360
	TGCTCAAGGT AAAATCTTTG AAGGATATGC AAACGTACTA TAATTTATAA AAAACGTCT	6420
	TTCCATTTAT CAACAATGGA AAGGCGTTTT TTGGTTCATC TGGTATTTTA TGACGTAATT	6480
45	AATAGGTTAT TTGATAATGA TAGTGTATGA ATGGCAATCT ATATAAATGT TTATATCTTT	6540
	TATACATGTA CATTATCACC TTCAAACCTT CACTCATATT ACTTTGGAAA TTTATTATAA	6600
	AATAGAAGTA TGGATGTATT TCTGAAATGA TACATTATTA AATAGATGAG AAAGTAAAG	6660
50	TTTTGAGCCA AGTACGCAAT TTAATATTAT AAGTTGCATA TAAACAGGA TGGGACATAA	6720
	ATCCCTAAAA AAACAGCAGT AAGATAATTT TCAATTAGAA AATATCTTAC TGCTGTTCTC	6780

55

1 CTTCTGACTG GCACTGCTCC CTCAGGAGTC TCGCCATTAA TACTACGTAT TAACATGTAA 6900
 5 TTTTACTTTT ACATACTTTA AAAAAATAAA ACACTTTGCC CAACTTGCAC ATAAATGTAA 6960
 AATTCAATAA AATGAATTTT CTGTGTTGGG TCCCTTCGTA TAATTTAATA AATACCACTA 7020
 AACTAAATTA ACGAGGTGCC TTATGTATAA AATTTATAAC ATGACCCAAC TTACACTACC 7080
 10 AATAGAAACC TCTGTTAGAA TTCCTCAAAA TGATATTTTCG CGATATGTTA ATGAAATTGT 7140
 TGAAACGATA CCTGATAGCG AATTCGATGA ATTCAGACAT CATCGTGGCG CAACATCCTA 7200
 TCATCCAAAA ATGATGTAA AAATCATCTT ATATGCATAT ACTCAATCTG TATTTTCTGG 7260
 15 TCGTAGAATA GAATAATTAC TTCATGACAG TATTCGAATG ATGTGGTTAG CTCAAGATCA 7320
 AACACCTTCT TATAAACTA TTAATCGTTT TAGAGTGAAT CCTAATACTG ATGCGTTAAT 7380
 TGAATCTTTA TTTATTCAGT TCCATAGTCA ATGTTTAAAG CAAAATCTTA TTGATAATAA 7440
 20 TTCAATTTT ATTGATGGTA CAAAAGTAGA AGCTAATGCC AATAGATATA CATTTGTGTG 7500
 GAAGAAAAGT ATTCAAATC ACGAATCGAA ATTGAACGAA AATTCAAAAA CATTATATCG 7560
 TGACTTAGTT GAAGAAAAA TAATACCAGA GATAAAAGAA GATGGAGATA GCGATTTAAC 7620
 25 AATAGAAGAA ATAGATTTAA TTGGTAGTCA TTTAGATAAA GAAATCGAAG ATTTAAATCA 7680
 TTCTATTGAG AACGAAGATT GTGCTCAAAT TAGAAAACAG ACCCGTAAAA AAATAACTGA 7740
 30 GATTAAGAAG TTCAAAAAGA AATTTGATGA TTATTCCGAA AGAAAAATA AATATGAAGA 7800
 ACAAAAATCG ATTCTTAAAG ATAGAAATAG TTTTCTAAA ACTGATCTGA TCATGATGCA 7860
 ACTTTTATGA GAATGAAGGA AGACCATATG AAAAATGGCC AACTTAAGCC AGGATACAAT 7920
 35 TTACAAATAG CGACAAATTC TCAAAAATGT TTTATCCTAT GACCTATTTT AA 7972

(2) INFORMATION FOR SEQ ID NO: 313:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3175 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

ATTTTTTAGT TAATTGTCTT TCTTAAATA ATTTTAGCTT TCATTAAATT AAACAATTTT 60
 50 ACAAGCTTGG AACACCAATC AAAATCCTAA GTTCTAAAAT GCAATATTAG TAGTCGTTGA 120
 CTGAATGAAC ATATGCTTAT AATATTTTTT TGCAATGCTA GTCAAGTTGA TTTATGCTCA 180
 CAAGGATATG CGATTTATAT TTTCTTACAA CAATGAAAAT GCCTGATACA ATGCGATCCT 240

	AATTAAATTA TTGTTGTTTT AACCATTGTA CTATCGTCGG AATCATAAAT CCTGTTGGCC	360
	CTTTTGGACC ATTATATGAA GCTTTATTAG TCGTTGCTGG ACCTGCAATA TCGAAATGAA	420
5	TGTGAGGTGT TTGACCACTA AAATGTGTTA CAAAACCTGC CGCAAATAGC GCTTTACCTT	480
	GTCCATTTCGT ATGGTTAACT AAATCAGCGA TATCACTGTG TTTAATACTT GCACGTTCCG	540
	TTGCAGTAAT CGGTAATTCA AATACCATTT CATCGACTTC AGAACTTATT TGTAATATAT	600
10	CGTTTAATAT CACTTTACTA TTCGATTCAA ATGCAGCAGC TTTATCATCG CCTAGTGCAa	660
	CAATTGCTGC ACCCGTTAAT GTAGCAAAGT CCATAATCAC ACTAGGCTGA TATTGATTTG	720
	CATAAAACAC AGCATCTGCA AGGACTAATC TACCTTCAGC GTCTGTATTC ATTACTTCTA	780
15	CAGTTTCACC ACTTAATGCT GTAAATACAT CATCTGGcTT CATTGATGCT TCATTTATCA	840
	TATTTTCAGC ACACGCAAGC ACTCCGACAA TATTTACAGG CAGTTGTAAA CGActAGCCG	900
20	CTTCAATGAT ACCAACGACA TTCGCAGCGC CACACATGTC AAAC TTCATT GTAGCCATGC	960
	CATTCTTCGT TTTAATACTA TAACCACCAG AATCATACGT TATACCTTTA CCAACTAAGG	1020
	CAATTGGTGC TTCATCTTTG TCTTTGCCAT TATATGTGAT GGTTACTAAa CTCGGTTTAT	1080
25	GCTTACTACC TTTACCTACT GCTTGTAATA GTCCGAATCC TTCAGAACT AAAGTGT CAT	1140
	AATCTTTAAC ATCTACTTTG ACCTTTGTAT TTTTAAATG ATTAACAATA TCTTCTGcAA	1200
	ATGTTTGTGG TGtTAATACA TTCGGTGGcA TATTACTAAA GTCTCyTGcC AAATTAATGG	1260
30	aTTGGcCAAT ACTGATACCC TCATGAATAA AATCTAATTC AATTAATGAT TCACTAATAA	1320
	GGkTTAAATT CGTCTTAAAC GGtGCCTTCT kACTTgrTTT ATAATGATCG aATyCATATG	1380
	kTGcACGCTC ACTTTGGAAT ACCGCATGCC ATTAATACAT CTGgATAACT GgATCcATAT	1440
35	TTTGgAnATA AATGaATCCa TAAGTAAATA CGTATCTTCa ATGTGTTCTG ACTTTATGTA	1500
	TTGGAAAAGA TGTCcCCaTA TTTTCAACAT ATCTTGATAA lACGTGTCTT TAAGTTTCCT	1560
40	AAACCAACTG TAATTAAACG ATATGTTTGA TCTTGACAT CAAATGCAGT TGTATAAATC	1620
	TTCCCAACTT TACTACCAAT AATATGTTGA TGTTTtagTC TTTCAAGTGA TTCTGTAATA	1680
	TCGATATGAT TAAAACTAAT GCGCTCTAAC TGATTTAAAT GTTCTGGTAT ACCAATAATC	1740
45	AATGTATTTA TTTCGTTGCT TAGTGTGTTA TTTAATTTAA AATTCATAAT GTACCTCCTT	1800
	AGATTTGATA TGTTACTCAC TTCAAATTGT ACAACAATAA AGCCCTcAGT GACACTGAAG	1860
	GACTTTATAA ATGAATAATT TAATTTTATG TGTTTtagCTG AAAC TTTATT TAAAGCCTAG	1920
50	AATTTACCTT TTTTGAATGC TAAACCGATA CCACCGATTT TGAATACCGC ACGTGTATCA	1980
	ATCACTTTTT TCATGAATGC TGCTTTTTTA CCAGCGATAG GTTTACCAA TACCATACCT	2040

55

GTTGATTAC CGTTTAAGAT GCGTTTAATG TTTTAGCAA CACTTTCACC TTGTTGCATT 2160
 GCAATTTGTG CTGTAGTTGG TAATGGACGT TCTTCTCCAG CTGGGATAAA CGCTGAACAG 2220
 5 TCACCAATAA CAAAAATGTT GTCGTAACCA TTGATTGTTA AATCTTGCTT TGTAACGATA 2280
 CGTCCACGTT TAACGCCTTC AAATGATTCT TCCATTAATT TACTACCACG TACACCAGCT 2340
 GCCCATACTG AAGTACCTGC ATTAAATTGT TGTTTTTTAC CATCTACTTC AACTACAAAA 2400
 10 CCTTTTTTCGT TACAAGCAAC GATTGGTGTA GCAATTTTAA ATTCAACACC GCGGTCTTCT 2460
 AAGTAGCTAA CTGCGTGGTT AACTAATTCT TCTGAGAACA TTGGTAACAT TTTAGGTGCT 2520
 GCTTCAACAC AAGTGATTTT AACTTTATTT TGATCCACAC CATATTTGCT ACATAATTCA 2580
 15 GGAATTCTGT CTGTTAATTC ACCTAAGAAT TCAACACCAG TGAATCCAGC ACCACCAACT 2640
 AAGATAGATA AATCGTTATC ATCTTTTTCT TTTGATGCTG CATAGTTAGC AAATTTGTCT 2700
 20 TCGATATGAC GTGATAATTC ACGTGCTGTG ATAACATTTT CAATTTGGAA AGCATGATCT 2760
 TTCATACCTT CGATGCCGAA TGTTCACCTA ACGAAACCTA ATGCTACTAC TAAAATATCA 2820
 AAGTCATAAA TACCTTGATT TGTCTTACC TTTTAGCAT CACGGTCAAT TTTTGTTACT 2880
 25 TCTGCTTGAA CAAAGTTCAC TTTGTCTTTC TTCAAGACAC TTTCCACAGG ATATAATACA 2940
 TCTTCATAGT TTAGTGATCC TGCTGATGCT TCATGTAACC ATGTTGCTTC ATAGTGATAT 3000
 TCATTTTTAT TAATAAGCGT AATTTCTGCT TCTTCTGTTG ATATCGCTTT TTGCAATTTA 3060
 30 GTTACAGTTT GTAAACCTGC ATAACCAGCA CCAAGTACAA GTACTTTTTT ACGATCTTGA 3120
 GCCATTTAAT TnCACCTAAG CnTTCATATT TTTTAAACCA AATGCTGATA ATTAC 3175

(2) INFORMATION FOR SEQ ID NO: 314:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 702 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:

45 CATCATTATT AAAGATTTTC AATCAATACA gAATCACAAT ACGTACGCAT TGTGCACGAT 60
 AAAAATACAG ATGTGTATAT TAACTATGAA CTACAAGAGC AACTAACGAA CAAAGCTTAC 120
 ATTGGTGATC ATATTTATGT TGAAGGGATA TGGCTCGAAG TACAAGCTGA TGGTTTAAAT 180
 50 GTATTGAGTC AGAATACAGT GGCATCGTCA TTAATTCGCT TAACACAAGA GATGCCACAT 240
 GCACAGGCAG ATGATTACAA TACGTACCAT CGTTCGCCAA GGATTATTCA CCGTGAACCG 300

TGGCGTTCCA TTATACCGCC ATTAGTAATG ATTGCTTTAA CTGTTGTCAT CTTTTTAGTG 420
 AGACCAATTG GTATTTATAT TTTAATGATG ATTGGTATGA GTACAGTAAC GATAGTATTT 480
 5 GGTATTACAA CGTATTTCTC TGAAAAGAAA AAGTATAACA AAGATGTTGA AAAACGAGAG 540
 AAAGATTACA AAGCTTATTT GGATAATAAA TCTAAAGAAA TTAATAAAGC GATTAAAGCA 600
 CAACGTTTTA GTTTGAATTA CCATTATCCA ACGGTGCTG AAATTAAAGA TATCGTTGAA 660
 10 ACGAAAGCAC CAAGAATATA TGAaAAAACC ATCGGCATCA TC 702

(2) INFORMATION FOR SEQ ID NO: 315:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4121 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

TATGTTCCGA CAACGAAACA AAGTGTAATT ACAAGAGCAA AGATAACTTT GAATGTTTGT 60
 25 AAACGTCCAT CTTTACCTTC AGTTAAATGC ATGAACATTA ATAATTGAAG TCCTGCTTGG 120
 ACGAATGCAA AGCCAAAGAT AATTGTCAAC TTCGCGTGGA ATGTTAATGA CGTGATAGT 180
 GTTACGTAAA CTGCTAAAAG CGTTAATACG ATAGATGCGA TAAATCCTAC AGTATGTTTC 240
 30 ATTATTGTAC TCATCCGCTA TACACCATCC CTATCATATA TACGGCAGTA AAGATGAAAA 300
 CCCAAACAAC ATCTAAGAAG TGCCAGTATA AACTTACTAT AAATAATTTT GGCGCATTAT 360
 ATTTGTCTAA TCCGCGTCGT TGGATTGGA TTAATAAACA AATGGCCCAA ACGATACCTA 420
 35 GCGATACGTG ACAACCATGC GTTCCTAATA GGATAAGAA ACTAGACCAG TAAGAACCAA 480
 TTGTTGGGTT AACGCCTTCT GATGCATAGT GTGCGAATTC ATAAATTTTG AATCCAACAA 540
 40 AGACTAAACC TAAAAGTAAC GTAATGATCA TCCAAAACAT CATTAACTTT TGTTTTCTT 600
 GGCGCATGTA GTAAATAGCA ATACCACATG TGTAAGAACT GAATAATAAT GCAAACGTCA 660
 TTATTAAAC AAGAGGCAAT TCAAATAACT CAGTAGTCAT TTTACCTGCA TAATCGCCAC 720
 45 CATGTTGCAA AGTTAATAGT GTCGCAAATA GGGTACCGAA TAACGCAAAT TCGGCTGTAA 780
 TGAAAATCCA AAAGCCAAGC TTATTTAATT CGCCTTCATG TGTGCGTGAA TCAATAGTGT 840
 TTGTATCATG ACTCATGACT TACAGCCTCC CTTTCTTTAA TACGCGCTTC TCTTAATCTA 900
 50 GCTTCAGTTT CTGCAACTTC AGCAGCAGGG ATATGATATC CGTGATCGAT TTGGAAACTG 960
 CGATAAATCA TAGTACCAAA AATACCGAAT AAACAAATTA ATGCTGGAAT TACAGTTTCG 1020

55

	TTGTTTGGCA TATGAATGTC TTTGTAATTA TGGTTGTCTA AGTAATGACG ACCATGTTCT	1140
	TTCATATCAA CAAATGTGTC GTAGTCATTC CAATCTGGTG TAATGGCAAA GTTGTATTTA	1200
5	GGTGAATTG CTGATGCAGT AGTCCACTCT AGAGTACGAC CAAGGCCATC CCAGTTATCT	1260
	CCAGTTGCTT CACGTGGAGA TTTGAAGTGA CTGTATACGA TACTAACAAAC AAGGAATAAG	1320
10	AATCCGATTG CCATTAATAA TGCACCGATA GTTGAAATTA AGTTTAATAA GAACCAACCA	1380
	TCTGATGGCA TATAAGTGTA TAAACGACGT GGCATACCAT CTAATCCAAG AATGAATTGT	1440
	GGTAAGAAAC AAACGTTAAA TCCGATCATG AAGAACCAGA AGCACCATT TTTTAATGTT	1500
15	TCGTTTAATT TGTAACCCAT CATCTTTGGA TACCAGAAGA TTAAACCAGC TAAGCAGGCA	1560
	AATACAACAC CAGTAACCAA TGTATAGTGG AAGTGAGCTA CTAAGAAGTA CGTATTGTGA	1620
	TATTGATAGT CAGCTGATGC CATTGCTAAC ATTACACCCG TAACACCACC TAAAAGGAAG	1680
20	TTAGGGATAA ATGCTAATGA GAATAGCATT GGTGACTCAA ATGTAATACG TCCTTTATAT	1740
	AATGTTAATA ACCAGTTAAA CAATTTTACA CCAGTTGGAA TACCAATCAG CATTGTTGAA	1800
	ATTGAGAAGA ATGAGTTGAT TAACGCACCA TTACCCATTG TGAAGAAATG GTGAACCCAA	1860
25	ACTAAGAAAC TAAGGAACGC GATACCGGCA GTTGCCCATTA CCATACTTTG ATGTCCGAAT	1920
	AAACGCTTAC GAGCGAATGT CGGGATAATT TCTGAGTAAA TACCAAATGC TGGAAGGATA	1980
30	ACGATATAAA CTTCAGGGTG CCCCCATACC CAGAAGAAGT TAGCCCAAAG CATTGGCATA	2040
	CCGCCATGTG CAACTGTGAA GAATGCTGTG TCAAATATTC TATCAGTTGT CATTAAATGCT	2100
	AACGCTACTG TTAAAGGAGG GAAAGCAAGA ATAACAATTA ATGTAGTAAT AAATGTTGTT	2160
35	ACTGTAAACA TTGGCATTG CATAAACTTC ATAGTTGGTG TTTTACATCT TAAAATTGTT	2220
	ACAAAGAAGT TGATACCTGT AGCTAAGGTA CCAAGCCCTG AAATTTGTAT AGCTATTAAAG	2280
	TAATAGTTAA CACCCGGACC AGGACTGAAT TCACCTGCTA GTGGCGCATA GTTTGTCCAA	2340
40	CCAGCTGCTG GTGAACCACC AATAATAAAT GACAGGTTGA ATAAATCAT ACCTGCAAAG	2400
	AATAGCCAGA AACTTACGTT GTTTAATACT GGAATGCAA CATCACGTGC TCCAATTTGT	2460
	AATGGAACAA CGATATTCCA TAAACCAAAG ATAAATGGCA TTGCCATGAA GATAATCATG	2520
45	ATTACACCAT GTGTACTAAA AATTTTCGTTA TAGTGGTTAG ATTCTAAAAA TTTGTTATCA	2580
	GGTACTGTTA ATTGCGCACG AATAAGTAAC GCATCAATAC CACCACGGAC GAACATTAAT	2640
50	ACGGCACAGA TTAAATACAT AATACCGATT TTCTTATGGT CTACAGATGT GAACCATTCT	2700
	TTGTAAAGAT ATTTCCATAA TTAAAGTAA GTAATTACTG CGATTAAACC AATAACTAAG	2760
55	AATGGGGCAC CAATTTGTGC CATTGTAATC ATCAGTTAC CTTTAACTAG TAATTGATCC	2820

TTGAAATTTT CTTCAATTTCT TTCGCATTTT TCGATTTCATC TTTCTTGAAC TCATTGTTAT 2940
 ATGGTTCGTC ATTTCCAAGA ATCATCAACT TCATACCATG TCGTTTATAG TTCGCATTTG 3000
 5 TAATTTGAGC TTTACGAGCA GGTATTAATG GTTTGTCTGA TACATCTTTA AACATATTTT 3060
 CTTCACTAGT GAAGTTTGGA TCTTTCAATT CGAAATTGAA ACGTTTATAT GCATAGAAGA 3120
 TGTATTCTGG ATCGGCTGCT GGATCAACAA ACGCCATATG TGTACCATTA AATTCTAAAG 3180
 10 CTTTATTAGG TGTGCTTGGT AATAATTGTT TATCAAATGT ATCTTGATCT AACGTTTTCT 3240
 TACCTTTAAC TTCTTTTACC CATTTGTCGT AGTCTTTTTG ACTAACGGCA TTTACTTTAA 3300
 ATGTTTGACG TGAGAATCCT TCACCATTGA AGTTAGAGTT ACGACCTCTG AACGTACCAG 3360
 15 TTTGAGATGC TTCTAACGTC CAATTCATTG TCATGCCAGT CATGGCATAT TTTTGACCAC 3420
 CTAATTGTGG AATCCAGAAA CTGTGCATTG TATCCATAGC TTGAAGCTTA AATACAACAG 3480
 GACGATCTTT AGGGATTGTT AATGTATTAA CAGTCTCTAT ATGTTTCATCT GGATAAGCAA 3540
 20 AGAACCATTG GTATCCTGCA CTTACTGCAT ATACAACCAT TGGATCTTTC TCACTCTTCG 3600
 GTGGTTTTTC GTAATCGTAT AAAGTTTTAA CTGTAGGAAT AGCTAAAGCA GCAACGATTA 3660
 25 TGATAGGTAT TACAAACCAT ATTGTTTCAA TGATGGCATT ATGGTGCATC TTACCAGATT 3720
 CGGCATTCTT ATTATACTA TACTTGTAAT TAAAAATGGC GAACATGCCA AGTACAACGA 3780
 AACAAATAAC AAGCATGAAG ACGATTGAAT AAAGAATCAA GAACTTCTGA CTACTTGCTA 3840
 30 CTGGCCCTTT TCGGTTGAAA ATTTCTATAT TTGAACAACC ACTAAGTAAA ATTAGTGTGC 3900
 CAAATAATAG AAGCAAAGAC TTAAATTTTG ACACTTTTTT GACCTCCTAA TACTACAAAT 3960
 GTAGGGCTTA ACATTAATTT TAAGTTATTA CACAATATT ACAAGGGCTT ATGGGAAAAA 4020
 35 AATTAATAAA ATTGTATCAA AAATGTTGAT AAATCAAGGT GTGACGTGGG TTCACACATT 4080
 TGTAAAAATT ATGTGTACAT TTTGTGACTA ATAGCGTTTT T 4121

(2) INFORMATION FOR SEQ ID NO: 316:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9310 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

50 CGAGTGAGTA CAAACATATT TTTATTTGCA AGGGGTAAAT GGCATATAAC TATCTTTTTT 60
 ATGTAAGCTG GTATAAAATT TATACTAATA GGAGGGATAG TATGAATATA GTAGGGCATC 120

55

EP 0 786 519 A2

	TTGGATTACG	ATTAGTTGAA	AAGTCGGTTA	ATCAAGACAA	TCCTTCAATG	TATCATTTGT	240
	TTTATGGGGA	CGAAGTAGGT	ACAGCCGGAA	CAATTTTAAG	CTTTTTTGAA	ATTCCCAATG	300
5	CGGGTCATAA	GCAGCCAGGT	ACTGAAACGA	TTTATCGATT	TTCATTATTA	GTACCAAATC	360
	AAGCGGCACT	TCATTATTTT	GAAAAACGTC	TTGAGAATAA	TGGTATTAAG	TCTGAACGTT	420
	TGTACTATCT	TGGACAAGAA	GGTGTGTGCT	TTAAAGATGA	AGACGACTTA	GAAATCATAT	480
10	TGCTTGTTAA	TGATAGTTTT	GAAGTACCAC	ATCAATGGCA	ACATAACGCT	TATAGTGAAA	540
	TACCTCAAGC	ATATCAAATT	TTAGGAATAG	GGCCAGTCGA	ATTAAGAGTT	AGAAATGCAG	600
	CGCGTAGCGT	AGAATTTTTG	GAAATGTCT	TAGGTTATCG	CAAAAGAGAT	AATAAATCAT	660
15	TCGATGTGCT	GACATTAGCA	CCACAAGGTT	TATATTCGGA	TTTTGTAGTT	ATTGAGCAAC	720
	AGGGACAACG	TGAAAGACCT	GGACGAGGTT	ATATCCATCA	TATTGCAGTT	AATACACCAC	780
20	AAATGAGTGA	CTTAGATGCA	ATTTACAAGA	AATTACAACA	ACAACCACAA	AGTAATTCAG	840
	GTATAATTGA	TCGCTATTTT	TTTAAATCAT	TATACTATCG	CCATAATTCA	ATTATGTATG	900
	AATTTGCGAC	TGAAGCGCCT	GGATTTACTA	TTGATACACC	TGTTGAACAA	TTAGGAAGTC	960
25	AATTGAACTT	GCCTGACTTT	TTAGAAGCAG	AACGTGAACA	AATTGAAAGT	AAGTTACACG	1020
	AAATATAAAG	GAGAATGTTT	AATGGCCAAA	TTAGAAATGA	ATAAAAATAC	GCCTCTTGAG	1080
	TTTGGTTTGT	ATTCCTTAGG	TGATCATTTA	TTGAATCCAT	TGAAAGGTGA	AAAAGTTAGT	1140
30	TATGAGCAAC	GTATTAATGA	AATTATTGAA	GCAAGTAAAT	TAGCAGATGA	AGCAGGTATT	1200
	GATGTTTTTG	CAGTTGGTGA	AAGTCATCAG	GAGCATTTTA	CAACACAGGC	ACATACGGTT	1260
	GTGTTAGGTG	CAATTGCCCA	AGCGACAAAG	CATATTAAAG	TTTCAAGTTC	TTCAACGATT	1320
35	ATTAGTGCAc	AGATCCTGTA	AGAGTATTTG	AAGACTTCGC	GACATTAGAT	TTGATTTCTC	1380
	ATGGTAGAGC	CGAAATTGTA	GCTGGCAGAG	CATCAAGAAC	AGGTATTTTT	GACTTGTTTG	1440
	GCTATGATTT	AAAAGACTAT	GATGAATTGT	TTGAAGAAAA	ATTAGGTTTA	CTTTTAGAGT	1500
40	TAAATAAAAC	TGAGCGTATT	ACTTGGTCTG	GAAAATATCG	TCCAGAACTT	AGAAATATGA	1560
	AAATATTCCC	AAGACCAATC	GATAATATAT	TGCCAATATG	GCGTGCTGTT	GGTGGTCCAC	1620
45	CTGCAAGTGC	TATTAAAGCG	GGAAAACAAG	GTGTGCCAAT	GATGATTACA	ACCCTTGGTG	1680
	GCCCAGCAAT	GAACTTTAAA	GGTCTATAG	ATGCTTATCG	TCAAGCGGCA	ACTGAAGCAG	1740
	GTTCGATgc	TTCCGCTAAG	TCTTTACCAG	TAAGTACAGC	GAGTCTGTTT	TATACAGCTG	1800
50	AAACAACCTCA	GGATGCTATG	AGAGAATTTT	ATCCACATTT	GAATACAGGG	ATGTCATTTA	1860
	TTCGTGGTGT	TGGTTATCCG	AAACAGCAAT	TTGCTAATTC	GTCAGATTAT	CGAGAAGCGC	1920
55							

EP 0 786 519 A2

	GTCATCAACG TTTTATGGCA CAGCTTGATT TTGGCGGTGT GCCATTTGAA AATGTTATGA	2040
	AGAATATTGA GTTAATTGGC AACGACATTA TACCGGCGAT TAAAAAGCAT TTATCAAAAT	2100
5	AGGAGGGGCG TCATCATGAA TATTGTATTA TTGTCAGGTT CCACAGTAGG TTCTAAAACG	2160
	AGAATTGCTA TGGATGATTT AAAAAATGAA CTAGAAGTCA TCAATGAGGG ACATCAAATA	2220
	GAGTTGATGG ATTTACGAGA ACTTGAATTA GAATTTAGCG TTGGAAAGAA TTATCTAGAT	2280
10	ACTACAGGAG ATGTATATAA ATTAACGACG TCGTTAATGC AGGCTGATGT GATTTTTATT	2340
	GGTTTTCCAA TTTTCAAGC TTCCATCCCT GGTGCTTTGA AAAATGTGTT TGATCTACTT	2400
	CCAGTCAATG CGTTTCGTGA CAAGGTAATA GGACTTGTAG CGACAGCAGG TTCTAGTAAA	2460
15	CATTATTTAA TTCCTGAAAT GCATTTAAAA CCAATATTGA GTTACATGAA AGCACATACG	2520
	ATGCAAACGT ATGTATTTAT TGAAGAGAAA GATTTTTCAA ATCAACAAAT TGTCAATGAT	2580
20	GATGTTGTAT TTCGGTTAAA AGCGTTGGCA CAATCCACAA TGCGAAGTGC CAAAGTACAA	2640
	CAACAAGTGT TTGAAGAAGA AAACAACCAA TACGACTTTT AAAGTATAAA AATAAGACGC	2700
	TCGGCACACT AAATTTGTAA GTGTTTGAGC GTCTTTTCAT ATTAAGTATA TAGCCAATGA	2760
25	ACGACGATAA AGGCAAGTGA TGACAAGCAT ATTGAGGTAA TAATGATTGT CATAAGCGGT	2820
	TTAAGTGCGC GATTTTTAAG ATCTTTAAAT GCAACATTTA ACCCTAAAGC AACCATGGCC	2880
	ATTAATAAGC AAATTGTTGA TACAGTATTT AAAATATTTA GCAATGCTGA CGGAATAGTT	2940
30	ACATATGTAT TCACTAAGGC CATAATGACA AATCCAATTA AAAAGTATGG AATGCTTATT	3000
	CGACCCCTGC TAGATGATTC TGATGAACGG AAACGCATAA TTAAAATAAG TACGATGGTT	3060
	AATGGAATCA GTAAGAATAC TCTACCAAGT TTACCAAGAA GTGCAATTTT AAGTGCATCA	3120
35	CTACCACCAA AGCCACCAGC TAAGACAACG TGTGCAATTT CATGAAGACT AACACCAGAC	3180
	CAAGCGCCAT AAACATTTGT CGTCATTGAA AAGATAGCGT AGATAGCTGT ATATATAAGT	3240
	GAAAATATCG TACCAATCAA TCGATGATA CCGATACTAA TAGCTGTATC CTTTTACGT	3300
40	GATTTGAATA TTGGAGCGAC TGCGGCAATA GCAGCAGCAC CACAAACGCC TGTGCCGACA	3360
	CCTAGTAATA ATGCGATGTT TTTGTCACCA TGCAACAGTT TGTGACAAA GAGCATCATT	3420
45	ACAATACTGA AAATAACGAC ACCTACATCG ATGGCTAATA GTTTACTACC TTGACCGATA	3480
	ATATCGAATA TATTGAGTTT AAGTCCATAT AGGATGATTG CAAATCTTAA TAAATATTTA	3540
	GATGAAAACG TAATACCTGA GCTATATTGT TCAGGATATC CTCTAAAGTG ACGATATAGA	3600
50	ATAGCGATTA ATATCGCGAT AGTTAATGCG CCAACCTTAT CTAGGATTGG CAATTTAGCT	3660
	GCTAAAAAGC TAAATAATGC GACTATAAAT GTTAGTGATA GCCCAATCAT AAAATGCTTA	3720

55

	ATTTTAAAT ATAAATTTGG AATGAATAAT AAAGTAGTGA TTAAATTAAG TTGTGTGATA	3840
	GGAAACTTGG ACATCAATCA AAGTAATAGG CACTACAACG CTTATTGGCG GGGCCCCAAC	3900
5	AAAGAAGCTG ACGAAAAGTC agCTTGcAAT AATGTGCAAG TTGGGGATGG GCCCCAACAT	3960
	AGAGAAATTG GGTCCGTAAT TTCTACAGAC AATGCAAGTT GGCGGGGCCC CAACATAGAG	4020
	AATTTGAAA AGAAATTCTA CAAGCAATGC AAGTTGGGGA AGGACAACAA ATTTAAGATA	4080
10	CAATGCGTAA CATTAAATATG TTATTATAAT GATAATTTAC AGAATTATAT GAAAAATGAA	4140
	TGAGGATGTG ATGGTATGTT TGGAAATGAAA GTGAATGAAC AAATAACATT AAAAAATTTA	4200
	GAAGCTCATG ACACAGAAGC GCTTTTCAAT TTAGTCAATC GTTCAAGAAA TTCACTTAGG	4260
15	GAATGGTTAC CTTGGGTAGA TGCAACTGAG CAACCATCAG ATACGCGTGC ATTTATTAAA	4320
	AGAGGACTTT TGCAATTTGC TGATGGTAAT GGATTTCACT GTGGCATTG GTATGAAGGA	4380
20	ACGCTAGTTG GTGTCATCGG TTTACATGAA ATTAATCACA TGCACAGAAA AACTTCATTA	4440
	GGGTACTATT TAGATAAAGA ATTTGAGGGT CATGGGATTA TGACACAAGC AGTTGAGGCA	4500
	TTGATAAAGT ATTGTTTCGA AGAGCTTGAC TTAAACCGAA TTGAGATTAG TGCCCGAGTT	4560
25	AATAATGAAA AAAGCCGGGC TATTCCTGAA AGGCTGGGAT TTAGTAGAGA AGGTATGTTA	4620
	CGTGACAATG AATTACTAAA TGGTATTTAT TCATCGAGTT ACATCTATAG TTTATTAAAA	4680
	TCAGAATACG ACCAAAAATG ACAAATTAGA CTTACAAAAG AGTGATGACA TTTAAATGG	4740
30	CAGCGCTCTT TTATTTAATT TTTGAAAATA AAAGGTTGTT GACAGTATTA TTTTATAACA	4800
	ATATAATGAT TTTGATAATT ATTATCAACT AGATGATGTT TATGGGAGGA TGCTTTAAAA	4860
	CAGCCGTTTT AAGTGTAATG TATTATTTTA GCGTGTAGGG AATGCGAAAA TAATATTTAT	4920
35	AAGAACACAT CTATGGGGAT AATAGAATTT CTATAATGAG GTGTCAAAAT GAAAAAGTTA	4980
	ACAACGCTAT TATTAGCATC AACGTTATTA ATTGCTGCAT GTGGGAACGA CGATAGTAAG	5040
	AAGGATGATT CAAAGACATC GAAAAAGAT GATGGTGTTA AAGCAGAATT AAAACAAGCA	5100
40	ACAAAAGCAT ATGATAAATA TACTGATGAA CAGTTAAATG AATTTTTAAA AGGTACAGAA	5160
	AAATTTGTTA AAGCGATTGA AAATAATGAT ATGGCCCAAG CAAAAGCGTT ATATCCAAAA	5220
45	GTTTCGTATG ATTATGAACG CTCTGAACCA GTTGCGAAG CATTTGGAGA TTTAGATCCT	5280
	AAAATTGATG CACGTCTTGC AGATATGAAA GAAGAGAAAA AGGAAAAAGA ATGGTCAGGA	5340
	TATCATAAGA TTGAAAAAGC ATTATACGAA GATAAGAAAA TTGATGATGT GACTAAAAAA	5400
50	GATGCACAAC AATTATTGAA AGATGCAAAA GAATTGCATG CCAAAGCTGA TACATTAGAT	5460
	ATCACACCAA AATTAATGTT ACAAGGTTCT GTTGACCTAT TAAATGAAGT TGCAACTTCT	5520

55

EP 0 786 519 A2

	GTTGAAGGCG CACAAAAAAT TTATGACTTA TTTAAACCTA TTTTAGAGAA AAAAGATAAA	5640
	AAATTAAGTG ATGATATCCA AATGAACCTC GATAAAGTGA ATCAATTATT GGATAAATAT	5700
5	AAAGATAACA ACGGCGGTTA TGAGTCATTT GAAAAAGTAT CGAAGAAAGA CCGTAAAGCA	5760
	TTTGCGGATG CTGTTAATGC ATTAGGAGAG CCACTAAGTA AAATGGCTGT GATTACTGAA	5820
	TGACAAATTA TGAACAAGTT AACGATAGTA CGCAATTTTC AAGACGTACA TTTTGTAAAA	5880
10	TGTTAGGTAT TGGCGGTGCC GGTGTTGCAA TTGGCGCAAG TGGTGTGGT AGCATGTGGT	5940
	CTTTCAAATC AATGTTCAAT ACACCAGAAG ATCCGGAAAA AGATGCGTAT GAATTTTATG	6000
	GTAAAGTGCA ACCAGGCATT ACCACACCCA CGCAAAAAAC ATGCAATTTT GTTGCCTTAG	6060
15	ATTTGAAGTC AAAAGATAGA GATGCAATTA AGGCAATGTT TAAAAAGTGG ACGGTTATGG	6120
	CTGATCGTAT GATGGATGGT GATACAGTTG GCAAGCCGAG TAACAATCCT TTAATGCCAC	6180
	CAGTAGATAC CGGTGAATCG ATAGGATTAG GTGCAAGCAA GTTAACGATT ACCTTTGGGA	6240
20	TTAGTAAGTC TTTGATGAAG AAAATTGGGT TATCTAGTAA AATTCCTGAT GCCTTTAAAG	6300
	ATTTACCGCA TTTTCCGAAT GATCAGTTAA TAGACGATTA CAGCGATGGT GATATTATGA	6360
25	TTCAAGCATG CTCAAATGAT TCGCAAGTAT CCTTTCATGC GGTTCATAAT TTAGTTCGTC	6420
	CATTCGAGA TATTGTTAAG GTACGTTGGG CGCAATCTGG TTTTATCTCT GCTAAAGGTA	6480
	AGGAAACACC TAGAAATTTA ATGGCATTTA AAGATGGAAC AATTAATCCT AGGAAGAATA	6540
30	ATCAACTTAA AGATTATGTG TTTATTGATG ACGGATGGGC GAAACATGGA ACTTATTGCG	6600
	TTGTCAGACG TATTCAAATA CACATTGAAA CGTGGGATCG TACTGCGCTG GAAGAACAAG	6660
	AGGCTACATT TGGTCGGAAA CGACATAGTG GTGCACCGTT AACAGGTGGG AAAGAGTTTG	6720
35	ATGAAATTGA CTTAAAAGCG AAAGATAGTC ATGGCGAGTA TATTATTGAT AAAGATGCCC	6780
	ATACGAGGCT AGCGAAAGAA GCAAATACGT CAATTTTACG TAGAGCCTTT AATTATGTGG	6840
	ATGGTACGGA TGACCGCACA GGTAACCTCG AAACAGGCTT ACTTTTTATT GCTTTTCAAA	6900
40	AAGCGACAAA ACAATTTATC GATATACAAA ATAATTTAGG TAGTAATGAT AAATTAAATG	6960
	AATATATTAC ACATAGAGGT TCTGCTTCAT TTTTAGTATT ACCAGGTGTT AGTAAGGGAG	7020
45	GATACCTTGG TGAAACATTA TTTGACTAAA TTTGTAGCAA TGCTAATAAC TGCTGCTATG	7080
	GTGTGTAGCT TTGGGTTACT GAAAAGTCAG GCAGCAGAAC AACAAAGTAT TAGTGATGTA	7140
	TATAGTGTGA TAACGGATGC GAAATCTGCA CTTTCTAATA ATTCGATATC GAATGACAA	7200
50	AAGCAGAAAG CAATTGAGCA AGTGGTAAGT GCAGTTAAGA AATTATCGCT TGAAGATAAT	7260
	AGTGAAAGTA ATGCTGTCAA ATCAGATGTG AGAAAGCTTG AAGATGCAAA AGCGAATGAT	7320

55

EP 0 786 519 A2

	GCTAGTAAAG ATGCGGGTTC TAAAATTAAA CTATTGCAAC AGCAAGTCGA TGCTAAAGAT	7440
	GCTGCGATGA CAAAAGCGAT TAAAGATAAA AATAAGCGG AATTAGAATC TTTGAACAAT	7500
5	AGTTTGAATC AGATTTGGAC AAGTAATGAA ACAGTGATTG GCAATTATGA CGCAAATCAA	7560
	TATGGACAAA TTGAAGTCGC ATTATTACAA CTTAGAATTG CAATTCATAA GTCACCATT	7620
	GATACGGCAA AAGTGTGACA TGCTTGGACA ACTTTTAAAT CAAATATTGA TCATGTCGAT	7680
10	AAAAAAGTA ATACGTCTGC AAATGATCAA TACCATGTAT CACAATTAAA TGATGCGTTA	7740
	GAGAAGGCGA TTAAAGCTAT CGACGACAAT CAATTGTGCG ATGCTGcaTg TGCGCTTACA	7800
	CATTTTATAG AAACCTGGCC GTATGTTGAA GGTCAAATTC AAATAAAGA CGGTGCTTTG	7860
15	TATACGAAAA TTGAAGATAA AATACCATAT TATCAAAGTG TATTAGACGA ACATAATAAA	7920
	GCACATGTGA AAGATGGTTT AGTAGATTTA AATAACCAA TTAAAGAGGT TGTGGGCCAT	7980
20	AGTTATAGCT TCGTCGATGT GATGATTATC TTTTACGTG AAGGGCTAGA AGTGTTGTTA	8040
	ATTGTAATGA CATTGACTAC CATGACGCGT AATGTAAAAG ATAAGAAAGG GACTGCAAGT	8100
	GTGATTGGTG GTGCAATTGC CGGACTTGTA CTGAGTATTA TCTTAGCAAT TACGTTTGTA	8160
25	GAAACTTTAG GGAATAGTGG CATTCTTCGT GAAAGTATGG AAGCGGGATT AGGTATCGTT	8220
	GCGGTCATAT TAATGTTTAT CGTTGGTGTT TGGATGCACA AACGTTCAAA TGCAAAACGT	8280
	TGGAATGACA TGATTAAAAA TATGTATGCT AATGCGATTA GTAATGGTAA TTTGGTATTG	8340
30	TTAGCGACGA TTGGTTTAAT ATCTGTGTTG CGTGAAGGTG TCGAGGTTAT CATTTTCTAT	8400
	ATGGGGATGA TAGGTGAGCT AGCGACCAA GATTTTATTA TTGGTATTGC TTTAGCTATC	8460
	GTTATTTTAA TCATCTTTGC ATTATTATTT AGATTTATAG TTAAATTAAAT ACCTATTTTC	8520
35	TATATATTTA GAGTGTGTC GATCTTTATT TTTATTATGG GATTCAAAAT GCTTGGCGTA	8580
	AGTATTCAAA AGTTACAATT ATTAGGTGCG ATGCCAAGAC ATGTTATTGA AGGATTCCCA	8640
	ACGATTAAC TGGTTGGGCTT TTATCCAAGT TATGAACCAT TGATAGCACA AGGTGCTTAT	8700
40	ATTATGGTAG TTGCTATCTT AATCTTTAAA TTTAAAAAAT AAAAAACAGG CCGAGTGCCT	8760
	GTTTTTTTTG TTGCTATATT GGAAATATTC GGTATTGCAG TATAACGATA ATCAGCAT	8820
45	TGATTGCTAT AAGGTTAATG TGTGGCGGT TTGCCTCGGC ATGTGAACTT AACGATGAAC	8880
	ATACTGAACT CAAAGAGCAA TATGAGTGGC AATGTGAGTA ATATATTTAA TGTTAAATCG	8940
	GGTGGTGCAA TGATACTTGC TAATACAAAG CAAGCGAAAT AAATATATTT ACGrTAATGT	9000
50	TTCAATGATG TGGTATCTAT AAGACCGAAT TTTGCAAGAC CCATAAATAA TATTGGTAAT	9060
	TGAAATAGAA GACCAAATGT GAATAACCAA CGTATGAGTT CAATCAAATA TGCTTTAAAG	9120

55

GGAAAGCCAA CATAAAATGC AAAAGCGACG CCAGCACAGA ATAATAACAC GCTGAAAAAA 9240
 CTATATTTAT AAATAAATTG ACGTTCATTA TTATGTAATC CAGGTGCAAT GAATGCCCCAC 9300
 5 AATTGATAAA 9310

(2) INFORMATION FOR SEQ ID NO: 317:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3458 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

ATTAGTATTA CACACTAATG TAAATAGATT GGTGGAGAA GAGATATTTG CTAATAAGTG 60
 CCTTGCTAAT AATGATGTTT AAATTATGAA TTCAATAAAA AAATTAATTG AAGCTGAGTT 120
 20 ATTAACAACG ACTAATGATT TTGAAGTTAG TATATATAAA AAGACAAGAC CTGAATTACA 180
 AAGTATTTTA AAAAGTTTTG GTATAAAAAC AACAGGTAAT AAACCTGACT TAATTAAACG 240
 25 TATTGACGAC AATTTTCATA TTATTAATAA CTTAGATTTA CCATATGTAT ATATACCAAT 300
 TAAGAAAGGA GAAGAAATTT TAAAGAAAAC AGAGTACTTA ACCTCTTTTA TACAAAGTTA 360
 TGGTGAAATT TCTCTTGAGC GTGCTTATTA TTTGGTTGAA AACTATATAG ATGAAAATTG 420
 30 CGATGATAAA GTTGACAGAA TATACAAGTT TGAATTTCAA AGAAAATATG ACAATGGCGA 480
 GTTTGATTTT AATCATGGAT ATAATTTCTGA ATTGAATATG TTGATAGATC ACTATAAAAG 540
 AGATGTAAAA GACTACGATA ATGCCAGAAA GTATTCAAAT ATTTATCTTT ACTTTGGTTT 600
 35 GAGAGATTTT TTAATAAAAT TAATGAGTAA TTATTCATAT TACGATAGTA AAGGGAATAT 660
 AGATTTGAAC GAAATACAAA ACGATCTGAA TAGATTTATC AACTCTAGCG CTTCTGGTAT 720
 GTACGAGCGA TTAATATATA ATGAAAATTT ATCCAATAAT ATTATGTTTG AATTATTTAA 780
 40 AAAGGACACA CAAGATTATA GTGATTTGGA AGAACAATTG ATTGAAAAGT TCATAAACTA 840
 TGTAGTGTCT AATGTAAAAA AAGAAAGTAG GAGTAATACT CTTATAGAGT TATCAAAAAT 900
 TTTAGAGAAC GGATATACAA TTGATAAAGA AGAATTTAAA AAAGAAGATG ATTATCTTTC 960
 45 TAAGTACATA TTTACTGACA TAGATTATTT GAAAAAGTTA GAATCAAAAA TAAACGTTGC 1020
 TATTGATATT CGAAGTGGAG AAATTCATTT GGTATTAGAT GATGATAGCC TTGATATATT 1080
 50 AATACAAAAT CAAAAATACG GCAATGAGTT TTGAGTCATA ACTAAATATA ATATGTTGAA 1140
 GAGAGGTTTT ATTAAATGG CTAAAATTGG TTATGATCGT GTATCAACGA AAGATTAATA 1200

EP 0 786 519 A2

	AACGTACAGA GCTTGTTAAG TGTTTAGATT ATTTACGAGA GAGCGACACA TTAGTTGTCT	1320
	ATCAACTTGA TCGGTTAGGT AGAACGACAA AACTATTaAT TGAATTaTCA CAATGATTCTG	1380
5	ATGATAACGG AATTGACTTA CAAATTAGTA ACATGAACAT TTCAACGAAA GACACAATGG	1440
	GCAAAATGTT TTTTACGATG ATGAGTGCAT TTTTCaGGAT TAGAAGTTAA TTTACTATGT	1500
	GAGTGATATAA AATAGACTTA GCAGCAACAA GAGCGAGAGG CCGAAAAAGC GGGCGCCCCCT	1560
10	CTTTACCAGA GAATAAAAAA TGAGAAATTA AATTTTTTATA TGATGAACAA ACGATAACAG	1620
	GGGAAGAAAT AGCTAGTTAG ACAAGGGTAT GTCACTCAAC TGTTTATCGA GTTGATTAAAG	1680
	AAATGAAAAA ACTTATACTA TGAATTACTG TTTAAAAGTG TGCATGTTAT AATATTTATT	1740
15	GAGCAAGTTG GATAGATGGT GGCTAATCTC TTAATAAAGG GGTGATGCCT ATGGTTATAG	1800
	TTGTTACTCC TAGaAAAGGA CTAGCATGTC TGATTTTGAA ATGCTTATGG TTGTATTAAC	1860
20	AATCATTGGT TTAGTATTGA TTAGTACTCA AGACCATAAA AAATAaCCTT CTATTCGCTT	1920
	TGACCGGCAT TTTTGAAGGC TATTTTTTAAA TAAAATATAA GGTCACCGTC TTTTAAACGG	1980
	GCTCATTAGG GTAACATGTT TCCGAGTGTT GCCCTTTTGG TGTTCaAGA GTTAATGATT	2040
25	TTCATCTTTT GCTTCTACTT GCTACAAATA TATTTTAAACA CATTTTCTT ATGAATTGTA	2100
	GTTCTGAACA TAATCAGAAT TAATAAAACC AACTTTCCAT ACAGCAGAAA ATACAATTAA	2160
	AAGTATAGAA TGTAATCAG AATTATATAA AAAGTATTGG AGACCTCTTC ATATATAGAT	2220
30	AATTCACTTA GTTATTTTATG AAAGAAGCCC CTaCAACTA AAGTTGAAAA ATAGAGGAAC	2280
	ACAGTTgGAT TACGCATCAA CTGCATaAGg CCCCTAAaAA CTAAAGTTGT AAGGGGcYCT	2340
	AAAATTTTATT TTGGTTGATT GTCTTCTGGT TTATCTGAAG TCATTGTTTT TGTTGTATTA	2400
35	TCATTTAATG AATCTTGTGC TTTTCTGTT TTTGTTTCAA CAGATGTTGT CACTTTATCA	2460
	TTTTCTACTT TTGTATTTTT TGTTTCAGTG ACAACTTTTT TATTCTTGTC AGCTGTTTTT	2520
40	ACTTTATCTT TAGTAATTTT TTGACCACGT TTTAAGTAAT ATTGAACGAT ACCCATTAAG	2580
	ATGATTGCAT GAATAACAAC GAATAAAATA ATTGTCATTA CTGTATAAAC TCTTATGATA	2640
	TTTTCTGCAA CACTTTGAGA GAGTGATTGC GTATTAAATG AAATCAAGTA ACCAACAGGT	2700
45	GTTTTAAATA AAATAATAAT TAAGTTTAAT AATAGTATTC CAATGAAAAA TTTGAAAAAT	2760
	GTTTTTTGAC CATTTTTTCAT TGCTTTAAAT CCATTAGCTA AGTGTGTTTT TACTTTATCG	2820
	TTAGTTGATT CaACGAAACT AGTCaTAAAG TTGAAAATAG GTATTAAACAA TAACCAAGTA	2880
50	ATAATTGCTT TAATCAATAA AACAACAATG ATTATGATGC TTTGAGTTGT AATAGCAATA	2940
	CCAATCAGGT GTGTACTATC TGCATATGAT GATTGTACTG AATTCATAAT CATTTCTGAT	3000

55

ATAGATACAA GACCGATCAG AACGCTCTTT AAATAATTAC CTTTTTTAAA TGCAATAAAT 3120
 AAATCGGTAA ATTTTACTTT TTCATGACTC ATTGCTCGTT TCATAACATT TGTAATTCCG 3180
 5 ATAAAAATTT GAACTAAAAC AAACAGTGAA ACTACAGCTG CTATAAGTAA TACAGCGATA 3240
 ACTTTTAAAT ATGCATCTAC TGGTGGTTTT TGCCCAAATT GTGAATAAAT TGCAATAGTT 3300
 TGTGCATACT TAGCAAGCTG GAAATTAGCA AGTAAATATA CAACTGCAAT TACAGCAAAT 3360
 10 GCTATAAGTG CATATATCAA TGTTATTAAT AATTGTGGTT TAGCATTTTT AAATGCTGAT 3420
 TTAAAGTAAG TAAACAAAGT GGTGCCTCCT TTTCTTCA 3458

(2) INFORMATION FOR SEQ ID NO: 318:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 695 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:

25 GCTTGCAAGA TGCTTTCATT AAAGACAAAG TCATCGATAT TATGAATATG TTCCAAAATC 60
 ACACTGATAT CACTTATACG TTGAATAAGT CTCATGCACA TATTATATGT ACGCCAGAAa 120
 TATTTGCTAA ATTGTTACAT ACGATTGCAA CTAGAAATAT CGACATTCTA TCTGCCAATT 180
 30 ATAGATCGTC ATCTATGAGC AAAGCTCGTA TATCATAATA AAAGACACAT CTATACGATG 240
 ATCATTTACA ATGATTGCGT ATAGTGATGT GTCTTTTTTT GTATTATTTA TCTTCGGACg 300
 ATTGATCCTC AACCATTTGT TCAAATGCTT CACGAACTTG TGGTACTGTC ATACCTACTA 360
 35 CAACTTGGAT ACTCTTACCA CTCTTTACTA ATCCGTGTGC CATTTGTTGA TGCGTGAAAT 420
 ATTCCGTATC TGCAACTTTA CTTTCATCAT AAACAGTTAA GCGTAATCTT GTTGTACAGT 480
 TAGTAACATC TTTGATATTT TCTTTGCCGC CTAAACCGTC GAGGTAATAT GCAGCTTTGT 540
 40 CTTCGTATTC ATTACCAGAT GAAGCACGTT TAGAATCTAC GCTATCACCT TTTTATTTT 600
 TGTAATCTTG CTTAGAAAAT AATTAACTT CCTCTTCTGT TTyCTTACGT CCAGGTAATG 660
 45 GAATATCAAA TTyCAAAATT AAGAATCTGA ATAAG 695

(2) INFORMATION FOR SEQ ID NO: 319:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 875 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

5 AATATTGaTa TTATATATAG TCGAATCaAT GATGACCCaC ATCGACcTAA TGCAAATAAT 60
 AAGACAATCA AGCAGTTAAA TGA~~CT~~TGTAT TGCTCGCGGA ATTTAAGATA TAATGAAATC 120
 GCATTCAAGA AATTTGATAG TCATTTGCTA TCAATTTcAG AAATATATTA TGAATTGCTA 180
 AATTTAATAA AAGCGAGTGA TCAGTATTAG AGAGAATAGA GCGTTAAGAC TCTATCGCCG 240
 10 AAGcGCAAGT AATTTATTAC GAAACTCTCA GGcAAAAGG~~r~~ TAATACTGTA ACGCGTTCCT 300
 GAATTGGTGA TTTATAAACA GGGTAGCGAT TGCTATCCTG TTTTATAAT TTTAAGGGGG 360
 TATTTCAATG TCAAGTGATT TAAAACAAAC ACCTTTATAT CAAAATTATG TTGATAGAGG 420
 15 TGCAAAAATT GTGGAATTCG GAGGATGGGC GATGCCTGTT CAATTTTCAA GTATTAAAGA 480
 GGAGCATAAT GCTGTTCGAT ACGAAATTGG CCTGTTTGAT GTTAGTCATA TGGGTGAAAT 540
 20 TGAAGTAACA GGTAAAGATG CTAGTCAGTT TGTGCAATAT TTATTATCAA ATGATACTGA 600
 TAATTTAACT ACTTCAAAG CATTATATAC TGCTTTATGT AATGAAGAAG GCGGTATTAT 660
 TGATGATTTA GTAATATATA AATTAGCTGA CGACAATTAT TTATTAGTTG TTAATGCTGC 720
 25 TAATACTGAA AAAGATTTTA ATTGGATTTT AAAACACAAA GAGAAATTTG ATGTTGAAGT 780
 ACAAATGTA TCAAACCAAT ATGGTCAATT AGCAATACAA GGACCAAAAG CtAGAGATTT 840
 AATTAATCAA TTAGTTGATG AAGAkGTAAC TGAAA 875

(2) INFORMATION FOR SEQ ID NO: 320:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5897 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

40 TCTTTTTACG AAGGCTACCG TTTAGGATCT GATAACTTAA CTGTAGGAGA AATTGTATTT 60
 AATACAGCGA TGACAGGTTA TCAAGAACT ATTTcAGATC CATCATATAC AGGTCAGATC 120
 45 ATTACTTTTA CGTATCCATT AATCGGTAAT TATGGTATCA ATAGAGACGA TTTTGAATCA 180
 TTAGTACCTA CATTAAACGG TATTGTAGTG AAAGAAGCGA GTGCGCATCC AAGTAATTTT 240
 AGACAGCAAA AGACACTTCA TGACGTTTTA GAATTGCATC AAATTCCAGG GATTGCAGGT 300
 50 GTTGATACAA GAAGTATTAC GCGTAAAATT CGACAACACG GTGTGTAAAG AGCTGGTTTT 360
 ACTGATCGAA AAGAAGATAT TGATCAACTT GTCAAACATT TACAACAAGT AGAATTACCT 420

	AGTGTGCTAC TTGTAGACTT TGTAAGAAG CAAAATATTG TTCGAGAATT AAACGTCAGA	540
	GGTTGTAACG TCACAGTTGT ACCATATACA ACTACTGCCG AAGAAATTTT AGCAATGGCT	600
5	CCAGATGGCG TTATGCTATC AAACGGACCA GGTAATCCTG AAGTTGTAGA ATGTGCGATT	660
	CCAATGATTG AAGGAATTTT AGGGAAAATT CCGTTCTTTG GTATCTGTCT AGGACATCAA	720
	CTTTTTCAT TATCTCAAGG AGCAAGCTCA TTTAAATGA AGTTTGGTCA TCGTGGTGCG	780
10	AACCATCCAG TTAAAAATTT AGAGACTGGA AAAGTTGATA TTACGAGTCA AAACCATGGA	840
	TATGCAATAG ATATAGATTG GTTAAAAAGT ACTGATTTAG AAGTTACTCA TCTTGCAATTA	900
15	AATGATGGTA CTGTAGAAGG TTTAAACAT AAAACATTAC CAGCATTTC TGTTCATAC	960
	CATCCTGAAG CAAATCCAGG ACCGTCAGAT TCAAACATC TATTTGATGA TTTTGTAGCA	1020
	ATGATGACTA ATTTTAAGGA AAAGGAGCGT CATATCAATG CCTAACGTA ATGATATCAA	1080
20	AACAATTTTA GTAATAGGGT CTGGGCCAAT TATCATAGGT CAAGCAGCTG AATTTGATTA	1140
	TGCTGGAACA CAAGCATGTC TAGCTTTAAA AGAAGAGGGA TATCGAGTTA TTCTTGTAAG	1200
	TTCAAATCCA GCGACAATCA TGACTGATAA GGAAATTGCG GATAAAGTAT ATATCGAACC	1260
25	GTAACTCAT GATTTTATAG CGCGAATTAT ACGTAAAGAG CAACCTGACG CTTTACTTCC	1320
	AACTTTAGGT GGTCAAACAG GTTTAAACAT GGCGATTCAA CTACACGAAA GTGGTGTGCT	1380
	TCAAGATAAT AACGTCCAAT TATTAGGAAC TGAGCTAACA TCAATTCAAC AAGCAGAAGA	1440
30	CCGTGAAATG TTTAGAACAT TAATGAATGA TTTAAACGTT CCTGTACCAG AGAGTGACAT	1500
	TGTAAATACA GTAGAGCAAG CCTTTAAATT CAAAGAGCAA GTGGGATACC CGCTAATTGT	1560
35	TAGACCGGCA TTTACGATGG GTGGTACCGG AGCGGTATT TGTCATAATG ATGAAGAATT	1620
	ACATGAAATC GTCTCAAATG GTCTTCATTA TAGTCCAGCA ACGCAATGTT TATTAGAAAA	1680
	ATCTATCGCA GGTTTTAAAG AAATCGAATA CGAGTAATgC GTGaTAAAAA CGATAATGCC	1740
40	ATCGTTGTAT GTAACATGGA AAATATTGAT CCAGTTGGTA TTCATACAGG CGATTCAATT	1800
	GTTGTGGCTC CTAGTCAAAC ATTATCAGAT GTTGAGTATC AAATGTTACG TGATGTTTCA	1860
	TTAAAGTTA TTCGAGCTTT AGGTATCGAA GGTGGTTGTA ATGTTCAATT AGCATTAGAT	1920
45	CCCCATTCAT TCGATTATTA TATTATAGAA GTAAATCCGC GTGTATCACG TTCATCAGCG	1980
	TTAgCTTCAA AAGCAACAGG ATATCCTATT GCAAAATTAG CTGCTAAAAT CGCGGTTGGT	2040
	CTAACATTAG ATGAAATGTT AAATCCAATT ACAGGAACAT CTTATGCAGC GTTTGAACCA	2100
50	ACTTTAGACT ATGTGATTTT AAAAATACCA AGATTTCTTT TTGATAAATT TGAAAAAGGA	2160
	GAACGAGAGC TTGGCACACA AATGAAAGCA ACAGGTGAAG TTATGGCCAT TGGTCGAACT	2220

55

	TTACCAAATG GTGAAAGCTT CGATCTTGAT TATATTAAAG AACGTATTTC ACACCAAGAT	2340
	GATGAACGAT TATTTTTCAT CGGCGAACAA TTAGAAGAGG CACAACATTA GAAGAAATTC	2400
5	ATAATATGAC TCAGATTGAT TACTTCTTCT TACACAAGTT CCAAAACATT ATTGATATTG	2460
	AGCATCAATT AAAAGAGCAT CAAGGTGATT TAGAATATCT TAAATATGCA AAAGATTATG	2520
10	GATTTAGTGA TAAAACAATA GCGCATCGCT TTAATATGAC GGAAGAAGAA GTATATCAAT	2580
	TGCGTATGGA AAATGATATT AAACCTGTTT ACAAGATGGT TGATACTTGC GCAGCTGAAT	2640
	TTGAATCTTC AACACCATAT TATTATGGTA CATACGAAAC TGAAAATGAA TCCATAGTTA	2700
15	CTGACAAAGA AAAAATCTTA GTATTAGGCT CTGGACCAAT TCGAATCGGC CAAGGTGTAG	2760
	AATTTGACTA TCGGACAGTT CACGCCGTTT GGGCAATTCA AAAAGCAGGG TACGAAGCGA	2820
	TAATTGTGAA TAACAATCCA GAAACAGTTT CAACAGACTT CTCAATTTCT GACAAATTAT	2880
20	ACTTTGAACC TTAACTGAA GAAGATGTGA TGAATATCAT TAATTTAGAA AAACCTAAAG	2940
	GTGTCGTTGT ACAATTTGGA GGACAAACAG CGATTAAATT AGCAGACAAA TTGGCTAAAC	3000
	ATGGTGTTAA AATACTGGT ACTTCACTAG AAAATCTAAA TCGTGCTGAA GATAGAAAAG	3060
25	AATTTGAAGC ACTATTAAGA AAAATTAACG TGCCACAGCC ACAAGGGAAA ACAGCTACAT	3120
	CACCTGAGGA AGCATTAGCG AATGCTGCAG AAATCGGATA TCCGGTTGTA GTAAGACCTT	3180
	CTTATGTATT AGGTGGTCGC GCAATGGAAA TTGTAGACAA TGACAAAGAG TTAGAAAAC	3240
30	ATATGACCCA GGCTGTAAAA GCGAGTCCGG AACATCCGGT ACTAGTCGAT AGATATTTAA	3300
	CTGGTAAAGA AATTGAAGTT GATGCGATTT GTGATGGAGA AACGGTCATT ATTCCAGGAA	3360
	TCATGGAACA TATTGAACGT GCTGGTGTGC ATAGTGGTGA CTCAATCGCT GTATATCCAC	3420
35	CACAACTTT GACAGAAGAC GAGTTAGCAA CACTTGAGGA CTATACTATA AAATTAGCTA	3480
	AAGGTTTAAA CATCATTGGC TTAATCAACA TTCAATTCGT TATAGCTCAC GATGGTGTGT	3540
40	ATGTTTTAGA AGTAAATCCA CGTTC TAGTA GAACGGTACC ATTCTTAAGT AAAATTACTG	3600
	ATATTCCAAT GGCACAATTA GCTATGCGAG CAATCATTGG GGAAAACTA ACAGATATGG	3660
	GTTATCAAGA AGGGGTTCAA CCATATGCTG AGGGTGTCTT TGTGAAAGCA CCAGTATTTA	3720
45	GTTTTAATAA ATTGAAAAAT GTTGATATTA CTTTAGGACC TGAAATGAAG TCAACAGGTG	3780
	AAGTGATGGG GAAAGATACT ACATTAGAAA AGGCGTTATT CAAAGGGTTA ACAGGTAGTG	3840
	GCGTTGAAGT TAAAGATCAC GGTACAGTAT TAATGACCGT CAGTGACAAA GATAAAGAGG	3900
50	AAGTTGTTAA ATTGGCACAA CGCTTAAATG AAGTTGGCTA TAAAATTTTA GCAACGCTCTG	3960
	GAAcAGCTAA TAAATTAGCT GAGTATGACA TACCTGCAGA AGTAGTAGGC AAAATTGGTG	4020

55

	TGACTAAAGG TAAAGAAGTA GAAAGGGATG GCTTCCAAAT TAGACGTACT ACAGTTGAAA	4140
	ATGGTATTCC ATGTTTGACA TCTTTAGATA CAGCTAATGC CTTAACGAAT GTAATTGAAA	4200
5	GTATGACATT TACAATGCGT CAAATGTAAA TCAATCAAAC TGTATCGGTG GGGCTGTAAT	4260
	TAACCATTTA CTAAAGAAG TTTATATTAC AGCCTCATT TTTAATGAA TTTCTTAATA	4320
10	TAAAGGGAGA CaTATATGAT GAAAGATTTA CCAATTATTG CATTAGATTT TGAATCAAAA	4380
	GAAAAAGTAA ATCAATTTTT AGATTTATTT GATGAATCAT TATTCGTAAA AGTAGGTATG	4440
	GAACTTTTTT ATCAAGAAGG TCCTCAATTA ATTAATGAGA TAAAAGAAAG AGGCCATGAT	4500
15	GTATTTTTAG ATTTAAACT GCATGATATT CCTAATACAG TTGGAAGGC GATGGAAGGA	4560
	CTAGCTAAAT TGAATGTTGA TCTGGTAAAT GTTCATGCTG CTGGTGGCGT AAAAATGATG	4620
	TCTGAGGCCA TTAAAGGATT AAGAAAACAT AATCAAGATA CAAAAATTAT TGCAGTAACA	4680
20	CAGCTTACGT CAACAACAGA AGACATGTTA CGACACGAAC AAAATATACA AACATCGATT	4740
	GAAGAGGCCG TTTTAAATTA TGCCAAGTTA GCAAAATGCAG CTGGTTTAGA TGGCGTTGTT	4800
	TGTTACCTC TTGAAAGTCG TATGTTGACT GAAAAGTTAG GTACATCATT TTTAAAGTA	4860
25	ACACCAGGTA TTAGACCTAA AGGTGCATCT CAAAATGACC AACACCGTAT TACGACACCG	4920
	GAAGAAGCAA GACAGCTTGG TTCGACGCAT ATTGTAGTCG GTAGACCGAT TACACAAAGT	4980
	GACAATCCAG TCGAAAGTTA TCATAAAATT AAAGAAAGTT GGTTAGTATA ATGGCTAAAG	5040
30	AAATTGCAAA ATCATTATTA GATATTGAAG CTGTAACATT ATCACCAAAT GATTTATATA	5100
	CATGGAGTTC AGGTATTAAA TCACCGATTT ACTGTGATAA CCGTGTTACG TTAGGTTATC	5160
35	CTTTAGTTCG AGGCGCAATC CGCGATGGTT TAATTAACCTT AATTAAAGAA CACTTTCCTG	5220
	AAGTAGAAGT TATTTCTGGT ACTGCAACAG CTGnTATTCC ACATGCAGCT TTTATTGCTG	5280
	AAAAATTAAA ATTACCAATG AATTATGTTT GTTCATCAAA TAAGATCAT GGTAAGCAAA	5340
40	ATCAAATCGA AGGTGCTAAA AGTGAAGGTA AAAAAGTAGT TGTGATAGAA GATTTAATTT	5400
	CGACAGGGGG ATCTTCAGTC ACAGCAGTTG AAGCCTTAAA ACTAGCAGGT GCAGAAGTAT	5460
	TAGGTGTTGT AGCTATCTTT ACTTACGGTT TGAAAAAAGC AGATGATACA TTTAGCAATA	5520
45	TTCAACTACC TTTTACACT TTAAGTGATT ACAATGAATT AATTGAAGTA GCTGAmAmTG	5580
	AAGGTAAAAT TTCTAGTGAA GATATCCAAA CATTAGTTGA ATGGAGAGAC AACTTAGCAT	5640
	AATATAGACA CTAGAAGGAG GAATTCAACA AATGAATGAC AAAACATCTA ATGATTTATA	5700
50	TGGGAAGATA AAACATTGTA ACGAATTTAT CAATCATTTCA AATGATTCCA ATCTATCTAG	5760
	TAGTCACGAT GTCGACGAAA GTTCAACGAA GCAAAAACAT ATAAAAAATA AAACAACAT	5820

55

TTAAAAACAA AAAAGCT

5897

(2) INFORMATION FOR SEQ ID NO: 321:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 7965 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

15	TCATTCTAAA TCAACTTATT TTCCATTGCA TAAATTGCTG CTTGTGTACG ATCGCTAACT	60
	TGTAATTAC TAAATATATG ACTGACATGT GTTTTAATTG TTTTTCAGA TACAAATAAA	120
	GTTTCTGCAA TCTCTTTATT TGTMTTACCT TTAACCATTT CACGTAACAC TTCAATTTCT	180
20	CTCTTTGACA ACTTATTCGT GTAGTGTGGT TTTTGGCTAA CTGTTTCGAA TACATCTTGT	240
	GCCTTAGGAT GTATCATTTT TTCACCGTTC ATAACCTGTC TAATAGTTTC AATTAATTGC	300
	TGAGGCTCAA CGTCTTTCAT TTCATAACCA TCAGCACCTT TATTGATTGC TGAAATTACA	360
25	TGTTTCATCAT CAACATAACT TGTTAATACC AAAACTTTAA TATCCGGATA ATGTGCCTTA	420
	ATATATTCCG TAATTTCAAT ACCATTCATG CCAGGCATCA CTAAATCTAA TAGCACAATA	480
	TCAGGGTGCT CATGCTCTTT TAAATATTCT AAAAATGTTT CTCCATCTGC AAAGTCTTGT	540
30	AAAACTTCTA TGTTTTCAaT CGTGGaTAAT AAAAATCGCA ATCCTTGtCG CACAATATAA	600
	KGGWCATCTA CTAATATnAC TTTGTTTCATG GGTATCTCC TTAAAtCAAG CTATTTTATA	660
	GGAATTGTGA ATTGTATTG TGTACCCTTT GTTGGCTGAG AATGAAAGGT CACTTTACCT	720
35	CTTAATAATT TAACTCTTTG TTTTATGTTA TTAATACCGT GTGATGAAGC TATCTGAACA	780
	TTATCGATCT CAAATCCTTG ACCATAATCA ATCAGCTCAA TATATAGTAT ATCGTTCATT	840
40	TGTTTTAATG TAAGATCCAT TTTATTCGTA TCAGCATGTT TCTTAACATT ATTAATACAC	900
	TCTTGTAATG CTCTGTATAT GTTTTCTTCG ATTTCATTAG ATAAATCGAT TAAACCTTCT	960
	ACATTTACAT TTAATTGTAT ATGCATTAAT TTAATATATG CTGTCAAAGC ATGAATTAAA	1020
45	CCTTGCTCAA GTCCAACCTGG CTTAAGTnGC CAAATCAATG CACGCATTTT ATTAAACGGCA	1080
	TTTTGACTCG TTtCCTCAAT CGTCTTGAAT GCTTGTTTAG CGATGGATTG GTTTGACATG	1140
	CCATACGCAG CATGTGctGT TAGTTTtACA GAAAATAACA TTTGATTTAC TGAATCATGT	1200
50	AAATCTCTAG CTAAACGATT ACGTTCATTA ATTTTTGCCG CTTCTTTTTT ACGGTCTGTT	1260
	AAATAAATAC GTTTGATGGC TGACCCTAAT TGAAATGCGA CAGACTCTAG CAACTCTAAA	1320

	TGACCCGATT TAAGTGGTAC CGTTGCATGA TGTGTAATAT TGTCATTTTG GCTAGGAAAT	1440
	GCTTTAGAGG CTAAGTTAAT ACGAGAACAA TTGACGATAT TCGACGCTTT CATTAGCCTA	1500
5	CGTTGATTAA ATGCTTTCAC ACACCAACAA GACCCATCTT TAATATAGTG ACAGTGGTCT	1560
	GCTGTCAAAG ATTGTGGTAG AGCCACATGT GATACAAGTT CATGTTCCACC TACGCTATTG	1620
	ATGAAAAATA TCCAGCCTGT CGTGAAATTA CTGCCCTCAA TTAAATATTT TAACGCACCT	1680
10	TGGGTCATGC TATACATTTT TGTTCCTTCG TTTAAAAATT CGGCAATCTC TTTTAATAAA	1740
	GCTAGTCGCG TCCTTTGTTC CATCAAATCG CTCCAATTCA TTTTACGTG TATTAACAT	1800
15	TATACATTGA GTTATTATAT TTTTAAATCT TAGACGTAAA CATGATAAAA TGGCCTTGAT	1860
	TACTCAATAG TTATATTTTG GAGAACTGAT TTGTGATATG ATATTAAAGA CTATAGGAGG	1920
	ATTTTATGAA ATTTAAAAATA CCAGAAAAC TTAATGACTT AAGTTTACGA GATATTTTCC	1980
20	AACAACTTAA GGTACCTAAA AAAGATTTAC ATCATTTAAA TATGTCTAAA GATATTACTA	2040
	TTAATGATAA ACCTGCGCGA TTAATGGATA AAGTGCATAC TGGCGACGAT GTATTTGTTC	2100
	CAACCATCGA TGAAAAAAGT AATTATGTTC CAAGTTATCG TTATGCACAA ATTAAATACG	2160
25	AAGACGATGA TATGGCAATC GTAATGAAAC CTAAAGGTGT TAAGACTCAC CCTAATGATT	2220
	TAAAAGAAAG CAATACTTTA ATGAATCATG TGATTTACAC TATTGATAGT GACTATGTCG	2280
	AACCAATTCA TCGACTGGAC CAGGAAACAG TAGGATTATT AATTGTTGCT AAAAATCCTT	2340
30	TAATGAAAAA AATTCTTGAT CGCATGTTAG AAGACAATGA TATTACGCGG ATATACAAAG	2400
	CAAATGTTAA GGCACTTTTA CCTTTAAAC CACAAACGAT TGATATGCCA ATTGGTAAAG	2460
	ATAAATTCCA TTCGAATAAA CGACGTGTGT CTCCTACTGG ACAGCGTGCA ATTACACACA	2520
35	TTTTAACTTC AAAAATGATA AAAGAAGCTG TGTGCCmACT TGAAATCAAG TTGGATACTG	2580
	GACGTACTCA TCAAATWCGT GTGCATTTAG CTGAAATTGG TCACCCTGTT ATTGGTGATC	2640
40	CTTTATATGG TGATTCAACG TTAAGACAAT TAGAACTTGA AAGTTACAAA ATAGAGTTTG	2700
	TGCATCCCTT GACTAAGGAA GTCATTTCCG TTTCTTTGGA TGAATAATT GATTAGTTTT	2760
	GCATGATATT LAACATGCA ATACCGCATT GTAACATAAT CAAGTATCAA CtTAAAcGGA	2820
45	TAGATGGAAA ATTATTAATT TTTTCAGATG TTCGGTTTTT TTGTTTTTTA CGATGCTTAG	2880
	GATTTTATAT TTTGATATTT TAGTAATTAT TCATTTTATA ACATCCTTGG ATAATGACTT	2940
	GTAGTCTTTT TCAACTGCGT TACGTGTATC TATGGACAAT ACATGACATC ATAAGATTTT	3000
50	TATCACAGGT TGTTTGGCCA ATACATGTAC AACAAATTCAT CATATAAAAA ATAGGTTCTA	3060
	TAATAAAACG GACTCCATGA AAAGTTTTTC CTTTTCATGG CTCTATATCA AATCAGACTT	3120

55

	CCAAACTTTA TTTTATATTA ATATTTAATT AATGAGGATC TACCATATCT TCTGGTTTAA	3240
	TCCATGCTTC AAATTGTTCT TCTGTAACAT ATCCAGTTTG AATTGCAGAT TCTTTTAAAG	3300
5	TTAAACCTTC TTTATGGGCT TTCTTAGCAA TTTGAGCTGC TTTTTCATAA CCAATATGTG	3360
	GATTTAATGC AGTAACTAAC ATTAATGATT GATTTAAATA ATTATCAATA TTCTCTTCGA	3420
10	TTGGTTCAAT GCCCACTGCA CAATTGTTAT TAAATGTTTC CATACCATCA GCTAAAAGAT	3480
	AAATTGATTG TAGTGTATTA TGCATAATAA CTGGTTTATA AACATTCAAT TCAAAGTTAC	3540
	CTTGGAAGT TGGGAACCAA CAACTGTATC ATTACCCATT ACTTGGACTG CnACCATTGT	3600
15	TAACATTTCA CATTGTGTAG GATTAACTTT ACCAGGCATA ATTGATGAAC CTGGTTCATT	3660
	TTCAGGGATA GAAATTTCTG CCAAACCAGC TCGTGGCCCT GAAGCCAACC ATCTCACATC	3720
	ATTAGCAATT TTCATTAAGT CTCCTGCTAA TGCCTTCAAT GTTCCATGCA ATTGAACAAC	3780
20	TTCATCATGC GCTGTAAGTG CGTGGAATTT ATTTTCAGAA GATACAAATG GATAACCCGT	3840
	ATTTTCTGAA ATATAATGTG CCACTTTATC ACCAAATTCa GGATGCGCAT TAATACCACT	3900
	ACCAACAGCC GTACCACCGA TGGCAAGATT TAAAATGTGC TTCTTAGATT CAGATAACAT	3960
25	TGTTTCACAA CGGTCAAGCA TATAACGCCA GCCACTAATC TCTTGTCTTA GTTTGATCGG	4020
	CGTTGCATCT TGTAATGTG TACGACCAAT TTTAATAATT GAATCAAATT TATCTTCTkT	4080
	TTCTTTCAAA GTATTTCTTA AAAGTTTTAA TGCAGGTTCT AATTTTGTTC CAACCTCTTG	4140
30	ATATAATGCA ACGTGCATAG CAGTTGGGaA TGTATCaTTC GaACTTTGAG ATTyATTTAC	4200
	ATCATCATTk GGGTGGATAC TTTCATCACT TTGATGATCT TTTAAATACA TATTAGCAAC	4260
	ATAACTTACT ACTTCGTTCA CATTCAATTT ACTTTGTGTA CCGCTTCCTG TTTGCCATAC	4320
35	AACTAGTGGG AAGTGTTTCAT CTAATTCACC TGATAAAATT TGATCACATG CGTATACAAT	4380
	GGCATCTTTC TTTGCCTCGC TTAATTTTCC TAAATCAAAA TTAGCTATTG CTGCTGCACG	4440
40	CTTTAGTTGT GCAAAACCAT AAACACTTTC GATTGGCATA CGCTCTTTAC CAACTGGGAA	4500
	ATTACGTTTA CTCTTTCTG TTTGAGCACC CCAATATTTA TCTGCAGGTA CTTCTATTTT	4560
	TCCAAAAGTA TCATGTTCAA TTCTTACTGA CATTCAATTT CTCCCCTTAT CACTGTTTAT	4620
45	TTAACTGTAG TATATCATTa AATAATTTAA TTGAGCAATT TATGATTAAA ACGTTTTCAT	4680
	AATTTGAAAT AAAAATACAC TAATCGCACG TGTTCAACCT TTATTACAGT GATACGGTCA	4740
	TACGATTAGT GTGTTATCTA TCATTATTTA GTTATTATTG AACTAAGTTT AATTACGATA	4800
50	CTTTGTTTTA GTAGCTTCAA CCGTAGCAAT AGCTGTAAGT ATATATAATA CAGCACTAAC	4860
	AATTGTCGTA TATGGATTa GAGCAACAAG CGTaCCTAAA ACTCCTGTTA AACTCGCATA	4920

55

	TACAATACCT GATTGATTAC TTTTAATGAA TGTTCGCGCA TTAACATCAT CAATTAATCC	5040
	TTTTGATAAA TTGAGTTGTA ATTTTATTAC TTTGAAAATA ACAGGTAAAT ATAATGCCCC	5100
5	AATTGCCAAT GGAAAAGCTT TAATTGATAT TAAACTTATA ATAAGTGTG CTATCAATAA	5160
	TTGAATCCAG TATTTTCCTA ACATAAATAT ATAAATCTCC TCTAATTTCA TTCTTCAATA	5220
10	GCAATCATA ATCTTGGCAT ATTAAGAAAC GCGGTTTAAT GATTTCATTA AAAATATTAC	5280
	TGATAGATGA CTTCCTTCAA TTATGTCTGG AGTAATTAAT TATCAATTCC GTTTAAATGG	5340
	TGTTTTAATA TTTAAATG AACTTTTGAT ATATTACTAT GTCTGGTACA CAAATCAATG	5400
15	TTTTATGCTT TACAAAGTTA TATTGGCAGT AGTTGACTGC AGTCCACAAC ATAGAGGCTT	5460
	CGGAATGTCA GCTTCTATTT CATGCAAGTT GGTGGAGCTC CAACATAGTG GAATTGGATT	5520
	CCCAATTTCT ACAGACATTG CAAATTGGGG AAACGGGCCA CAACTCAGA AACTGGTGGA	5580
20	AAGTCAGCTT AAAATAACAT GCAAGTTGGC GAGGCCCCAA AATAGTGAGA TCGGATTTCT	5640
	AATTTCTACA GACATTGCAA ATTAGGGAAA CGGGCCACAA ACTCAGAAGT TGGTGGAAG	5700
	TCAGCTTAAA ATAACATGCA AGTTGGCGGG GCCACAACAT AGAAAAATTG GATCCTCAAT	5760
25	TTCTACAAAC AATGTAAGTT GGGGAAACAG CCCCAACACT GAAACTAGCA GAAAGTCAGC	5820
	TTCTATGAAT ATAATAAAAA AGCTAGGTAA CAAAATGCTA CCTAACTTCA TATTCAAGAT	5880
	AATCAATCCT ATTTGATATA TGTTCTATAC TATACATTAT TTACATGATA AATAACTGAA	5940
30	TATTACACAA TTATAATACT TTACTGACTG TCTTCTTCAG AATTCTTTTC TTGATCATT	6000
	TGATCAGAGA TTTGTTCCaT TTCTTTACCT AATTCTTTTA AATCTTCAA ATCCGTTACC	6060
35	ATACTGTTTT CTTCCTCATG ATAATTTAAT TTTGGATCTT TGTCTTAGA CATAATCAAT	6120
	ACCTCACAGA TTTTAAATTA AGCAAAGCGT GACGTAAAGT AAGCTTTAAC ATCTTCAGGT	6180
	AAACCTGCAG CCGCTTCTTT ATCAAGAATA ACATTTACCA TTCTATGTGC TTTTAAATCG	6240
40	GCTGGTTCGA AGCTTGTTTT ACCATTTTCT TGATATAATT TTTCAACTAC ATCTCGTTTA	6300
	TTAGCACCTG TCACTACTAA GAAAATTTCT CTTGCTTCCA TTAGTCCTTG ACGAATACTA	6360
	ACATTTAACT TACCTTGCTC ATCGATAGAA ACAACTTGTA ATGTTAATTT CCCTTTATTT	6420
45	TCTTTAGTTT TAATCTTATC AGCGATTAAAT TCGATTGCAT CTTTTTCATA AGCAATTGGA	6480
	TAAACTTGAC CTGCTGGTAC ACCTAACGCT TCGAAATATG ATTTTTTATC GTCATAATCT	6540
	AAAATATTTA TTTGGCTAAA ATCAACAGCA TGTTTTTCaA CATTTTCTT TAATTCATCT	6600
50	AGAACTGGCG CTTGATCTGT ATCTAAATGA AAACCTGCAA TTGTAGTAGG ATTATTGTTA	6660
	AATTGCTTTC TAATAATATC AGCAGCATAT TCTGCTACAA GTTGACTATT GTCAAAGACT	6720
55		

5 GATATTAATA ATTATTATAC CCTAACTTTC AATATATCAA ACCATTTAAC TTTAACATGC 6840
 TTATACTCTA AATATAGCAC TTAAGCATCA TTTTATAAT GAAAATGAGT AAATTTTAAT 6900
 10 TCAATCCCGG TAAATCTTGT TGACGTAACG CTTCAATAAT TAACAACGCA GCAGTATTTG 6960
 ATAAATTTAA TGAACGAATA TGTTCACTCA TAGGAATTCT TAACGCTGTG TCTTGATATT 7020
 TCTCTTTCAC CCAGTCTGGT AATCCTGTCG TTTCTTTTCC AAAAATGAAG TAAAAATCTT 7080
 15 TGTCATGATT TGAAAAATCA AAATCACTAT AAGTCTTTTT ACCAAATTTT GTTAATAAGT 7140
 AATACTCGCC ATTTGTGACT TCAAAAAATG CTTCAATACT ATCATGATAC GTAATATTCA 7200
 CAAATTCCTA ATAATCTAAA CCGGCTCTTT TTAACATTTT ATCATCAGTT CTAAATCCAA 7260
 20 GAGGTTTAAT TAAATGTAAA TGTGTGTTTG TACCTGCACA CGTACGCAA TGTTACCAGT 7320
 ATTAGCTGGG ATTTCTGGTT GATATAAAAC GATATGATTT GTCATATTAC TATTCTCTCC 7380
 TTGTGTCTAA TCCTTTTATC ATTTCACTCT GAACCTCTGC ATCCTCTTGA TCATAATTAG 7440
 CATTGATAAA ATCTCTTGCT TCTTCCCCAA GAATTTGACC AATGGCCCAA TAAGCAGTTG 7500
 CTCGAATCAA CGGTCTTTCA TCTGTTGTTG CAACTTTTTT CAATTCTGGA ATTGCATCCA 7560
 25 CTTCAATAAA ATGCGCCAAT GCTAAAATAG CATTTCGTTG TATCGGCTTT TTACCACGCC 7620
 AAGCACCTGC AAGGTGACCA TATGTTTGTG TGAATTCTTT ATTAGACATA CGTAGCAAAG 7680
 GTACTAATCT TGGCTTTAAA ATTTCTGGTT CAAAATGAT GTCATCTTGT TCGGTATTAA 7740
 30 TACCTCTATT TTTCGGACAA ACTTGTTGAC ACGTATCGCA ACCATATAAT CTATTCCTAA 7800
 TTTTATAACG ATATTGGTCA GGCATATAGC CTTTGTGTTG CGTTAAAAAA CTAATGCATT 7860
 TCTGACTATT TAATTGGCCA TTCCAACTA ATGCACTTGT TGGACAACGA TCAACACAAA 7920
 35 TTGTACAATG CACCACAGCT ATCTAATAAT GGATCATCAG GTTCC 7965

(2) INFORMATION FOR SEQ ID NO: 322:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1302 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:

50 GCCCTGTTGG AGAAATCACC TTTATACGAT GGTGAAAAA GCATCATTAG GTACAATCGT 60
 GGCAATTATA GTTATTTTGC TGTTTTTAAG AAACATTCGT ACGACGGCAA TTTCTATTAT 120
 ATCGATTCCG TTATCACTTC TTATGGCGCT TATTGCTCTG AAATTGAGTG ATGTTTCATT 180

55

	TGTAGTTGTT GAAAATATTT ATCGACGCTT AACAGATTCA GAAGAACAAC TAAAAGGTGA	300
	AAATTTAATT ATCAGTGCGA CAACTGAAGT ATTTAAACCA ATAATGTCAT CGACACTAGT	360
5	TACTATTATC GTCTTCTTAC CACTTGTGTT TGTATCAGGT TCAGTAGGCG AAATGTTTGTAG	420
	ACCTTTTGCA TTGGCTATTG CATTTAGTTT ATTAGCATCG TTATTAGTGT CAATTACACT	480
	CGTTCCAGCG TTGGCAGCTA CACTATTTAA AAAAGGCGTT AAACGTCGTA ATAAACAACA	540
10	TCAAGAAGGA TTAGGTGTTG TTAGTACAAC TTATAAAAAA GTATTACATT GGTCATTAAA	600
	TCATAAGTGG ATTGTAATTA TATTAAGTAC ATTAATTTTG GTTGCAACTA TTGTATTTGG	660
	AGGACCGAGA CTAGGCACTA GCTTTATTTT AGCAGGTGAC GATAAAATTT TAGCTATTAC	720
15	TTATACACCG AAGCCTGGTG AAACGGAGCA AGCAGTGTG AATCATGCGA AAGATGTTGA	780
	AAAATATTTA AAACAGAAAA AGCATGTAAA AACAATTCAA TACTCAGTTG GCGGTAGTAG	840
	TCCAGTAGAT CCAACGGGTA GTACAAATAG TATGGCAATC ATGGTTGAAT ATGATAATGA	900
20	CACGCCTAAT TTTGATGTAG AAGCGGATAA GGTATTAAAA CATGCAGATG GCTTTAAACA	960
	TCCTGGAGAG TGGAAAAATC AAGATTTAGG AACAGGTGCA GGTAATAAAT CTGTAGAGGT	1020
25	TACTGTAAAA GGTCCATCAA TGGATGCCAT AAAATCAACT GTAAAAGATA TTGAACAGAA	1080
	AATGAAACAG GTTAAAGGAC TAGCCAATGT CAAATCTGAT TTATCGCAA CATATGATCA	1140
	GTATGAAATT AAAGTCGATC AAAATAAAGC GGCAGAAAAT GGTATTTCTG CAAGTCAACT	1200
30	TGCAATGCAC TTGAATGAAA ACTTACCAGA AAAACAGTT ACGACTGTTA AAGAAAATGG	1260
	TAAAACTGTT GATGTTAAAG TCAAACAAAA TAAGCAAACA GC	1302

(2) INFORMATION FOR SEQ ID NO: 323:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1003 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

45	ATATATATTA ATTTAAACGT GTTTCACATG TACCAGTGTT AATGACAGAT AATGCTGCGT	60
	TTAAACCACC TTCAACAAGG TTTTGTACTG CTTTCATCAGA GAAGAAAGCA ATATGTGGTG	120
	TTACTAAAAT TCTTTTCATGT TCGATTAACT CTAATAATGT TTTATCGTCA ATGTCTTTAT	180
50	TAGTCCAGTC ATTTGTGAAG TATGCTGCTT CATTTTCATA AGTATCAATC GCAGCACCTA	240
	ACAAAGTACC ATCGTTCCTT GCAGCGATTA AATCAGGTGT ATTGATGACT GCACCACGTG	300

55

5 GATAGCTTTC TTTGTTTCGCT GGAACATGTA AAGAAATAAT ATCGGCATCT TTAATAGCTT 420
 CTTTAACACT ATCTTTATAA GTTAAAAAGT CTAAATCTTT ATTAGGATAG GCGTCATAAG 480
 CTGTAATTGT AGCACCAAAT CCTGCATATA TtTTAGCTGT AGCAGCACCG ATACGACCCG 540
 TACCAATAAT TGCAACAGTC ATATTTTTTAA CTGGTTTGA CATGATTTCT GCTTGCCAAG 600
 10 TAAATCATG TGCTTGTA CA GCGCGTTCAA TATCTGGGAA GCGACGCACT AATTGTAGGG 660
 CGATAGATAC AGAATACTCT GCAATtGTTT CAGGTGAATA ACTAGGAACG TTAGATATCA 720
 CAATATTGTG TTTTTTAGCT AAATCTAAAT CACATATC AAATCCAGCA GTACGTTGTG 780
 15 CAATTTGTTT AATACCGTAA GATTCTAATT TAGGATAAAC GTCATTTTCT AACTTACCAA 840
 ATTGCATTGT AGTTACGCCA TCGTAATCTT TTAATTGAtC GACTGTAGCA CTTGATAATA 900
 GCTCTTTAGA AGTAGTTACT TCGACATTAT TCTtTtTCCC CCAATTTAAT GCCATCTCTT 960
 20 TCTCATAATC ACGCGTaCCA AAGAACATAA TTCTCGTCAT TAT 1003

(2) INFORMATION FOR SEQ ID NO: 324:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5030 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:

30 GGCTTTTATA GTGTTTACGA TTATTATGAT TACTTATATT CACGCTGTTA AAAATTATCC 60
 35 AAATAACCGT ACTGTTTCATT ATGTTTATAC AGCTGCATTT ATACTTGTTA TATTACAAGT 120
 TATCACAGGT GCATTATCTA TTATGACAAA TGTTAACCTG ATAATCGCAC TTTTCCATGC 180
 ATTATTTATC ACTTATTTAT TTGGTATGAC AACATACTTT ATCATGCTAA TGTTACGATC 240
 40 AGTAAGAAGT GACATGCAAT AACaAAAAAG cmaAACCGTAA TTTTAATGGC ACGCCCATTA 300
 AAATTACGGT nTTTTATATC AATATTTAAA AtTAAACCTA AGCCATGTAA AAACGAGATT 360
 45 ACACGTCAAT TGTTGTGTAA TCTCGTTTAA TnTTAATCAT TTTAGTCAGT TGCTTTTTTCA 420
 ATTTTCGATTA ATAAATCGCC TGTCGCTATT GTGTCACCAT TATTTACAGT TACTTGTTTA 480
 ATCACACCGT CAAATGGTGC TTGAATTGTT GTTTCCATTT TCATAGCTTC AGTAATTAGC 540
 50 AACGGCTGAT TAGCTTTCAC AGTTTCACCT AACTAACCT TGACTTCAGT TACTGAACCT 600
 GGCAATTTGAG CACCGATATG ACTTGATTA CTCTTATCTG CTTTGGCTT AACGTTTCGA 660
 55 TTTGTATGCA CATTTTCATC TTTAATGTAA ATACGTCTCG CTTGACCATT CATCGCATAG 720

	CGTTTACCTT	TATCGATTTC	GATTTCTACT	GTTTCACCAT	TACGCATTCC	AAAGAAGAAT	840
	GTAGGCGTAT	CAAGTAACGA	TAAGTTTCCG	TATTGATTTC	TAGTTTGAAT	ATATTGTTCA	900
5	TATACTTTTG	GATATAGTAC	ATAACTAATA	ATATCTTGCT	CCGTAACAGG	ACCTTGTTGC	960
	TCTTCTTCAA	GCAACTCAG	GACTTTTTC	AAATCAACTG	GCTCTAGATA	TTCACCTGGA	1020
	CGAGCTGTTA	GTGCTTCTTG	GCCTTTTAAA	ATAACCGCTT	GTAAATCTTT	ATTAAAACCA	1080
10	TTTACAGGTT	GTCCTATTTT	ACCTTTGAAG	AACGACACTA	CTGATTCTGG	GAAATCTAAT	1140
	TTATAGCCAT	CTGTAATCAC	GGATTGTTCA	TCAAGATCAT	TTTGTACCAT	ATAAAGTGCC	1200
15	ATATCACCAA	CTACTTTAGA	CGATGGTGTT	ACTTTTACGA	TATCACCAA	TAAGAAATTC	1260
	ACTCTGCGAT	ACATATCTTT	GACTTCATCA	AATCTTTGCG	CTAAACCTAA	ACTTTTAGCT	1320
	TGTTGACTTA	AATTCGAATA	CTGTCCACCA	GGCATTTCAT	GTTGATAAAT	TTCAGTATTC	1380
20	GGTGATTTGA	TATCACTTTT	AAAGTCTGAA	TAATAAGTAC	GTACAGTTGA	CCAATAATGA	1440
	CTAAGTGACT	CCATACCTTC	AATATCAGTT	CTAAGGTGGC	GTGGGAAGCC	ATTTAATGCA	1500
	TAATATAACG	AATTGGCGCT	TGGCTGACTT	GTAAACCAC	TCATTGAAGC	AACAGCAGTA	1560
25	TCAATGATAT	CGACACCAGC	ATCTATTGCT	TGTTTGTATG	TTAATAAACC	ATTACCACTT	1620
	GTATCATGAG	TGTGAAGATG	AATTGGTAAA	TCTACAGCTG	ATTTTAACTC	ACCAATCAAT	1680
30	TCGTAAGCGG	CTTTAGGTTT	TAATAAGCCT	GCCATATCTT	TAATCGCTAA	AATATGGAAA	1740
	CCTTCACGTT	CTAACTCTTT	AGCTAGTTTG	ACATAATACT	CTAAAGTATA	AATGTTTGAT	1800
	CGCTCAGGAT	TTAAATGTC	ACCTGTATAA	CAAATAGTAC	CTTCTGAGAT	TTGCCCCGCT	1860
35	TCTTGTACTG	CTTCATTGGC	AACTTTCATT	TGATCTACCC	AGTTTAATGA	ATCGAAAATT	1920
	CTAAAGACAT	CTATGCCTGC	TTTAGCACTT	TCTTGACGA	ATTTATGAAT	AACATTATCA	1980
	GGATAGTTTT	TATAACCAAC	TGCGTTTGAA	GCACGTAACA	ACATTTGGAA	TAATACATTT	2040
40	GGAATAGCTT	TACGTAGACG	TTCAAGTCGT	TCCCATGGGT	TTTCCTTCAA	GAAATTATAT	2100
	GCCACATCAA	ATGTAGCACC	GCCCCACATT	TCTAGTGAGA	AACCATCTTT	AAATACGTCC	2160
	GCTGTTTTTG	ATGCGATATT	AATCATATCC	TTAGTTCTAA	CTCGTGTAGC	TAATAATGAT	2220
45	TGGTGTGCAT	CTCTAAAGGT	TGTATCTGTT	AGTAAGACAT	CATCCTGCTT	TTTAACCCAT	2280
	TCAGCTACAC	CTTTTGGACC	TACTTCATCA	AGCAATTGTT	TCGTACCACT	AAATGAAGCG	2340
50	ATTTTACTTG	AAGATACAGT	TGGAATTGAT	GCTAATTCAT	AGTCTGGTTT	CGGACGTTTC	2400
	TCAACATTTG	GGAAACCATT	AATTGTTACA	TTACCTATAT	ATTCTAATGT	TTTAGTACCT	2460
55	CTATCTAGAG	ACGGCTGAAT	GTCGAAAAGT	TCTGGTGTTC	CTTCAATAAA	TTTAGTTGTG	2520

	TTAACACCAC	GAATACGCAT	TTCTCGTAAT	GAGCGTACCA	TTTTTTCTTC	TGCTTGTTTA	2640
	AATGATATCG	CGTGTGTAGA	TAATTTTACG	AGTAATGAAT	CATAATAAGG	TGATATCTCA	2700
5	GCACCTTGGA	AACCATCTCC	AGCATCAAGA	CGTACACCAA	AGCCCCCGCT	TGAACGATAA	2760
	GCAATGATTG	TTCCAGTATC	CGGCATGAAA	TCATTTAACG	GATCTTCTGT	TGTAATACGA	2820
	CATTGGATGG	CATAGCCTAA	TGTTGTAATA	TCTTTTGTG	GCGGCATATT	AATCTCTTCA	2880
10	CCAAATAAAT	CGGCACCTGC	TGCAACTAAA	ATTTGTGTCT	TAACAATATC	AATTCCTGTT	2940
	ACCATCTCTG	TAATTGTATG	CTCTACTTGT	ACACGAGGGT	TAACCTCTAT	AAAGAAGAAT	3000
15	TCGTCACCAG	ATACTAGAAA	TTCAACAGTA	CCTGCATTGA	CATATTTAAT	ATTTTCCATC	3060
	AATTGAATTG	CAGCATCACA	AATACGTTGA	CGTAATGTTG	ATGATAATCC	AACTGATGGT	3120
	GCAACTTCTA	CAACTTTTTG	ATGACGACGT	TGTACTGAAC	AATCACGTTT	AAATAAGTGT	3180
20	ACGATATTTT	CATGTTCTGC	ACCTATGACT	TGTACTTCAA	TATGCTTTGG	ATTATCAATG	3240
	TATCTTTCTA	TGTAAACTTC	ACTATTACCA	AATGATTTTT	CAGCTTCTGA	TTTTGCTCTA	3300
	TGGAAAGCAT	CTTCTAATTC	ACTTCTTCA	CGAACGATTC	TCATACCTTT	ACCGCCGCCA	3360
25	CCACTTGTGG	CTTTAATCAT	TAGCGGGAAA	CCAGCTTCTT	CTGCAAATTC	TTTTGCTAAT	3420
	TCATATGATT	TAATTGGACC	GTCTGTACCA	GGAATAACTG	GTAAATCTGC	CTTGATAGCC	3480
30	GTTGTACGAG	CTTTAACTTT	ATCTCCAAAC	ATATCTAAAT	GTTCTAAATG	AGGACCAATA	3540
	AATTTAATTC	CTTCTTCTGC	ACAACGACGC	GCAAATTGTT	CATTTTCACT	TAAAAATCCA	3600
	TAGCCAGGAT	GAATCGCATC	CACATTCGCT	TGTTTTGCTA	CATCAATGAT	ACGCTCAATA	3660
35	TTTAAATAAC	TTTCAGCAGG	ACCTAAATCA	CTTCCAACTA	AATAGGATTC	ATCTGCTTTA	3720
	TATCTATGTA	ATGAACTTTT	GTCTTCATTC	GAATAAATTG	CAACTGTGCT	GATGTCTAAT	3780
	TCTGCCGCCG	CTCTGAATAT	ACGAATTGCA	ATTTCTCCAC	GGTTAGCAAC	AAGTAACTTT	3840
40	TTTATTTGTT	TCAATAGCGA	TACACTCCTC	AACTATTAG	AATTTTCTAA	CTAATTAGAT	3900
	AATAAAATTT	TATCTTAAAG	CGCTCTGTTT	TGCTATAGTC	TTGTTTCTMAA	TTTTCAAAAT	3960
45	TTaACATyCT	tGAGACAATT	AAAaCCyCCG	CTTcmGaAAT	AATAAATTCA	AAAATGACTA	4020
	TGCAACAACA	GGTAGTTCCA	CGTTTTTGTT	GTGAAACATT	TTCGATTTCT	ACAACTCTAA	4080
	AAAATTAAAA	ATAAAATTGC	AAAACATCAA	CATTTATTAT	CAATAGCGAT	AACTTTATCT	4140
50	TATCATCATG	ATTCTAATTT	CGCCACCACA	TTTAGTAATT	TTTAGTCATA	AAATTTAGTT	4200
	ATAATTATAC	GTTGTTTTGT	TTATAAAATT	TGATAATCak	GAGTAATCtC	GTAATATCAA	4260
55	AACaAAAAGG	AAGTTAAGCG	TTGTTTGGTT	GcCTAACTTC	CGTTATTGAA	CTCATCcAGT	4320

TCTCGTACTA AATATTGGCT AGTATTTTTT TAATTAAATT GTCTTCTTAT ATCAACTTTT 4440
 TGTGTGTGTT TCTTTCGTTG CTGGTCTACT TTGATTGTGT TACCTACAAT CAGAAGTAAA 4500
 5 CCCATAGCAA TACTTAAACT AATCATTGAT GATCCACCAA AGCTGATAAA TGGCAATGGC 4560
 ACACCAGTTA ATGGAATTGT TGCCGAAATA CCGCCAATGT TTACAAACGT TTGACTTCCA 4620
 AAGTATGTGG CAATCCCAAC ACACACAAGT TTATAAAAAT ATGATGATGT TTTATTTGCA 4680
 10 AACTGGAAGG CACGATATAC AATAAAGAAC TCTAAAGTAA TAACTAGCAA TCCTCCGATT 4740
 AAACCTAATT CTTGCGAAAT AATTGCAAAA ATAAATCTG TATGTGGTTC TGGTAAATAG 4800
 CCCAATTTCA TTGCACTATT TCCTAATCCT TTTCCAAATA CGCCACCGTT ACCTATCGCA 4860
 15 AGCAATGAAT TGGAAATATG GTATCCAGTT CCTGATTCGA ATTGGAATGG ATCTGTTAGC 4920
 GTACTAAATC TGGCAGTTAA ATAAC TTGGT AACCAACCAG CCATTAATGC AATGACAAAT 4980
 20 ACTACTAAGA ATCCTAGCAC TGCTGGTATA CCAATCTTA GGACTTTGTT 5030

(2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 1389 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

CTTGTTAATC CGAAAATTAT TAGTCAATCA AATGAAACAA TAACAGACTT nGAAGGTTCA 60
 35 ATTACATTGC CAGATGTTTA CGGCGAAGTG ACAAGAAGTA AAATGATAGT TGTCGAAAGT 120
 TATGACGTCA ATGGGAACAA AGTTGaACTA ACTGCACATG aAGATGTAGC AAGAATGATT 180
 TTGCATATTA TAGATCAAAT GAACGGTaTC CCTTTTACAG AACGTGCGGA CCGTATTTTA 240
 40 ACAGATAAAG AAGTGGAGGC ATATTTTATA AATGACTAAA ATAATATTTA TGGGTACACC 300
 AGACTTTTCA ACAACTGTTT TAGAAATGCT TATTGCAGAA CATGATGTCA TTGCAGTCGT 360
 AACGCAACCA GATCGACCTG TTGGACGTAA ACGTGTTATG ACACCACCAC CAGTTAAAAA 420
 45 AGTTGCAATG AAATATGATT TACCTGTATA TCAACCTGAA AAATTAAGTG GATCAGAAGA 480
 ATTAGAACAA TTGCTTCAAT TAGATGTAGA TTTAATTGTA ACTGCTGCTT TTGGACAATT 540
 50 ATTACCTGAA TCATTGTTGG CATTACCAA TCTTGGGGCA ATTAATGTAC ATGCATCATT 600
 GTTACCGAAG TATAGAGGTG GTGCACCAAT TCATCAGGCA ATTATCGATG GTGAACAAGA 660
 AACCGGCATA ACAATTATGT ATATGGTTAA AAAATTAGAT GCGGGTAATA TTATTTGCA 720

55

	ATTAGGGGCA GATTTATTAA AAGAACTTT ACCATCTATT ATAGAGGGCA CAAATGAAAG	840
	CGTACCTCAA GATGATACGC AAGCAACATT TGCTTCCAAT ATTCGACGCG AAGATGAGCG	900
5	AATTAGCTGG AATAAACCAG GAAGACAAGT GTTTAATCAA ATTCGTGGAT TATCACCATG	960
	GCCAGTTGCT TATACAAC TAAGATGACAC TAACTTGAAA ATATACGATG CTGAACTCGT	1020
	TGAGACTAAT AAGATAAACG AGCCTGGAAC CATTATAGAA ACGACTAAAA AAGCCATTAT	1080
10	TGTTGCTACA AATGATAATG AAGCTGTTGC AATTAAAGAT ATGCAATTAG CTGGGAAAAA	1140
	GAGAATGTTA GCTGCCAATT ATTTAAGTGG TGCGCAAAAC AACTAGTAG GGAAGAACT	1200
15	TATATGATAG AAAACGTGAG AAGTCTTGCT TTTGACACGA TTCAAGATAT ATTAAATGAA	1260
	GGTGCGTATA GTAACCTGCG TATCAATGAA GTGTTGTGAG AAAATGAATT AAATGCAATG	1320
	GATAAGGCTT TATTTACAGA AATTGTCTAC GGAACCGTTA AAAGAAAATA TACGTTAGAT	1380
20	TTTTATTTA	1389

(2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2746 base pairs.
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

	TTTGCTAATA ACAATAAAGC CAAAGCCGAT TCACACTCTA AACAGCTAGA AATTAATGTT	60
35	AAGAGTGACA AAGTACCTCA AAAAGTAAAA GATCTAGCAC AACAACAATT TGCTGGTTAT	120
	GCTAAAGCAT TAGATAAACA AAGTAATGCA AAAACTGGTA AATATGAACT TGGCGAaCTT	180
	TTaAAATTTA TAAATTTAAT GGTGAAGAAG ATAATAGTTA CTATTATCCA GTTATAAAAG	240
40	ACGGTAAAAT TGTTTATACT TTAACACTTA GTCCTAAAAA TAAAGATGAT TTAAACAAAT	300
	CAAAAGAAGA CATGAATTAC AGTGTTAAAA TTTCAAACCTT CATCGCTAAA GATTTAGACC	360
45	AAATTAAAGA TAAAnATTCA AATATCACTG TTCTTACTGA TGAAAAGGG kTTTATTTTG	420
	aAGAAGATGG cmAAGTTAGA TTAGTAAAAG CTACGCCTCT ACCTGGTAAT GTAAAAGAnA	480
	AAGAAAGTGC TAAACAGTT TCAGCAAAAT TGAaACAAGA GTTaAAAAAT ACAGTAACAC	540
50	CTACTAAAGT TGAAGAAAAC GAaGCGATrC AAGAAGATCA AGTTCAATAT GAAAATACAT	600
	TAAAAAACTT CAAAATTwGA GaACAACAAT TCGATAACTC ATGGTGTGCA GGATTCAGTA	660
55	TGGCAGCATT ATTAAATGCa ACTAAAAATa CAGACACTTA TAATGCACAT GATATTATGC	720

	AAATGATTGA ATACGGTAAA TCACAAGGCA GAGATATTCA TTATCAAGAA GCGGTACCAT	840
	CATATGAACA AGTTGATCAA CTTACAAAAG ATAATGTAGG AATTATGATC CTTGCACmAA	900
5	GTGTATCTCA AAACCCCTAAT GACCCACATT TAGGACATGC GCTAGCAGTT GTTGGTAATG	960
	CTAAAATTAA TGACCAAGAA AAACCTATTT ACTGGAATCC TTGGGATACA GAATTATCAA	1020
	TCCAAGATGC AGATTCAAGC CTATTACATT TATCATTCAA TCGTGATTAT AACTGGTATG	1080
10	GTTCAATGAT AGGTTACKAA AAAGTAATAT AGATATTGAT TAAAGGCAGG TAAAACTATG	1140
	TATCAACTAC AATTTATAAA TTTAGTTTAC GACACAACCA AACTCACACA TCTAGAACAA	1200
15	ACCAATATCA ATTTATTTCAT TGGTAATTGG AGTAATCATC AATTACAAAA ATCAATTTGT	1260
	ATACGTCATG GCGATGATAC AAGTCACAAT CAATATCATA TTCTTTTTAT AGATACGGCA	1320
	CATCAACGCA TTAAATTTTC ATCTATTGAT AATGAAGAAA TCATTTATAT TCTTGATTAT	1380
20	GATGATACAC AGCATATCCT CATGCAAACG TCATCCAAAC AAGGTATTGG CACTTCGCGA	1440
	CCAATCGTTT ATGAGCGCTT AGTATAACTA ATTTAAATGA TTTCACCTCA TAAAGCGGGT	1500
	TGGCGAGAAT TCAATTTCTC ACCAGCTCGT TTTTTCATTG TAATAATAAT CTTTAACATT	1560
25	TATTCTTTCT CTATTAATTT TTCTCAAACCT ATCTTATCTT TATGATAATT AATTAAAATG	1620
	CCCTTTTAAA TTCTTATAAA ATAAAAAGC CACCTATCGT CGCTAATAAA CGACGCAAGT	1680
30	GACTTAATAT CATATTCAAA ATAACCTATG GGAATTTAGG GAATTGATCG AAGTCAGGAT	1740
	CACGTTTTTC TTAAACGCA TCACGGCCTT CTTTCGCTTC ATCAGTTGTG TAATAAAGCA	1800
	ATGTTGCATC CCCAGCCATT TGTGTAAAC CAGCTAAACC ATCTGTGTCA GCATTCATAG	1860
35	CTGCTTTAAG GAATCGTAAC GCTGTTGGTG AGTGTTCAT AATCTCTTA CACCATTGCA	1920
	CAGTTTCATC TTCAACTTTC TCTAAAGGTA CCACTGTATT TACTAGACCC ATATCTAAAG	1980
	CTTCTTGTC ATTGTATTGA CGACATAAGT ACCAAATTTT CCGTGCTTTC TTATGTCCAA	2040
40	CGATACGTGC TAAATATCCT GAACCATAAC CCGCATCAAA TGAACCTACT TTAGGACCAG	2100
	TTTGTCCAAA AATAGCATTG TCAGCAGCAA TCGTTAAGTC ACAAACAACA TTTAGTACAT	2160
45	TACCGCCACC TACAGCATAA CCTTTTACCA TCGCGATAAC CGGTTTTGGA ATAATACGAA	2220
	TTAAACGCTG TAAATCTAAT ACATTTAAGC GAGGGATTG GTCTTCACCT ACATAACCAC	2280
	CATGTCCACG TTTCTTCTGG TCACCACCAG AACAGAATGC TAAATCACCT TCACCAGTTA	2340
50	ATACGATAAC TGAAACGTTT TGATCATCAC GTGCACGTGA AAATGCGTCA ATCATTTCAG	2400
	CAACTGTTTT AGGTGTAAAC GCATTGCGTA CTTCAGGGCG ATTTATTGTT ACCTTAGCAA	2460
55	TCCCTTCGTA AAATTCATAT TTGATTTTCAT CATATTCTCT AAGTGTTCCT CATTGTCTGT	2520

TCCACATGAA TTGTATGACC TGTAGCAGAA ATTAATTTAC ATTTACTATT AGGAATTAAA 2640
 TTTGCCATTT TTTTCGCAAT CTGTACAAAT TTTTCATCAT ATTCTCCAGC TAATATTAAT 2700
 GTTGGTACTT TAATTTCTTt CAGCGCGGCC ATAAGTTTGG CATTTC 2746

(2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

TAATGTTTAG TTTATTAACA GTAAGTTCGT ATATCAATGT TTAGTGCTCC CAAAATTGA 60
 AGTTTGAATT TTAAAAGCAT CTTGTAGAAT TTAGTTGTAT TTTTTTCAA GAAATTCATT 120
 TTGATTATTT TTGATAATGA GCATTTTAAT AGTAATACAT GTTTATAGTG TGTAATATAT 180
 GTCTATACTA GTAGTAACTA TATAGAGAAA GTAGGAATAA ACTATGTCAC AAGATGTAAA 240
 TGAATTAAGT AAGCAACCAA CGCCAGATAA AGCAGAAGAT AACGCATTTT TCCCATCACC 300
 ATATTCCCTT AGTCAATATA CAGCACCTAA AACAGATTTT GATGGTGTG AACACAAAGG 360
 TGCCTATAAA GATGGTAAAT GGAAAGTATT GATGATTGCT GCTGAAGAGC GATATGTATT 420
 ATTGGAAAAT GGAAAAATGT TCTCTACGGG TAATCATCCT GTTGAAATGT TATTACCTTT 480
 ACATCATTTA ATGGAAGCAG GTTTTGACGT TGATGTTGCG ACATTATCTG GTTATCCAGT 540
 TAAATTAGAA TTATGGGCTA TGCCAACTGA AGACGAGGCA GTTATAAGTA CTTATAATAA 600
 ATTGAAAGAA AAATTAAAAC AGCCAAAAAA ATTAGCAGAT GTGATTAAAA ATGAATTAGG 660
 ACCTGATTCA GACTATTTAT CTGTCTTTAT CCCAGGCGGA CATGCTGCAG TTGTTGGTAT 720
 TTCTGAAAGT GAGGACGTTT AACAAACATT AGATTGGGCA TTAGACAATG ACCGCTTTAT 780
 AGTTACATTA TGTCATGGAC CAGCAGCACT ACTTTCAGCA GGGCTTAACA GAGAAAAATC 840
 TCCATTAGAA GGATACTCTG TTTGTGTCTt CCCTGaCTCA TTAGATGAAG GTGCAATAT 900

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3642 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	ATCTGAnTnG AGATAGTGAT AATGTGTCAC CcAATTTTaA AcCtTTATTT TTaAGcGtTT	60
	CaTCaGCTAA CaCtTCaTTa TcTtTAGtCG cTTTATGcCC tTcTATTAAA CtCGGaACTa	120
5	AAAATGaTga CtTTTCAACa CCGAACaCTA AAACaTTGtC tTTTtGATGG CcATTAGACA	180
	CAATTTCCCC TGTTTGCTTC AAAGTAGCTT GCTTCTTGTA TTTATTTTCA ATATCTTTCT	240
	TGTTAAAAAC AGATTGTTGC ACAGTTTGAT TGGCATCTTT ATTTAGAACA ATGGCATCTG	300
10	CTTGCCACTT ATCAATGCCT TCTTTATTCA TATTGATAAG ACCATTGCGC AATCCAGATA	360
	ATAAAAATAG CAAGTAACTA ATCATCGTTA ACACACCAAT AATTAGTCCA AACTTCAATT	420
15	TGTTGCGCCG TATTTcATTc CAAGCTAAAA ACATGCATTT CTCTCCCTAC TACTATGATT	480
	TAAACATTGT TTATATTCTT AGATGCACGT ACGTCGTGTT GCGCTCTGTA ATGTTATACA	540
	TACACTTATC CTTcATTATA CCCGAACTTT TTATATTAAA AGCAAATTTA TGGAAAATGC	600
20	AATTAATTGT CTATTATTTT TGTACGGTAC ATTTAAAAATT AAGGATCAAT TTAAAAACGC	660
	CTACATATAC CTTTAAGTAC ATGTAGACGT CCAATTCATA TATTATTAA CTTCGcCTGT	720
	TTTAGGATCG AATTGCTTAA TAGcATTTTT ACGTAATTTA TCTTTTGCTT TrTCACTTGC	780
25	TTTATAGTTA TTGTTGTAaa TCGTAGCTTC CCAACTACCA TAcATTGGGT TAGGGAAAAT	840
	GATATATTTC TTACCGAAAT CGTCTTTATG TTTTTCaATT AATGCTTCAC GAGATTcAGC	900
30	TGTAGCTTCT TTTGGATCTG TAAAGTCTAA TAAATTATCT CCAAATAGCA TGACAAGTTT	960
	ATGATCCTTT TGAACCATTt GTCTGCGTGA TTCTTTACTC TTATCATCTT TAcCTTTTAg	1020
	TAAAATATGA CTCTTCTTAG CTTGAGGGAT ACCTTGTTGT TTtAAGTTCT TTTGTGTTGC	1080
35	CTTTAAATCT TTTTCTTTAT CTCTATCAGA AATATAGTAG ATATCGACAC CTTTTTTGTC	1140
	AGCATATTTc AAGAATTCTT TTGCGCCATA GACAGGTTTA GCTTTAGCAG cTTGTACCCA	1200
	TTcATGCCAA CCTTCTGGGA AAGGTTTATT ATGTATTGAT GCATAGCCTT GATATGGAGA	1260
40	ATTATCTAAA ACTGTTTCAT CTAAATCCAA AGCAATAGCT AACTTATGTT TAcCTTTATT	1320
	CTTTTTAATC TCTTTATCTA ACTGTGTCTT TGCACtGTTA TAACCTTGTA AATATAATGC	1380
45	TTTTGCTTCA GCTGAATTTT GATACCAAGC CACTGCCATA ATATTTTGAT TACCAAGATT	1440
	CGCCTTTTGT GATGCTGGTA TAGAAGCTTG TTGCGTTTGT TGAACCTCAG CAGAACTTTT	1500
	GGCAAACGCT GTAGAATTtG TCGTTTGTGG TGCTGAAACT GTaACCGCTA CCGATAATGA	1560
50	TGCTATTGCA ATATACTTTG AAATTTTATT CATCTTATCA CCTCATGATT AATATTtAAA	1620
	ATACAGTTAA AATTATAAAT GCATTTATTT AATATTGCTA TACTATGAAA AGATATTtAA	1680
55	TATTATTCTT TGgAAAAGCT AACAAATATG TGAACATTTC ATATAAGCAT GATTtAATGG	1740

	GTGACTAGGC	CTTCCTATCA	GACATATTCA	CTCATCCACG	TATCATTATG	TGTACAGTGT	1860
	GCTATCTCTT	ATTTACCTAT	TGGAACAACC	ATAAACTCAT	CCATAGTTTA	CCTTTTATAA	1920
5	ATAGCAGTCC	TCACTCATAC	AATTTCTCAT	AAAAATCACA	ACGCTCCAAC	GTAATTTCCAA	1980
	CTTACTTTCA	CCTATTTTAA	TTCATAAAAA	CGACACTTTA	ATTGTCATTA	TCCAATAATA	2040
	GCAAGACGTT	ATTATTGCAA	TCTTTTTTAT	AAAATAATAG	AATCATAGTA	TTGTCAITTA	2100
10	AAGATAAAGT	AAGAACGTTT	TTATTTTTC	GATTTTTTAA	ATTATTATGA	ATATCTAGTT	2160
	TTAGGAAGGA	AATTACATTG	AAAAACAAG	TTATTATTTC	GGGCCTCATG	TTATTTTCAC	2220
15	TATTTTTTGG	AGCCGGAAAT	TTAATATTCC	CGCCCATGCT	TGGCCATACA	GCGGGTCAAA	2280
	ATATGTGGAT	TGGTATGCTA	GGCTTTGCCC	TTACAGGCAT	ATdACTCCCC	TTTATTACTG	2340
	TTATTGTTGT	TGCATTTTAT	GATGAAGGTG	TTGAAAGTGT	AGGCAATCGT	ATACATCCAT	2400
20	GGTTCGGGTT	TATTTTTGCT	GTCGTGATTT	ACATGTCTAT	CGGAGCATT	TACGGTATTC	2460
	CACGTGCTGC	AAATGTGCGG	TACGAAATTG	GTACAAGACA	CATTTTACCT	GTGCATAACC	2520
	AATGGACTTT	AATTATATTC	GCAGCAATCT	TTTTTGCCAT	CGTTTACTGG	ATTAGTTTAA	2580
25	ATCCATCGAA	AATCGTTGAT	AATTTAGGTA	AATTATTAAC	ACCGTTATTA	CTATTAATGG	2640
	TCGCTCTATT	AAGTATTGCT	GTCATTTTCA	ACCCTGAATC	TGCACTAAGT	GCACCTAAGG	2700
30	ATAAATATAT	AACACATCCT	TTCATTTTCA	GAAGTTTGGA	AGGCTATTTT	ACAATGGATC	2760
	TTGTTGCTGC	GTTAGCTTTT	TCCGTAGTCA	TTGTCAATGG	CTATAAGTTT	AAAGGCCTCA	2820
	CAGATCGCAT	GAAAATTTTA	AAATATGTCT	GCTTTTTCAG	TCTTATTGCA	GCCATATTAC	2880
35	TTGGAATGAT	TTACTTTTGA	CTTGCAATAC	TTGGGGCATC	AACAGCTCCA	GGAAACTTTA	2940
	AAGATGGTAC	AGATATATTG	ACGTACAAC	CATTACGATT	ATTTGGTTTC	TTCCGGTAACC	3000
	TCGTATTTGG	AATGACGGTT	ATCCTTGTCAT	GCCTAACAAC	ATGTATAGGA	CTCGTCAATG	3060
40	tTGCGCCACA	TTTACTAAGA	AACACGTACC	TAAGTTTTCT	TATAAAATAT	TCGCACTTAT	3120
	TTTctCTATC	ATAGGGTTCT	TATTTACAAC	ACTTGGTTTA	GAAATGATTT	TAAAAATTGC	3180
45	TGTCCCATTA	TTGACTTTAA	TATATCCCGT	GTCGATTGCA	CTTGACTTCA	TATCATTTGC	3240
	TAACATGTTT	AGCACATTCA	GATTCAGTTG	GGCCTATCGA	CTCGCAACTG	TTATTACATT	3300
	GATTATTTCA	ATTTTACAAA	TACTAAATAG	TTTCAACTTA	TTACACGGTG	TTATTTTGAA	3360
50	ATCGTTTATG	ATGTTACCTT	TAGCAGATAT	CGATTTAGCT	TGGCTTGATC	CATTCATGCT	3420
	CTTTGCTATT	ATCGGTTTCA	TAATCGATGT	ATTTATACGC	CGTCCGAAAC	AAGCGACAAC	3480
55	TTAATAAATG	CTCACTGCCT	AGTAATGATT	GACCCATCGT	TACTAGGCTT	TTtCATATGA	3540

TAAAAAATCC TAGCTGTTAT TCAAAAATAT TAGTTTTTAA AA

3642

(2) INFORMATION FOR SEQ ID NO: 329:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2187 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

15	TTATTGATAT TGAAAATTCA AAAACTGCAA CACATCTTAT CACTAAAAAC CAACACTATC	60
	ATCAAAGTGA CACTCATTTT GAACAGTATA AGAAATTTAT TTAGATTCA GGTATATCAT	120
	CAACACAATT TGTATATAAT AACCTGTCTG TAAGCGGATT TAAATATACT AATGATGGTA	180
20	AGAATCCAAT TCAATTATCT GACATAGTGT ATCACTTAAT CGCATTATTA CGATATGGCG	240
	GTGGCATTAG CTATCAATTA TTAGATGACC ATTCAAATTA TATTCCTTG TACAACAAAT	300
	ATGGTAGCCC CCTACCATTA ATGCATTTAT ATAAATGTT TAGACCTTTT GTTAATGAAG	360
25	ATATTGAAAT TACAAATAAT TATGTATTGA GTCGTAAAGA TAATAACTAC CATTTCTTAT	420
	TATTCATATA AATTAATGAT CGATATATGT CAGACGTAAA ACAAGATTTT ATTTCCATA	480
30	ATGAATTACC TCAAGACTCT TTGATGATTA TTAAACATT GAATCATGAA CATGGTTCAA	540
	TTCAACATTT GCTTCCAATA AGCGATCAAC TTGTTTATAT AGAAAAAGAA ATTTTAGATG	600
	AATTAGACAA AACGAATTAC CCTAAAACGG AGCTGTCAGT TCAAGAAGAA ACTGGTCGTA	660
35	CATTTGAACT CAAGTTAAAT CACGACGAGG TTAAATATAT TTGCTTTAAA CCAAGCTAAA	720
	TACTAACAGT CCTCTTGTGT TTAGTTTCTT ACGTTAAAGG CTATTTATAT CATAAGGAGA	780
	TGATTTGTAA TGAGTAACTC ACAAGCAATT CAAGCAATTG AAAACGTGTT AGTAACGTCA	840
40	AAAGTTGGTG TATTATCAAC TGCATATAAT AATAAACCCTA ATAGTAGATA TATGGTCTTT	900
	TATAATGATG GTCTTACTTT ATATACTAAA ACGAATATCC ATTCTGCTAA GGTCAAAGAA	960
45	ATTAAAGATA ATCCAGCAGC ATATGTTTTG TTAGGCTATA ACGACACAAC TAATCGTAGT	1020
	TTCGTTGAGA TGGAAGCGAC AATCGAAATC GTTACAGAAC AAGAAGTGAT TGATTGGCTA	1080
	TGGGAAACAC AAGACAAAAG CTTTTCAGT TCAAAGAAG ACCCAGAGTT ATGTGTTTTA	1140
50	AGAGTAGTTC CGCAATCCAT TAAGCTAATG AATGATAAAT CATTAGATAC ACCTATCAAA	1200
	ATCGATTTAT AACACAAAGT GTATATAGGA AATAACTTTT ATGAATTCTA GATATAACAA	1260
55	TGTTAAATAC TTAATAAATC TCGCTATAAT TAAAGTGTTT AATATGTTTA CAATTCAATT	1320

TTTAAATTAA TTTTATGTAA TATAAATACT GCATTTGCAA ACTGTTGCAC TTTTAGGTAT 1440
 AACAGAATTA ACTACATTTA AGGAGATTGA TGAACCATGA AAAAGAAAAA AGGTTTTGGT 1500
 5 CTTGGTATTA GTTTAATCGC CATCATGTTA ATTGTATGTA TTGTATTAGT AATCATGATG 1560
 ATGACTGGCG GAAAGAAAGA TACATACTAT GGAATTATGA AAGATAATAC TACTATTGAA 1620
 AAAATGATTA GTGAAAAAGA TGAAAGTATT GAAAAAATG TTAAATTACC TTCAGATTCA 1680
 10 GATGTTAAAG TTAAAAAGG TGATTTTGTA ATTGTTTATA AATTAGCAGA TTCAGATAAA 1740
 ATTGTTAAAG TTAAAAAGT TGACCATGAC GATGTACCAC ATGGTTTAAT GATGAAAATT 1800
 15 CATGACATGG GCAAAATGCA CATGAAACAC TAATTGTAAT TTAAATTACA AATTTTAGTT 1860
 GCCATCAAGG TATATACGAG TAAAAGCAGC GGTAAGTTGA TTTCCAATTT GGAATCATTT 1920
 TACTGCTGCT TTTTATATTT GAAATACTTT CATATTGAAT AGCTCCACTT GCCGTTTCGCC 1980
 20 TgcgCTTTGC GCATGCATAA AAGCCCCTAA CAACCTGAGG TCACTGCGCT CCGGTTTCGCC 2040
 TGcgCTTTAG CGCATGCATA AAAGCCCCTA ACAACCTGAG GTCACTACGC TTCGGTTCGC 2100
 CTGCGCTTTA GCGCATGCAT AAAAGTCCCT AACACCTGA GGTCACTACG CTTCCGTTTCG 2160
 25 CCTGCGCTTT AnCATGGCCA TAAAAGC 2187

(2) INFORMATION FOR SEQ ID NO: 330:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1788 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

CCnCCTTTTT AAACCTGGnG AAATGtmcAG tTTTGATGGt ATTGGGtTCT TTAGTATTAT 60
 40 GCTTAAGATA GAGTGAATA CGCTCCTGTT GTTCTTTATA GTATATTGCT TTTTGTTTTT 120
 CTTTTTCGTC ATATTTCACT TTAAATAAA TGA CTGATGC AACTATACAT ATACATAAAA 180
 TGACACCAAT AATTATAAAA ACATGTTTTT TGTTCAATTA AGTAACTCCT AAAATGTGGT 240
 45 GGAAAATGAA AAAATTTTTA TAATCTATAA TTATGAACAT TACAGATTAT AAACCATAAC 300
 ACTAACATCG TCGCCTTCAT TAACTTGATT GTTAAATCA GCAATTACTG AGAATTGTAC 360
 50 AAGGAATAAT TGCCTATTAT GCCCTCATGT AATTATTGCC TTAATAACAT TAACAAAATC 420
 GATAGCTATT ACATTAAATG CCTATACCCC AGACCTCAA CACCTTTTTA TACAGGACGC 480
 ACTGTAATTT CATyyACGTT AACGTGCTTT GGTGTGTGTA ATGCATATAA TACTGCTTCT 540

5 GTATCTACCA TTCCTGGAGA AATGCTTGT ACTTTAACGC CTGTCTTTGC CAACTCTTTT 660
 TCTAATCCTT GAGTAATAGT GTGAACTGCT GCTTTCGTGC CACTATAAAT CGTACTACTT 720
 10 TTCGTTACTT CAAAGCCAGA AATAGATGcA ATGTTAATAA GATGGCCACT TGATTGTTCT 780
 AACATAGTTG GTAATGCAGC CTGTGCCGTA TATAAAGTGC CTTTGATATT CACATCAATC 840
 ATACTATCCC ACTCATCTAC TTGATAATCA GTAATCTTAG ACGACAACAT TTGCCCCGCA 900
 CTATTGATAA CAATATCCAA ACCACCGAAT GTTTGTTGTG CAATTTTTTAT CAATTCATCG 960
 ACTTCTTCTT TATTCGTTAC ATCTGTTGGC ACTACCTTCA CACTATCTTG TGACAATTGA 1020
 15 TTCGCTACGT TTTGTAATTT TTCTTTATTT CTACCTGCTA AGACAACTTT TGCCCCCTTCT 1080
 TCATGTAGTA ATGTTGCAAT TGCTTCTCCA ATACCACTAC CTGCACCTGT AACTACTGCT 1140
 ACTTTATCTG TTAATACTGT CATAATGATC GACTCCTTTG ATTCTTTTTA TTTTTTCAGG 1200
 20 GTAAATCATA AATACATATT ACTTTTAAAA AGCGTATCAC AATTCATATA ACGGTCATAA 1260
 TAACTCGCTT CATTTTCATA GATAAATTAC ATTACAAGCC ATTCGAAACA TACAATTAAT 1320
 CGTTGCTTAT ATTTTTTATT TTTAAAAATG TTGAAAAATC GTCACCTTCTT TATTGTAAAA 1380
 25 ACATTATATT AGTAATAAAG TTAATACTGT GnATTTaTCA TTCGATTGAA TGATTAGAGG 1440
 GAGGAATAAA ACgTGACATA TCATGAGCGT GTTTTAGCAT TAAGAGCAGA AAGTAAAGA 1500
 30 ACCGCATTTG ATTTTCGATT CGAAGATTTA TTTAGCAAAG AAGAATGGcT AAGTATGTCT 1560
 CTTGCAGAAA GACAAAAAGC TGAAAAAGCA TTTGACACG AgTTAAAAAT ATGGACGATG 1620
 TAAGAATGCC CTTCTCAAGT GTCCATGACG CCCAAGTAAA ATTATATAAT GTTGTATATT 1680
 35 CTTATAACGG CaTTAAACGT AATTTTAAAC AAGTTGAAAA TGGAAGGATT CTAATATCAT 1740
 TTCGTTTATA TATnGCAGAC CATGGATAGA ATTTTnTATG GTnAATCC 1788

(2) INFORMATION FOR SEQ ID NO: 331:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1341 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

50 TAAGCCAAAT CCATTGTTAC CAGAAGAAGT TCGCTTCATC TTAAAACAAA TGGGTCTTAA 60
 AGAAAAGACT ATCGATGTTG AACTCGAAGT TGGCGAGCAA GTTCGTATTA AATCAGGTCC 120
 55 ATTTGCGAAT CAAGTTGGTG AAGTTCAAGA AATTGAAACA GaTAAGTTTA AGCTAACAGT 180

5 GCTTTAATTA ACAATTAAAG TTATTAACT AACCAAAGA TAAAAAGAG TATTGATTTT 300
 TTAATTAGAA AAGTGTTAAA ATTATGTGGT cGcGCTTTTA GAGCGCCCAT TTCGTcACGA 360
 AATGTTAAGA GTGGGAGGGC AAAACTGAGC CCTGTGACCA CATCACGATA TCAAGGAGGT 420
 GCACATCGTG GCTAAAAAAG TAGATAAAGT TGTTAAATTA CAAATTCCTG CAGGTAAAGC 480
 10 GAATCCAGCA CCACCAGTTG GTCCAGCATT AGGTCAAGCA GGTGTGAACA TCATGGGATT 540
 CTGTAAAGAG TTCAATGCAC GTACTCAAGA TCAAGCAGGT TTAATTATTC CGGTAGAAAT 600
 CAGTGTTTAT GAAGATCGTT CATTACATT TATTACaAAA ACTCCACCGG CTCCaGTATT 660
 15 ACTTAAAAAA GCAGCTGGTA TTGAAAAAGG TTCAGGCGAA CCAAACAAAA CTAAAGTTGC 720
 TACAGTAACT AAAGATCAAG TACGCGAAAT TGCTAACAGC AAAATGCAAG ACTTAAACGC 780
 TGCTGACGAA GAAGCAGCTA TGCCTATTAT CGAAGGTACT GCACGTAGTA TGGGTATCGT 840
 20 TGTAGAATAA TTTTACGAAT ATTAAATTG ATTACATGAT TTAAACGATG AAGCAGATAA 900
 CAGAGATAAT AATGATGAAT TATAAATATA ATCTGAATGA CTAGATTAAT GATTGATTTA 960
 TTCATAAGAT TAATTCTTCT GTTGTCTGcy CTTAACTTGC ATATAGCAAG TAATGTGGGA 1020
 25 GGAAATTCCG CTAAAACCAC TAAAGGAGGA ACTATAAATG GCTAAAAAAG GTAAAAAGTA 1080
 TCAAGAAGCA GCTAGTAAAG TTGACCGTAC TCAGCACTAC AGTGTTGAAG AAGCAATTAA 1140
 30 ATTAGCTAAA GAAACAAGCA TTGCTAACTT TGACGCTTCT GTTGAAGTTG CATTCGGTTT 1200
 AGGAATTGAT ACACGTAAAA ATGACCAACA AATCCGTGGT GCAGTTGTAT TACCAAACGG 1260
 AACTGGTAAA TCACAAAGTG TATTAGTATT CGCTAAAGGT GACAAAATTG CTGAAGCTGA 1320
 35 AGCAGCAGGT GCTGACTATG T 1341

(2) INFORMATION FOR SEQ ID NO: 332:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5136 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

CTCTAAATCT TCAATAGGTA ACTTCATTTT AATAATCCC ATTTGAACAA TTGCTTCTTC 60
 50 ATCATTGACA ATTTCTAAAA CTTCAACCCTT TTGACCGTAA GAKAATACTT TkACTTCATC 120
 ACCAGCAACA ATTTTATCGT ATTTTGTCTT TTGTACATTT TGCTTTATAG ATTTGCTTTC 180
 ATAATGATCA TCTAATCGTT TCTTCTTATC AATCAATTCA TGTTCTTTAA CATCTGCACC 240

	CTTAATCTTC TGATTTGCTT TKTCTTTTCGC TTCCTCTATT AGAGACTTTT CATAATTTTG	360
	GAATTGTTGA TACTGCTTAG ATAAATCATC ATGCACTTGT CCGCKTskTT TACAAGACGG	420
5	TCCAGKTCTA ACCTCTGTGT CTCTACACGT TTGTAATTAC GCTCTAATGA TTCAATCATT	480
	TCATTTATTT CTTTTTCATC AGTACCAATC ATCGTCTTAG CCTTATTAAT AATATTCAAA	540
	CTAAGACCTA ACTTTTTAGA AATGTCAAAA GCATTTGAAC GACCCGGCAC ACCCATTAAAT	600
10	AACTTATACG TTGGACTCAA AGTATCTACA TCAAATTCTA CACTCGCATT CATAACGCCT	660
	TCTCGATTAT AACTATATGC TTTAAGTTCA GGATAGTGCG TCGTTGCCAT TACTAGAGAA	720
	CCAATTTTTTC TAACATGATC TAAAATGCTC ATTGCTAATG CAGCACCTTC ACTCGGATCT	780
15	GTACCTGCAC CTAATTCATC AAATAAACT AAATATGTT TGTCTGCATG CTTTAAAT	840
	TCAACTATAT TCGTCATATG AGATGAAAAA GTTGATAATG ATTGTTCTAT TGATTGTTCA	900
20	TCTCCGATAT CGCAATATAC ATTTTTAAAT AACTTAACT GACTACCATC AAGTGTGGGA	960
	ATCAACAATC CTGATTGAGC CATAACAATA ATTAAACCTA ATGTTTTTAA TGTTACAGTT	1020
	TTACCACCTG TATTCGGTCC TGTAATAATT ACCGTTTCAA TATCTTCCAT AAATTCGATG	1080
25	GTATTAGCTA CAACAGTCTC ACGATTTAAT AATGGATGGT ATGCTTTAGG TAAATATACA	1140
	GTACGGTCCCT CTTTAAATAT CGGCTTTGTT CCTTTAATAC TTCTACTATA TCTCGCTTTT	1200
	GCGATTAAAA AATCTAACTG ACCCATGACT TGTTCTGCCA CAAGTAGTGC ATCTTTGTCC	1260
30	GCAGCCACAT AACCAGTTAG TTGCGTTAAA ATGCGTTCTT TTTCAATTGC TTCGTCATGA	1320
	CGTAATCGAC TAATTTGATT ATTCATTTCA ACAACTGATG ATGGCTCAAT ATACAATGTT	1380
35	TGTCCTGAAG CAGATTGATC ATGTACAATC CCATTAAAAT CTTGTCGATA TTCAGCTTTG	1440
	ACAGGTATAA CGTTTCTTTC ATTCCTAACT GTTACAATAG CATCTGATAA TTTTTTCTGA	1500
	TTTGCTTGGC TTTTAACAAT ACGGTCCAAA TTTTGTCTAA TACGTTGATT CGTGCTAGAA	1560
40	ATTTTACTTC TAATCCCTTG CAATTCATAA CTCGCATTAT CATATAAATC ATACGTATCG	1620
	CATGTTTCAT TTATTGTTG AAAAAGATCA GTTAACACAG GTAATTGATT CATCTTGTC	1680
	TCTAATATTG GGTATTTAAC ACCTTCATCT TCTTCAACCA ATTGATTATA AAATGTCTTG	1740
45	AATTGATTTT GTAATTGAAT TAATCTTTTT ATCAAGTTAA GCTCTGATAC ATTTAAACG	1800
	CCGCCAATAT CAGCGCGATG AATGAATGCT GATACTTTAG ATAAGCCACT CAAGCTTGGT	1860
50	AAACGATGCT TATTATAGAT TTGAGCAATC TCATCCGTTT CTTCCATTTG AAAAACAACC	1920
	GTTTCAAAAT TAGTAGCTGG CATCATTTGA TTGACCTTTT CCAAGCCTAA GTCACATAA	1980
	GTTTCATTGG CAACGAGTGA TTTTATTTTT TCAAATTCTA AGACGTCTAA TGTTTTTTGT	2040

55

	TCACGCGATA ATGCGTTAAT CACTCTATCT TTTGTTACAA ATCCTTTTTG CGCAGTTGtA	2160
	CGCCATAATT CATAAAATCT AAATGATTTG TATGATGCGC ATCAGTGTTA ATAGTTAATT	2220
5	TCACATTTGG ATATTTACGA ACGATATCAG CGCTCAGATC CAGTCGATGT GGATTGGCAT	2280
	TAATTTCTAA TACTGTATTC GTTTCTTCAG CTAATGCCAT TAATTGTTCA ATATTCGGTT	2340
10	TATAACCATC TCTTCTACCT ATAATACGCC CTGTTGGATG CGCTATATGT CGCACGTATG	2400
	GATTGCGACA TGCATTAGCT AATCGTTCCA TAATTTGTTT TTCTGATTGG TTAAAGCTTT	2460
	GATGAATAGC TCCAATTACA TAATCAAGTT GTGCTAAAT TTCATCATCA TAATCCAGCG	2520
15	AGCCATCAGG TAATATATCC ATTTCTGTAC CTGAATAAAT ATCAATTTCA CTATATTCTT	2580
	TATCTAAAGC CTTAATTTCT TCGTTTTGTC TTAAAAGTCT TTCTACTTGT AAGCCATTAG	2640
	CAACACGTAA ACTTTGTGAA TGATCAGTAA TTACCATGAA TTTATAACCT TTTGCGATAT	2700
20	TTGCTTCTAC CATGTCTCGA ATAGAAAACG CACCATCACT ATACGTTGTA TGCATATGAA	2760
	TATCACCATT AATATCATCT ATTGTAATGA TATTACTTAG ATCTTTATCA AATTCGCTAC	2820
	CATCTTCTCG CATAGCAGGT GGTATAAAAT TCACATTAAAT ATGTTTCATAT ATCTTGGCTT	2880
25	CACTATCATA TTGAATTAAT GTACCATCAG CTTGTTCAAT TCCATATTCA CTTACTTTTT	2940
	CATCACGTGC TTTAGCAAGT TGTGGAATC TTATATTATG TTCTTTTGAC CCAGTAAAT	3000
30	GCTGCAATGT ATGATAAAAA GCACTTGGTT CAATTAATCG AAAATCGACA CCAATCGTTT	3060
	CATCATCATA CGCTAATTCT AATGAACTT TTGTGTTCCC CACTGCAACT TCTTTTACTT	3120
	TATTGGGAAT ATTTAATAAT TGCTGCTGCA CTGCTTTTGG GTTATCGGTA CTTATTATGA	3180
35	AATCTAAATC TTTGCTCATT TCTTTAAAC GACGGAAGCT TCCTGCAGAT GAATATTGAT	3240
	CGATATAATT TAATGTATCT ATATAATCAA TGATTTCTTG ATTAAGTCTT CTCATTGAT	3300
	CAATTGGATA TCTATCTTTC TTAGCACCAA GTTGTTTCAC AGCTTCTAAT ATGTTTGT	3360
40	CCGTTTTCTT AGCAAATCCG CTTAATTCAC TAACTTTTCC ATTTTCACAA GCAACTTGAA	3420
	GTGACGCTTT ATCAACAATA TTCAACTCTT TATATAGCTT AGCAATTTTC TTGCTTCCAA	3480
45	GTCCTTGAAT TTTCAAAAGT GGAATAAGAC CTTCCGGAAC TTCTTCTGT AATTGCTGTA	3540
	AATACTGAGA TTCACCGGTC TCACGGTAAT CATTGATTAC TTCTGCAACA CCTTTACCAA	3600
	TGCCTTTTAA CTCCGTtACA TCAGATATTT CATCTAATGG TCGTTCATCT AATTCAAGAC	3660
50	TTTGAGCTGC TTTTCGaTAC GCTGaTATTT TAAAAGTATT TTCCCCTTTT AATTCCATAT	3720
	AAGTAGCAAT TTGTTCTAAT AGTTTGATAA CATCTTTTTT TGTCATAATA ACACTCCATA	3780
55	AAAAGAAGAC CAGGACGTAT CATTAAATATA TACCTTTGTC CTGACCTCTT ATGATAATTT	3840

TAGATATTTT AAGCTGATGT TGTAATGCTT CGTTAGGATA TAATGCCAAT AGATATAACG 3960
 TAAAGTGTA GACAATTATC GTCATAAACA CACCAACTAT CATTCCCATT GCACGACTGA 4020
 5 AAATATGAAT GTTTGTATAC GCTATTATTT TATCAAAAGT TACGATAATT AGATATAAAA 4080
 TGAACCTACA AAACAATGTA ATCATTAAAA AAGCTACAAT CGCTTCAAAT CGATTTTGTGTA 4140
 10 GATGATTAAA ATGAAACGCA AAAGTTGTAT TAAATGCTGT TGTTTTAGGA TATGGAATAA 4200
 ATACAATTAA TCTTCTACA ATAGATTTGT AAAATTGACT GGCAATCCAC AATGATACAA 4260
 nCGTTGCACT CAAATGTATC ATAGATAACC AAAAACCTCG TCTGAATCCA ACGATGACAA 4320
 15 AATACACAAA GAAAATGATT ATGATAAAAT CAATGACCAT TTATTGCTCA CGCTGCTGCA 4380
 ATTTGTGAAT TTGTTGTTTC AAACGTCGAT TTTTTTCTT CTAGTAGTAC TTTTTCATGC 4440
 ATAATATTCA CAGCAGTTAG TATTGCTTTT CTTGAAGTAT CTAAACCTGC TGCTTTATAC 4500
 20 CCTAATTCTT TTATTTTATC ATCAACTAAA TGTGCTACAT ATCGTATGTG CTCTGGGTTA 4560
 TCTTCCCCAA CAATTGTAAA AAGCTGATCA TTAATTGATA CATTTACCTT GTTTTTAAAC 4620
 TGTGTCATTT ATAATTTCTC CTGATCCTTT TTTTAAAATC TAAATTCACG TTATAAAATA 4680
 25 TGACTGGATA GTTTGTCTGA ATTTGATACT AATATTGTTA TATTGCAATT ATGATAAAAC 4740
 AACAACACAA TCTCTATAGA TGACTTAATG TTCTTTTAT AATGAAATAA TGTAAGAAT 4800
 TTTCTATTCA ATACTTTATC ATGTTTAAAT TGTGTCAC TAACATTTTC ATAAACATTA 4860
 TACATGACCA CTATGTATTT TGTAAGTATC CGCAATTAAT TCTTTACAAC ATACATAAAT 4920
 GTTTCTGACG TTATTATCAT TTATGATATG ATTATTTTTG CTAAAGACAA TGAAATTTTA 4980
 35 TGAAAGGATT TACACAAATG GCGAATATCG TTTTAAATTT GTCGGATAAA GACATAACGA 5040
 CATTAAATGTC ACGCATTCTT TTTGATACTG AGAATTTACC TCAAGGnATG AnAGCACGTG 5100
 CAAAGTATCC AAAATACAAC TGTAATATT TACCAT 5136

40 (2) INFORMATION FOR SEQ ID NO: 333:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 4239 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

GGCCAAAATT GCACCTCCAC TTTCTTTTGA ACAAATCTAT TTAAACGCA CATTCCCATT 60
 ATGACAAAAT AAATGTGAGT nACATTTGTT TGTATTTTAA CATGACTACT AACGCAAACT 120

55

	GGTATCTTTC AAAGATAAAA TCTTAATAAT TTCTTAGTAA ACTCTTTTCT CTAGATTTAT	240
	CACAATATTA TATAGACCTA TTTTATTTTG ACGTAAGTTG CTAGTATCTT CAAACAAAAA	300
5	ACCTTTATAA AAATTCATAC CTTTATGCTA TCGCTGTAGG CTCATTAACT TGTTACATAT	360
	AATTCTTAAC TATCCTTTGA TGATTGTTTT ATTAGATTGT TTCGTTGATG GATACTTTCA	420
10	CGAATTTCTA TAGTTCAATG CTAATAAAAA AACAGCCCAA AACTTTAATT TGTMTTGGAC	480
	TGTTTTATAA TTATGCTTGC GATGGTGTTC TAGTTTCTGA AGTTTGTTCA GCAATGTCAT	540
	ATTTAAACTC TTTACCATCA TGATCTACTG TAACTTTCTT ACCTTCAATT TGATTACCAT	600
15	CTAATATTAA TTCACTTAAA TTATCTTCGA TAGTTTTTTG TATCGCTCTA ATTAATGGTC	660
	TTGCACCATA TTCTGGATCA TATCCTTCTT CTGCGATTTT GTCTTTTCGCT TTATCAGTTA	720
	CAATAATATT TATGTTTTGT TCAGATAATC GATTTGTTAA TTTATTAACC ATCATTGTTA	780
20	CAATTTCTTT TAATTCCTCT TTTGTTAGTT TATGGAATAC AATGATATCA TCTACACGGT	840
	TTAAAAATTC TGGACGGAAT GAATTTTTTA ATTCTTTTAA CATCGTTTTT CGAATTGTTT	900
	CATAATCTTG TCCATCACTT GAACCACCGA ATCCAGCAA TCGTTGATCT TGTAATTCTT	960
25	GTGCCCCAAC GTTTGATGTC ATTATGATAA TTGTATTTCT GAAATCAACT GTACGTCCTT	1020
	TTGTATCTGT CAAATGTCCA TCATCTAAAA CTTGTAATAG AATATTAAAT ACATCTGGAT	1080
30	GAGCTTTTTT AATTTTCATCA AATAAAATTA CAGAATATGG TTTACGTCTA ACTTTTTTCAG	1140
	TTAATTGTCC ACCATCATCA TGACCAACAT ATCCTGGAGG AGCACCAACT AATCGGCTCA	1200
	CTGCGTGTTC TTCCATAAAT TCACTCATGT CTACACGGAT CATCGCATCA TCATCGCCAA	1260
35	ACATTGATTC AGCTAAAGCT CTAGCTAATT CAGTTTTTACC AACACCAGTT GGTCCAAGGA	1320
	AGATAAAGCT ACCAATTGGT CGTTTAGGAT CTTTTAACCC TGCACGGGCA CGTCTAACCG	1380
	CTTTACTGAT TGAATTAACA GCATCTTTTT GCCCAATAAC TCTCTCATGT AATGTATCTT	1440
40	CTAGACTAAG AAGTTTTTCA GATTCTGTTT CATTGATTTT AGTTAATGGG ATACCTGTCC	1500
	ATCCTGCAAT AACTTCAGCA ATATCTTCTT CTGACAATGA AGTTGACATG CCATTTTGTG	1560
45	CATTCTTCCA TTCATTTTTA GCTTCTTCAT ATTGCTTTTC AAGTTTTGTT TGTTTATCAC	1620
	GCAGgTTAGC AGCATTTTCA AACTCTTGAG CATGTACTGC GGCATCTTTT TCATTTTTAA	1680
	CTTTTTCAAT TTCTTGTTCA ATTTCTTTTA AATTATTAGG TGTCGTATGA CTCTTAAGTC	1740
50	TTACTTTAGA ACTTGCTTCA TCAATTAAAT CAATTGCTTT ATCTGGTAAG AAACGATCTG	1800
	AAACGTATCT GTTACTTAAT TTAACAGCTG CTTCAATAGC TTCGTCTGAA ATATTAATAC	1860
55	GATGGTGTGC TTCGTAACGA TCTCTTAATC CTTTTAAAT AGCAACTGTA TCTACTACTG	1920

	TTTTGCGATA	TTCATCTAAT	G TAGTAGCAC	CAATACATTG	TAATTCACCA	CGTGCTAATG	2040
	CCGGCTTCAA	AATATTCGAA	GCATCGATAG	CACCTTCAGC	ACCACCAGCA	CCAACTAAAG	2100
5	TATGCAACTC	ATCAATAAAT	AGGATGACAT	TACCTGCTTG	TTGGATTCTT	TCCATAACCT	2160
	TTTTCAGACG	CTCTTCAAAT	TCACCACGAT	ATTTAGTACC	TGCAACTACT	GTTCCCATAT	2220
	CTAAAGACAT	AACACGCTTA	TCTTTTAATG	TCTCTGGTAC	CTCATTATTC	ACTAtGGCtT	2280
10	GCGCTAAACC	TTCAGCAATA	GCAGTTTTAC	CAACACCTGG	CTCTCCAATA	AGCACAGGAT	2340
	TGTTTTTCGT	ACGTCTACTT	AATACTTCAA	TTACACGTGT	AATTTCTTTA	TCACGTCCTA	2400
15	TAACAGGATC	TAATGTACCG	TCTTTGGCAA	TGACTGTAA	GTCACGAGCT	AAACTATCTA	2460
	AAGTTGGAGT	ATTATTTGAC	TTACTAGCTT	GTGCATTTTT	ATTACTCATT	TCAGGGTTTC	2520
	CTAAAGCTTT	CACAACTTGT	GCACGTGCTT	TAGTAATATT	TAAATCTAGA	TTTGCAAAAA	2580
20	CTCTTGCTGC	AACACCTTCA	TTTTCACGAA	TCAAGCCTAA	TAAAATATGT	TCCGTTCCAA	2640
	CAAAATTGTG	ATGTAATTTT	CTAGCTTCAT	CCATCGATAA	TTCAATGACT	TTTTTAGCTC	2700
	TAGGTGTATA	ATGCAATGTA	CCAACATGAT	CTTGACCATG	TCCGATTAA	TTTTCAACTT	2760
25	CTTCAATTAC	TTTATCTTCA	GTGATATTAA	AACCTTCTAA	TACTTTTGCA	GCAATTCTCT	2820
	CAGGTTCTTT	CATTAACCCC	AATAATAGGT	GTTCTGTTCC	TATATTTGAA	TGATTTAAAC	2880
	GAATTGCTTC	TTCTTGGGCA	TGTGCTAATA	cGCGCTGTGC	ACGCTCAGTT	AATCTACCAA	2940
30	ATAACATAAA	TAATGACCTC	CTACTTTATA	TGTTCTCTTA	GTATATCTGC	TCGTTTTTCT	3000
	TTTACAGATT	TGTCATCTTC	TTCATCTAAT	AAAAATGGTG	ACTGTATAGC	TACCATCAAT	3060
35	TCATTAAATT	TAAAGTTTTG	TAATTCAATG	TAATTTAAAT	CTATACCAAG	TTTAACTcGC	3120
	TTAATCTATA	AGAAGCCTCT	TCCATAGTTA	TCATTCTACA	GTTTTGTAAA	ATACCTAGCG	3180
	AGCGAAAAAC	ACGGTCTTGT	GTTTCTAATT	GATTATAAGT	GTCTAACTTT	TGTCGTATTT	3240
40	GTTTTTCTTC	ATGAATGATT	TGATTAACAA	CTTCTGTAA	TGTTTCTATG	ATTTCTAACT	3300
	CAGATTTACC	AAGTGTAAGT	TGGTTGGATA	CTTGATAAGT	ATGTCCATAA	ACTTGCGAAC	3360
	CTTCACCGTA	AATACCTCTG	ATTGTATATC	CAAACGATT	AATGGTTTGA	GCAATCCGTG	3420
45	TCATTCTTTT	CATAATAGAT	AGACCTGGCA	AATGTAGCAT	CACGCTTGCT	CTCATACCAG	3480
	TACCTATATT	GGTAGGACAT	G TAGTTAAAT	AACCAAGTTG	TTCATCATAA	CTTATATCAA	3540
50	GGCTTCGATC	TAATTCATCA	TCAATTGATG	AAGCTTGATT	ATATAAAGCC	TGTAATGTCTG	3600
	TGTCAGTTCC	CATAGCTTGA	ATACGAATAT	GGTCCTCTTC	ATTTATCATG	ACACTTAAAG	3660
	ATTCATCGTC	ATTCACTAAT	ACTGCGGCTG	CTGGTTGTTT	TATTAGTTCA	GGACTAATCA	3720

55

CAAAGTTTGG CAAGGCATCT TGTACCTCAT TTATAACTCT AAATCCCATC ATTTTCAGTA 3840
 GCATACATTA GTGGATGCAC ATGATTTTCT AAAcTACGCG CTAACCGAAT TCTAGAAGAC 3900
 5 ATAACAATTG GTGTTTCTTC ATTACTTTTC ATCCATTGGC TGATATTATC ATGAATATTA 3960
 TCGGTCAATCA TGTTCACCT CACTCTCAGC TTTTAGTGCT TTAATTTTCAT CTCTAACAAT 4020
 GGCTGCTTCC TCAAAATCTT GGATTTCAAT AAGTTTTTTC AAATATTCAT TCTTTTCTTC 4080
 10 GATTTTTTCGC TTAAAGCTA TCTTTTATG TGAAGAATGT GGTGTCTTTC CAACGTGCTC 4140
 AAATTGTCCA CCTTGAAGTC TGCGGACGAT ATCAATGATG TCATCTTTAA ATGTTGcATA 4200
 15 ACAATTAGCA CACCCAAATT TACCAACATG TGCAATATC 4239

(2) INFORMATION FOR SEQ ID NO: 334:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1245 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

CAAAACTTTG ATTATGTCGT GTCAGTTATT AATGCATACA GTAATGCATC GTTGTTATTT 60
 30 TTTATTTATT TACTATTTAT TATGTACCAA ATCAGCCTTc CAGTAAAATA GGTCAAATTA 120
 CTGATTTTCT AAATATAAAA TGCCTCCTAA TAACATACTA TTAGTACATC ATTAAGAGGC 180
 TCTTGTTGTTA TTTGCATACT AAGCGCTCAA ATTTAAATTT AAGATGAAGA TTCTTGCAAT 240
 35 AATATTTCTA TATTTGTCGA TATATCCAAT GGATCTTCCA TTGGCAAGAA TCGATTTACA 300
 ACATTTCCCTT GTCGATCGAT TATAAATTTT GTGAAATTC ATTTTATTGG GGaCCCAAAG 360
 ATTCCTGGTT GTTCaTTCTT TAAATGCGTa AATAACGGAT GTTCATCGTT CCCGTTTACA 420
 40 GATATTTTAG CTAACACTGG AAATGTTACA CCAAATTTCT CACGACTAAT TTTCAAGATT 480
 TCTTCATTAG AACCTGGTTG TCGATTGTCA AAATATTAT TCGGAAAAC CAACACTACA 540
 AACCCACGAT CCTTATATTT TTGAAATAAA GTCTCTAGTT TTTTCAATTG TTCGCTATAT 600
 45 ATACATTCTG TTGCAGTATT AACAACTAAA ATCACTTTAC CTCTAAATGC TTCTAATTTA 660
 TAAGTTAAGC CTTTATAATC ACTTACTTCG ATATCATACA CATTTCTATT ATTCATAAGA 720
 50 CACCCCTACA CAGCCTTTTT TATATTGAAT ATGTTCTTTT TAGAATGTTT TGATAAAATA 780
 AGTGCGCGTT TACACCGTGA ACACACATTA TATAGCGTGA TACATTTTTC GAGCACACGA 840
 TAAATAATGT TCGAGTTTAT GTTGTGCTC AACCTATCCG ATTTACCGTC TTTTTTCACC 900

55

CCGTTATAAC CCCTCATTAT AATCATCCTT ATTTTCTATT TTTAAAAAGA CAATTAGACC 1020
 GCTCTTTTAAA CTATAGaTTA ATACTTAAGk TAAACTCATA CATACTGATA CCATACGTTA 1080
 5 GATCTAACAA TTTAAAATTC GTTATAACTA TGGATTAAAG AGCTGCCCAA CTCATATAAT 1140
 CCTTAAAAAC TTCACATGTG ATTGTkTATT AAGCCCTCCT TTATCnTATT AAATATCCTT 1200
 10 ATAACCCTTT TAAAATTAAA CTGACACACT CATAATTGT TACAC 1245

(2) INFORMATION FOR SEQ ID NO: 335:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

CTTGTAAGG TAGTTGTTCA TTTAAATTAA AACAGTATGC TAAGTTnGTG CTTATATTTT 60
 GCAACTTCAA TTCGACAGGC TCATCTTGAT ACATGAATGC CTCAATTTCA TCATGTGATA 120
 25 ACTTTTTTCAA AATATCAATT AAATGrATAC TAAAAACGAT AAATAACATA TCCCAATTTG 180
 AATTCAGCCC TAGCGATTTT AATTTGTTTA AAATCTTATC TTTTGGAAA ATTGGATTCC 240
 TAATGTCTTG TATATCGTTA TCAGTCAAAG TTTCCCAATC TATATGTGAA TGAAGACCTA 300
 30 AATAACACTT ATCCATTAAT AATTCATATA CCGTTAATGC AGAGACATCG AAACAACGTT 360
 CTTCACTTAx AAAAACGCCA TCAACATCAA ATAAAATTTT CTTACAATC CCCACTCCAT 420
 35 TTCTGaAAAT TCAGaTATAA ATCATTCTAC TATTTGACTA AAAAAAGCGC AAACCCTATT 480
 GAAGTaGATT TGCGCTTTAG CTGTTnAAAT TTTATAAATG TnTTTCAATT TCATCAGCAA 540
 CCTGCTGTAC GTGTGTACCG ACAATAACTT GAGTTGAATG TTTGCCATTA ACAGTAACAC 600
 40 CAACTGCACC GCGTTTTTTA ATCTTCTGTT TATCAATAAT AGATGTGTCT TTAAACTCTA 660
 GACGCAACCT TGTTGCACAA TTGGTTAAAT TAACAATATT CTCTTGACCG CCTAAACCTT 720
 CTAATATTTG TATAGCATGT TGATGATATT TACTTTGTTT AATATCATTT TCACCAGGAG 780
 45 CAATATTATC TTTTACAAC TTTGGGTCAA CTAATTCATT TTCACCTCTA CCAATCGTAT 840
 TCAAGTTAAA TACTTGGATT ACTACACGGA AAATCACATA GTATAAGATG AAAAATACAA 900
 50 CACCTTGAAC AAGCAACATC AATGGATGAT TTGATACTGG ATTAATTAGT GATAACACAT 960
 AATCTATCAA ACCTGCACTA AATGAAAATC CAGCTGTCCA ATGGAATGTA GCTGCGATAA 1020
 ATAAAGATAA TCCTGTTAAT AACGCATGAA CAACATATAA GATTGGCACA ACAAACATAA 1080

ACCAACCGTA AACTTGTTTT TTCTGAGTAG TTTTAGCTGT ATGATACATT GcTAACGCAG 1200
 CCGCTGGAAT ACCGAACATC ATGATTGGGA AGAATCCCGC TTGATAGCGT CCTGTAATAC 1260
 5 CTTTTATAGC ATCTTTGCCA CTTTGGAATT TACCAATATC ATTAATACCA ATCGTATCAA 1320
 ACCAGAACAC ACTATTCAGT GCATGATGTA ATCCTGTAGG AATTAATAAT CTATTGGCAA 1380
 CACCATATAT GAAAGCTCC 1399

10

(2) INFORMATION FOR SEQ ID NO: 336:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1329 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

TATAGTTnTA TTATTTAGCG AAGCATTAAAT ACTACCACCG GTTATAACAA ATGTATTTTG 60
 CGTATTAAAT TGAATGGTAG GACCAATCAA AGTATATTCA ATCGCTGGAC CATCATTTGT 120
 25 AATTAATGAC TCGCAACCT TAAACTAAA TTGATCCATG GCACCTGCGC CTGAAAATCC 180
 AATATGTTCA TAACCTATTC TTCCTAGATC TTGTACCGTT GAAAAGAGAC CTGGTTGTAA 240
 AATCTTAATT GACATTTTCA ATCACCACCC AGTCATCAAC ATTAAAGTTG CCATCTGATA 300
 30 TATCTCTTTC GATTGTATA AATTTCTGTT CATCTATTGC ATAAAATTGT ATCCATTCTC 360
 CTGCTTCGTA CATTGACATT GGTTCACGCT CGCTGCTAAA TACTTTTAAc GGTGTGCGTC 420
 CAATAATTTG CCATCCGCCA GGAGAATCTG ATGGATATAG TCCTGTTTGA TTATTCGCAA 480
 35 TACCTACAGA ACCTGCATGA ATTTTAAACC TTGGCTGATT ACGTCTAGGT GTATGTAGTT 540
 GTTCATCAAG TcCGCCTAAG TATGGAAATC CTGGCATAAA TCCTAGCATA TATATTAAAT 600
 40 AAGGTTTACT TGTATGTTTT TCAATAACTT GCTCAACAGT TATTCGATTA TGCTTTGCTA 660
 CTTCTTCAAT ATCTGGTCCA TATGTACCAC CATATTGAAC AGGTATTTTA ATAATACGAT 720
 TGGTTTGATT CACAGCATGA ACATTTTTTT CATTAAATTT GTTAAGTTCT AAATTTTCAA 780
 45 TTAATTTAGA AGATGTTATA GCTTGTTTAT CAAAATATAT TAGAACTGCT CGATACGAAG 840
 GGACAATATC TTGAATTTCT AATATTTCTT TTTCTCGTAT CCACCGTACC ATTGCTGTGA 900
 CATTACGATA TGTCTCTTCG GATATTTTAT TTTCAAAATA AATCATAATT GTCTGCTCGT 960
 50 TAATAAATCT TACATCCACT TTAAATCCCC CTTTGTATTG CAATAAACCA GTATTGAATA 1020
 CCTTTTCATT GTATCATTGA GAAGCACAAG TTGTTTAATA AGTAATTCAA ATCGCATATA 1080

55

TTAATATTGT AACTCTTACA CTAATTTAGG TTCTGCTATC ATTCCGGTCTG ATGGAAAATT 1200
 TTTACTTTTC ATCTGTCCGA TTTTTTGATT TTGAATATAA AAAAGCACGA CCGAAGTATC 1260
 5 ATTAACACAC TTCAATCGCG CAATTAAATA ATCTATTTGA TCATTTATTG GATATTAAAC 1320
 ATTTTACG 1329

(2) INFORMATION FOR SEQ ID NO: 337:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3421 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

20 GAATAGTGCC ATTTGGAAAG ACATAAAGAA TCCAGAAGCA CCTTTAGCGA ATATGCCGTG 60
 TTCATATAGA TTTGTGAAAC TGGCATGCCC GAATTGAGTT TTAAATGCAA ATAGAATCAT 120
 GACGAAACCA ACTACTATTA AACCAATAAT TGTGCTATT TTAATGATAG AGAACCAAAA 180
 25 TTCTAATTCT CCGAAAAGTC TTGCGCTAAG TAGGTTGAAT GACATTAATA ACAGTACACA 240
 AAATAGTGCA CTTATCCAGT TTGGAATTC TGGGAACCAA AAGCTAACAT ATTTTGCCAC 300
 AGCCGTTACT TCAGCCATAC CTGTAATAAT CCAACAGAAC CAGTATGTCC ATCCGGTAAC 360
 30 AAATCCTGCA AAAGGCCCAA TATATGTATT GGTACATCT GCGAAAGATT TAAATTCAGT 420
 ATTCTGTATA ATGATTTCTC CTAAACCTCG CATAACATA AATAACATA ATCCTATAAT 480
 GATGTATGTT AATAGAATTG AAGGGCCGGT TAATGCAATC GTTTGACCAG CACCTAAGAA 540
 35 TAAGCCTGTA CCAATTGCAC CGCCAATTGC AATTAATTGT ATGTGGCGAT TGCTCAGTTC 600
 CCTTTGTAAT TTTTCAGCCA TAATACATCT CCCTTAAATA TAGATATGTT TATTATGCAC 660
 40 TTATATTGAG ATATATACAA TTATTTTCGG TAAAATGTG TAAAATTCCA TGTTAATATA 720
 CTTTGTTTTT TATAATCATA TATAATAACC AATTGAAAAT TTAATTCTAT TGTAATAATC 780
 ATGGATTATT CACATCTTGA AAAAGCTTTA ATGGTGCTAT TTGTGGCTAT TCTGTGACAT 840
 45 TTACATAGAT TTACAAAAAA ATTGTTGCAC ATATAATGCC AGTtTTTATA TTTCAAAAC 900
 GAAATGCGTT TAcTATAATA TTAGTTGAAA GCCATTTTCAT AAAGAAACAG TAAAGGGGAA 960
 ATTTtATCaTA GCmGaATTAC AAAGAGGTTT AGAAGGGGTT ATCGCmGCGG AGACTAAAAT 1020
 50 AAGTTCAATT ATTGAAAGTC AATTGACTTA TGCCGGCTAT GATATTGATG ATCTAGCTGA 1080
 AAATGCGCAA TTTGAAGAAG TTATTTTCCT ATTATGGAAC TATAGATTGC CAAACGAAGA 1140

55

	TACACATTTT	GAGGAGTATG	TTACAGATCA	CGTGCATCCA	ATGACAGCAT	TACGTACGTC	1260
	ATTATCATAT	ATTGCACATT	TCGATCCTGA	TGCTGAAAAT	GAATCAGATG	AAAATCGTTA	1320
5	TGAAAGAGCA	ATGCGTATAC	AGGCTAAAGT	AGCATCATT	GTTACAGCGT	TTGCTCGAGT	1380
	AAGACAAGAT	AAAGAACCAC	TTAAGCCTAA	TCCTGACTTA	AGTTATGCGG	CAAACCTCCT	1440
	ATATATGTTA	CGTGGGGAAT	TACCAACAGA	TATAGAAGTA	GAAGCCTTCA	ATAAAGCACT	1500
10	TATTTTACAC	GCTGATCATG	AGTTGAACGC	ATCTGCATTT	ACGGCACGTT	GTGCGGTATC	1560
	ATCATTGTCA	GATATGTACT	CAGGTATTGT	AGCAGCCGTA	GtTCTCTGAA	AGGGCCATTA	1620
	CATGGTGGTG	CAAACGAACA	AGTTATGACG	ATGTTATCTG	aGATTGGGTC	AaTTGAAAAT	1680
15	GTTGATGCTT	ACTTAGATGA	AAAATTTGCT	AATAAAGrTA	AAGTAATGGG	cTTCGGTCAT	1740
	CGTGTATATA	AAGATGGTGm	tCCTAGaGCG	AAaTATTTaA	GaGAAaTGAG	CCGTCAaATT	1800
20	mCGAAAGACG	CTGGTCGTGA	AGAATTATTT	GAAaTGTCAG	TGAAaTGGA	AAAmCGTATG	1860
	GCAGAAGAAA	AAGGATTAAT	TCCTAATGTT	GATTTTTATA	GTGCGAGTGT	TTATCACTGT	1920
	ATGGAATAC	CTCATGACTT	ATTCACGCCA	ATCTTTGCTG	TAAGTCGTTT	TGCAGGATGG	1980
25	ATTGCTCATA	TTTTAGAACA	ATATAAAGAT	AATAGAATTA	TGCGTCCTAG	AGCGAAATAT	2040
	ATTGGCGAAA	CGAATCGTAA	GTATATCCCG	CTTGraGaaa	GAAAmTAATC	AATACAAATT	2100
	AAAAATGAAG	ATGTAAAATT	TGGAGGTAAA	ATAACTATGA	CTGCAGAAAA	AATTACTCAA	2160
30	GGAAGTGAAG	GATTAAACGT	ACCTAATGAA	CCAATTATCC	CATTTATTAT	CGGTGATGGA	2220
	ATTGGACCGG	ATATTTGGAA	GGCAGCAAGC	CGAGTTATAG	ATGCTGCTGT	TGAGAAAsCC	2280
35	TATAATGGCG	AAAAACGCaT	TGAATGGAAA	GAAGTGCTAG	CTGGCCAAAA	AGCATTTGAT	2340
	ACAACTGGTG	AATGGTTACC	TCAAGAAACA	CTTGATACAA	TTAAGAATA	TTTAATTGCT	2400
	GTTAAAGGAC	CTTTAACAAC	ACCAATTGGT	GGTGGTATTA	GATCATTTAA	TGTGGCTTTA	2460
40	CGCCAAGAAT	TAGATTTATT	TACTTGCTTA	AGACCGGTAC	GTTGGTTTAA	AGGAGTACCA	2520
	TCACCTGTTA	AACGTCCACA	AGATGTTGAT	ATGTTATTTT	TCCGTGAAAA	TACTGAAGAC	2580
	ATTTATGCTG	GTATTGAATT	TAAAGAAGGT	ACAACAGAAG	TTAAAAAGGT	AATTGACTTC	2640
45	TTACAAAACG	AAATGGGTGC	GACAAACATT	CGATTCCCAG	AACTTCAGG	TATTGGTATT	2700
	AAACCAGTTT	CTAAGAAGG	AACTGAGCGA	TTAGTTAGAG	CAGCTATACA	ATATGCTATC	2760
50	GATAATAACC	GTAAATCAGT	TACTTTAGTT	CATAAAGGTA	ATATTATGAA	ATTTACAGAA	2820
	GGCTCATTTA	AGCAGTGGGG	TTACGATTTA	GCATTATCTG	AATTGGTGA	TCAAGTATTC	2880
	ACTTGGCAAC	AATATGACGA	AATTGTTGAA	AATGAAGGCA	GAGATGCTGC	TAATGCTGCT	2940
55							

TTACAACAAA TTTTAACTCG TCCAGCTGAG CATGATGTTG TAGCAACTAT GAACTTGAAT 3060
 GGTGACTATA TTTAGATGTC TTTAGCTGCA CAAGTTGGTG GTATTGGTAT TGCGCCAGGT 3120
 5 GCAAACATTA ATTATGAAAC AGGTCATGCT ATTTTGAAG CAACACATGG TACAGCTCCA 3180
 AAATATGCAG GTTTAAATAA AGTGAATCCA TCTTCAGTAA TTTTAAGTTC TGTATTAATG 3240
 TTAGAACATT TAGGATGGCA AGAAGCGGCA GATAAGATTA CAGATTCAAT TGAAGATACA 3300
 10 ATTGCTTCAA AAGTTGTTAC TTATGACTTT GCCCGTTTAA TGGaTGGtGC TGAAGAAGTT 3360
 TCTACATCAG CATTTCAGA TGAATTGATT GdAAATTTAA AATAAGCAGA ATAGAATTAG 3420
 15 G 3421

(2) INFORMATION FOR SEQ ID NO: 338:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3173 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

CCCTnGATAC CCAAAGTCC CTAATACCAC TGGCTAAACC TAATGGATAG TACCATTTAT 60
 TTTCCAATAA ATAAACCAAC TGCAATnGCT ATaACTCTAA ATATAATAGA GATAATCmCA 120
 30 TTAATTGGAT TAATACGCCA AATACTAGTA ATAATAGGCT AGATAATAAT CCACCTAAAA 180
 AGTACTTTTT AATTCCAAAG AAAGCTAATA TCAATAATGC TGCCGGTGCA GATAATTGAA 240
 35 AATCTAATCC TGGTATAATG GACGGTATTT TCAAAAGTGC CAAAATGGTT AAAATCGCAG 300
 CAATGACACT AATTGAGTA ATATCTTTTG ATGTCATACT AAAACCCCTA TACCGTTTCA 360
 TAAACAACTT GCTTCGGTGT GCTTTCTAAA AATGATATGT AATGATTTAA ATCAATACAA 420
 40 TCGTCCACAA ATATTATTCT GCCTCCATAT CTCGTATTAA CTGGTTTAAT ATCAAATAAT 480
 CGATGGTAAC CAATTTTAGC AGCGGCAAAA TAACCTGTGC TATACGTAA GTCATCGGAC 540
 ACGCAAAGTT CTCCTTTGAC ATACGGATGC GCATTGATAC AACTAGCAAT TGCTAAGGCA 600
 45 TCAGTCACTC TTTCAATTAAG ATCACCTTTA TTATTTATAT CTTCAAACGA AAAATGTGTT 660
 GCCCTAATCC CCCTTTGTCC AAATGAATCT AAACGTTTAC CAGATATAGC AGATAGAATA 720
 ATAGCTCCTG TATAAACCGT TTCATTTTAA ATATATGTCA TCCCTTGATT TAGCGCTTGT 780
 50 TCAGTGACAC CACATTCTTG TGTTAAATGT TGGAGATTG CTTTATCATC CTCAATAATT 840
 TGTAATGCTT TTATTGTTG AATCGGTTCC ATGATTTTTT GTATTTTTAT ATTGAGAAAA 900

	ATTTTGTCAA ACTCACAAAT CGTTTCAGCA CCACTAATAT GAACATCTTG ATTGCTAGAA	1020
	CGCATTTTTTA TACTATACAT GACGATCACC TcAATCTTCT TGaTGCAAAA TTTCAAACAA	1080
5	CCTATCTATA TCTTGTTcAG TATGAAAATA CGACAATGAT ATTCTTAACA TTGGCTTAGT	1140
	CACAGTtGGA TACCTTAAAT AACTTGTAaa CACATGATGC TTTAATAATG TTTGATGAAT	1200
	GTTCTCAGCC GCTTCTATGT CATCAAACtC AATAAACTTA ATCGGCGAGT TTGaACTATT	1260
10	ATAATKAACA TTGAGTGCTT TTAACtTTTTG GTTAAaATAT TTACTCAAAC TATTTAATTT	1320
	AGTGCGTCTA TCATCAGCAT TTATTAACtT TTCAATGTTT CTTTTTATAA AATACAAATT	1380
	ATAAATTGGC AAACtACTTG AGTAGATGAG TGGTCTACCG TGATTAATTA aCATATCCtT	1440
15	CaCATCaTTT GaACTKaAAA TcACACCCCC GTATGCACCA CATGCTTTAG ATAAACTAGA	1500
	AGTGAGTATA TCTACACCTT GATAATTcGA GTAAtTCTCT ATtCCAAAAC TATGTGAAAC	1560
20	ATCGAGTATC AGTGTTGCGT TAnATTTATG CTTTAATGAG ACTAATTGAC CAATATCCAC	1620
	AACGTCGCCA TTCGTTGAAA ATACACTATC AGATATGATT ATTTTTGGTA TATTTTGATT	1680
	AGGGTATTTT TCTAACCTTT TTTCTAAATC AGCAATATCT AAATGCTTAT ATATCACTTT	1740
25	TTCTAAACCA CTTAACtTAA TACCGTCAAT AATACTCGCA TGATTTTCTT GATCTGAAAA	1800
	CACGACACAA TTTGTATTTT TGAAAATATT AAATAACGCC AAATTAGCAT CATAACCACT	1860
	ATTTAAGATA GTACATGcAC TATATCCGAG CCAACCTGCT AACATTGTTT CAATTTCTTC	1920
30	ATAAGCTGTC GAACTTCCAC TAATTAATCT TGAACtTGAT AAGTGATAAC TATACTCCG	1980
	CATAAATCTT TCGAAATCAT CcTTATCAAA CGCTATTTGA CCTAATCCTA AATAATCATT	2040
	AGATGTATAG TTCGTACATC TCTTATTTTc TACTTCAATA TACTGTCTAT CTATATACCC	2100
35	TACCGATTTA AGCGACCGAT ATAACCCTTT CTGTTGTAAT AAATCAATTT GCTCTTGAAA	2160
	CTTCATTCTT GTTTTCCtTA TTTTCACAAG TGTCATAATC AATTTCAAAG CCTAAATCAT	2220
40	TAATCATATC GTAGTCTAAT TGGTTCGGTT GCCCACCAGT AATTAGATAA TnCACCGACA	2280
	AATATTGAAT TCGCCGCTTT TAATGCTAAT GGCTGTAACG AACGTAAGTT GACCTCTCTT	2340
	CCTCCAGCAA TACGAATTTc TTTCGTAGGA TTGATTAAAT GGAATAATGC TACGATTCTT	2400
45	AAACATTTCA TTGGTGTTAA ATCATCCATG CTTCCAAACT TTGTGCCTTT GATTGGATGC	2460
	AAAAAATTAA TCGGAATACT GTCGGCATCC ATTTCTTTTA AAGCAAATGC CATATCAACA	2520
	ATATCTTGAT TAGATTCTCC CATACCACAA ATCAGCCAG AACATGGTGA TATATTATTC	2580
50	GCTTTCATTA GTTCTATCGT ATCTGTTCTA TCTTTATAAC TATGCGTTGT CACGACGTTA	2640
	TCATGGTAAT TTTCACTTGT ATTAATATTG TGGTTATATC TGTCTACACC AGCTGACTTA	2700

55

TGTTGAGATT TAATCGTTCT TACAGTATTA CTAATATGAT CAACTTCTTT ATCGCTCGGT 2820
 CCTCTACCAC TCATAACAAT ACAATATGTT CCAATATGAT TATCATGTGC CACCTTTGCT 2880
 5 CCATCGATAA TTTGTTCCCTC TGGGAATTAAA GCATATCGCT GTTTTTGTTT AATATCTCGT 2940
 GATTGTCCAC AGTACCCACA ATTTTCAGGA CATATACCAC TTTTAGCATT TAAAATCATG 3000
 10 TTTAATTTTA CTTTTTTACC AAAATAATGT TTTCTTAAAA TGTACGCCTC ATTTAATAAA 3060
 TCTAAGGTAT CAATATTAGT ATCCTCATAA ATTTTCAATA CAGTCTCTTT TGTtAATTGT 3120
 tCCCCTTGTA ATATGCGTTT AGCCAAATTC ATATTAAACAC TTCCTATCTA AAA 3173

(2) INFORMATION FOR SEQ ID NO: 339:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1694 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:

25 CGATTATCCA TTAATACAAC CCTAAGTAAA TGTATAAAAA TTATCTTCCA CAAACTTCAA 60
 CAAAAGCCTA AATAAATTAC AGCAATTTAT CAAATATTGC TTACTTTGAT TTTATGAAAT 120
 30 nACTTAATTC TAACACATAC TAAATCATCA TATACTAATT CGAAATCAAA TGCATTTAGA 180
 GATAATCaAA ATGCGGAAAC ATCTCCaATA ATCAATAATC TATTCCCAAT AAATATGAAT 240
 GTTCTCAACA ATACATTATT TATATCTCTT TACACTGTCA TCGACAAAAA CTAAATCTTT 300
 35 CACTTTCAAT TTCGAACGTG GTTCTACGAC ATTTGCTGCT ATATCATTTA ATGGGATTAA 360
 AACAAATGCA CGTTCATTCA TTCTCGGATG TGGCACCGAC AGTTTTGGTA AATCTATCAT 420
 TTCTTCTCCA TACAACAAA TATCCACATC TAAAGTTCTA GGACCCCATC GTTCCTTTCT 480
 40 AATACGGTGT AAACATTCTT CTGTCKTCAA ACAACATTCC AACAGTTGTA ATACTGTGAG 540
 TGTTGTTTsA ATTTCAACAC ACAAATTTAA AAAGTTAGGT TGCTCAGTAT ACCCAACTGG 600
 TGCTGTTTCA TAAATCGGAG AAATAtTAGA TACGTTAATA CCATCATATT CATTCaAAAT 660
 45 CTTyATAGCA TCGTTTTAACT GGCTTTCTCT ATCACCAaTa TTACTACCTA AyCCTAAGTA 720
 TGCTTGAATC ATyTATTCTC CCTCACTATT TCGATACCTA CTCCATCATA ATGACCCGGA 780
 ATCGGTGGGT TTTCTTTAGT GATTCTCACT TTCGTTTCCA TTACACGATT ATATTGTGAA 840
 50 TTTATACGAT TTGCAATACG TTCAGCTAGA TGCTCAAGTA AATTAAACGGC CTTACCTTCC 900
 ATAATTGATT TAACCTCTTc GAACACTTCA CCATAATGAA CTGTATCAAT AACATTATCA 960

ATTTTCATTTT CAGCTGATAA AGCACCATGA TATCCATAAA AGCGCATACC TTTAAGAAAG 1080
 ATTGTGTCTT GCATTTTTCAT TCTCCTTTAA AAAATCTATA CCTTTAGCTA ATTTAGCATT 1140
 5 CAACTCGACA TTATGAACGC GTACTGCTCT AACGCCTTTC ATAATACCAT ATGCAGTCGT 1200
 AGCTGCAGTT ACTTCATCTC TTTCAACCGG TGTGTATCA TAACCCATCA TCTCTTTAGT 1260
 GAAACGTTTC CGGCTTGTCG CTAATAAAAC TGGATATTCT GTTGCAACAA GTTCATCCAG 1320
 10 TCTTGCCATA ACTTCGGCTT CTTCATTTCT AGTTTTAGCG AAACCTATAC CTGGATCTAG 1380
 CCAAATTTTA TTTGAAGGTA TACCAGCTAT TTTAGCTTGA TGTGCTTGTG CTAACAAAGA 1440
 TGTTAACATT TCTTCGACAA CCGGTTTCATC ACGATTACCA TTTCCATTAT GCATTAAAAT 1500
 15 AATTTCCGCG TCATATTTAG CTACAATTTG GAACATACGA TGATCATACA GACCGCCcAT 1560
 tGATCATTAA TCATATCAAC GGCTAATTTT AAACATGCTT CAGCAACCTC ACTTCGAAAT 1620
 20 GATCGACTG AAATTTTTTA CATCAaAACC GACAATAGCT TCaCAACAG TAATACTCTG 1680
 TTCCATCTCT TCTG 1694

(2) INFORMATION FOR SEQ ID NO: 340:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1358 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

AGCATTTCTT TTCTATAAAC ATTTAATTGA ACATTATTAA GTACACTATT ACTATAGTCA 60
 35 CTATATTGAA CaCATACCTC ATTTAATTCT AATAGCGGTT CAGATTTGTA CTTATTATCA 120
 TTATTTGCAG ATGTTTCATC TATCCATTTT TTCACTTTAA ATTTAACATG TTCACTCATA 180
 40 CAAACGTCAC GTAAATTTCG TAAGTTATCA ATGGATTGCA CATCTACTTC TGCATATTTA 240
 AGCGCTGTAC AGTATAATGG TTCACGTATG CCTGCTTCTT TAAGCTTAGA TGATTTTAGC 300
 AAATCACTAG GCGTTGTATT AGCGATGATT TTTCCATCTT TAAAAAGAAG AACTCTATCA 360
 45 AACGTATCAT CTAATGATTC TTCTAATCGA TGTCGACAA TAATCATCGT TGACTTTGTT 420
 TCTTCATGAA TATTGTTTAA CAATCTCAGC GTTTCATGTC CTGTCGCAGG ATCTAAATTG 480
 GCCAGTGGCT CATCCAATAT TAAAATAGGC GTACGATGGA TTAATATACC ACCTAATGAA 540
 50 ACGCGTTGTT TTTGACCTCC AGATAAATCT TCGGGTCGGT GATTTAAATG TTCTATCATG 600
 CCAACTTTTT CAGCCCAATA ACTTACATTT TTCTTCATAT CATCTTGTTT AACACAATTA 660

55

TCTTGTA AAA CTGTACCAAC AACATTAGAT CTATCATGTA AACAACTAAC GGTTGCATCT 780
 TGATTATTTA TATATAGTTC CCCAGTTATG TTACCTTTAG TTTTAAATGG AATTAAATCCG 840
 5 TTTATGCAAT TTGCAAAAGT CGATTTACCA CTACCCGAAG CACCAACTAC TAATACTTTT 900
 TCTCCTGGAT AAATATCAAC ATTTATATTC TGTAATGTAG GTGTTGCTTG ACTATGATAT 960
 TGAAAACTAA AGTCTTTGAA CGAGATAATT GGTTCACTCA TGATATATCA TTACCTTTCT 1020
 10 ATATTCATTT ACATATCTGA TTCAACAAA TAACTATTCC TTACGTAAAC TACCTTTTTT 1080
 AATTGAGAT GAAGCATATG CTTTAAATA TATTGTCCCA ATAATGCCAA CTGAAATAAT 1140
 ATTTAATACT GCAGAGATAA CACCTTGTGT ATAAACCTTG TTAGCTGGTT CGTTATAAAT 1200
 15 CAAAATATCT AATGTTGGTG CAATAAGTGC CCAGCAAATA ATATTGCGAA TAATTTGACC 1260
 GATATTAAAA TAAACCATCG ATTTCCTAGA AAATCGGCCT GAAGAAAGAT TTAATTTTAG 1320
 20 TCCAATCCAG CCATATAAAC AGCCTATAAT TCCCAGAG 1358

(2) INFORMATION FOR SEQ ID NO: 341:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4557 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

TAGAAGAATT GGAGAAAATG CTAATTCAAT TGTCAACATT CCATAGTTAT CATGATTTAG 60
 AGTTTCTATT TGTGACACGT GAAGATGAAG TTGAAACATT GAAATGGGCA CGTTGGTTGC 120
 35 CACATATGAC ATTGAGAGGG CAAAACATTA GAGGATTTGT TTACAATCAA CGAACGCGTG 180
 ACCAAATTTT AACGTCAATT TATAGCATGA TTAAAGAACG TATCCAAGCT GTGCTGAACG 240
 40 CAGCAGAAGT AATGAGCAAA TTATTTTCAC ACCGCAATTA GTGTTTGTCA TTACAGATAT 300
 GTCATTAATT ATTGATCATG TCATTTTAGA ATATGTAAAC CAAGATTTAT CAGAATATGG 360
 TATTTCAATTA ATCTTTGTTG AAGATGtGAT TGaAAGTTTG CCAGAGCATG TAGATACCAT 420
 45 TATTGATATC AAGTCTCGTA CTGAAGGCGA ACTGATTACG AAAGAAAAAG AATTAGTTCA 480
 ATTGAAATTT ACACCTGAAA ATATTGtTAA CGTCGATAAA GAATATATCG CGCGACGTTT 540
 GGCGAATTTG ATACACGTCG AACATTTGAA AAATGCAATT CCTGATAGTA TTACATTTT 600
 50 AGAGATGTAT AACGTGAAAG AAGTAGATCA GCTTGATGTG GTTAATCGAT GGAGACAAAA 660
 CGAAACATAC AAAACGATGG CAGTACCTTT AGGTGTAAGA GGTAAAGATG ATATTTTATC 720

EP 0 786 519 A2

	AGGGAAATCT GAGATTATCC AATCATACAT TTTATCTTTA GCTATTAATT TTCACCCCTCA	840
	TGAAGTTGCA TTCCTATTGA TTGACTATAA AGGTGGGGGT ATGGCGAACT TATTTAAAGA	900
5	TTTAGTCCAT TTAGTTGGTA CGATTACAAA CTTAGATGGC GATGAAGCGA TGCGTGCCTT	960
	AACATCAATC AAAGCCGAAT TGAGAAAACG TCAACGTTTA TTCGGAGAGC ATGATGTTAA	1020
	CCATATTAAT CAATACCATA AGTTATTTAA AGAAGGTATT GCGACAGAAC CAATGCCACA	1080
10	TTTATTCATT ATTTCCGATG AGTTTGCCGA ATTAAATCA GAACAACCTG ATTTTATGAA	1140
	AGAACTTGTA TCAACGGCAC GTATTGGACG TTCGTTAGGT ATTCATTTAA TACTTGCGAC	1200
	ACAAAAACCA TCGGGTGTG TTGaTGACCA AATTMGTTCT AACTCTAAAT TTAAGTTGGC	1260
15	ATTAAAGTA CAAGATAGAC AAGACAGTAA TGAAATTTTA AAAACACCAG ATGCAGCAGA	1320
	CATTACmTTA CcAggTCGTG CGTATTTACA AGTTGGTAaT AATGAmATTT ATGAATTATt	1380
20	CCAATCTGCA TGGAGTGGTG CAACATATGA CATCGAAGGC GATAAATTAG AAGTTGAAGA	1440
	TAAGACGATT TACATGATTA ATGACTATGG TCAACTTCAA GCAATCAACA AAGACTTGAG	1500
	TGGACTTGAA GATGAAGAAA CGAAAGAAAA TCAAAGTGGG TTAGAAGCGG TCATAGATCA	1560
25	TATCGAATCT ATTACAACAC GATTAGAAAT CGAAGAAGTT AAGCGTCCAT GGCTACCACC	1620
	ATTGCCAGAA AATGTATATC AaGAAGATTT AGTAGAAACa GATTTTCAGAA AATTATGGTC	1680
	AGATGATGCA AAAGAAGTGG AATTAACATT AGGACTTAAa GACGTACCAG AAGAACAATA	1740
30	TCAAGGACCG ATGGTATTGC AATTGAAAAA AGCTGGGCAC ATCGCGTTAA TCGGAAGTCC	1800
	AGGATATGGT AGAACAACGT TCTTACACAA CATTATTTTC GATGTTGCAA GACACCATCG	1860
	TCCTGATCAA GCACACATGT ACTTGTTCTGA TTTCCGGTACC AATGGTTTGA TGCCAGTTAC	1920
35	AGACATACCA CATGTCGCTG ATTACTTTAC AGTAGATCAA GAAGACAAGA TTGCTAAGGC	1980
	GATACGTATA TTTAATGATG AAATTGATCG TCGTAAGAAG ATTTTAAGTC AGTATCGTGT	2040
40	CACTAGTATT TCTGAATATC GAAAATTAAC TGGTGAAACA ATTCCGCATG TCTTTATTCT	2100
	TATTGATAAC TTTGACGCAG TAAAAGATTC ACCTTTCCAA GAAGTTTTTG AAAATATGAT	2160
	GATTAAAATG ACGCGTGAAG GGCTAGCATT AGACATGCAA GTAACCTTAA CTGCTTCAAG	2220
45	AGCTAACGCT ATGAAAACAC CAATGTACAT TAATATGAAA ACGCGTATCG CAATGTTTTT	2280
	ATATGATAAA TCAGAGGTGT CGAACGTAGT AGGACAGCAA AAATTTGCGG TTAAAGATGT	2340
	TGTGGGTCTGA GCATTGTTAA GTAGTGATGA CAACGTATCA TTCCATATTG GCCAACCAT	2400
50	TAAACATGAT GAGACCAAAT CATATAATGA TCAAATTAAT GATGAAGTAT CGGCGATGAC	2460
	AGAATTTTAT AAAGGTGAAC ACCAAATGAT ATtCCTATGA TGCCAGATGA AATTAAATAT	2520

55

	GGATTAGATT ATGAAGGTGT TACACTACAA AAAATTAAAT TAACTGAACC AGCAATGATT	2640
	TCATCAGAAA ATCCGAGAGA AATTGCGCAT ATTGCTGAAA TTATGATGAA AGAAATTGAC	2700
5	ATATTAAATG AAAAATATGC GATTTGTATC GCAGACTCAA GTGGAGAGTT TAAAGCTTAT	2760
	AGGCATCAAG TGGCTAACTT TGCCGAAGAA AGAGAAGACA TTAAAGCGAT TCATCAACTA	2820
	ATGATTGAAG ACTTAAAGCA AAGAGAAATG GACGGCCCAT TTGAAAAGA TTCACTTTAT	2880
10	ATTATCAATG ATTTTAAAC ATTTATTGAT TGCACGTATA TTCCGGAAGA TGATGTTAAA	2940
	AAGCTTATTA CAAAAGGACC AGAACTTGGC TTGAACATTT TATTTGTCGG CATTCATAAA	3000
	GAATTAALAG ATGCTTATGA TAAACAGATT GATGTTGCAC GTAAAATGAT TAACCAATTT	3060
15	AGTATAGGTA TTCGTATTTC AGACCAACAA TTCTTTAAAT TTAGATTTAT TCAACGAGAA	3120
	CCTGTTATTA AAGAAAATGA AGCATATATG GTCGCAAACC AAGCTTATCA AAAGATTAGA	3180
20	TGGTTTAAAT AGCAATGAAT TAAATAGGAG GGAGGTATGT TATGAATTTT AATGATATTG	3240
	AAACAATGGT TAAGTCGAAA TTTAAAGATA TTAAAAAGCA TGCTGAAGAG ATTGCGCATG	3300
	AAATTGAAGT TCGTTCTGGA TATTTAAGAA AAGCTGAACA ATATAAGCGA TTAGAATTTA	3360
25	ATTTGAGTTT TGCCTAGAT GATATTGAAA GCACAGCAAA GGACGTACAA ACTGCAAAAT	3420
	CTAGTGCTAA TAAGGACAGT GTAACGTGTA AGGGAAAGGC GCCCAATACG TTATATATTG	3480
	AAAAAGAAA TTTGATGAAA CAAAAGCTTG AAATGTTGGG TGAAGATATC GATAAAAATA	3540
30	AAGAATCCCT CCAAAAAGCT AAGGAAATTG CTGGCGAAAA GGCAAGTGAA TATTTTAATA	3600
	AAGCAATGAA TTAATATTGA GGTGAAGATA TGGGTGGATA TAAAGGTATT AAAGCAGATG	3660
	GTGGCAAGGT TGATCAAGCG AAACAATTAG CGGCAAAAAC AGCTAAAGAT ATTGAAGCAT	3720
35	GTCAAAAGCA AACGCAACAG CTCGCTGAGT ATATCGAAGG TAGTGATTGG GAAGGACAGT	3780
	TCGCCAATAA GGTGAAAGAT GTGTTACTCA TTATGGCAAA GTTTCAAGAA GAATTAGTAC	3840
40	AACCGATGGC TGACCATCAA AAAGCAATTG ATAACCTAAG TCAAAATCTA GCGAAATACG	3900
	ATACATTATC AATTAAGCAA GGGCTTGATA GGGTGAACCC ATGATGAAAG ATGTTAAGCG	3960
	AATAGATTAT TTTTCTTACG AAGAATTAAC AATTTTAGGT GGTAGTAAAT TGCCTCTCGT	4020
45	AAATTTTGAA TTGTTTGATC CATCAAATTT TGAAGAAGCT AAAGCTGCTT TAATTGAAAA	4080
	GGAATTAGTA ACAGAGAATG ACAAGTTAAC TGATGCAGGT TTTAAAGTGG CTACATTAGT	4140
	CAGAGAGTAT ATTAGCGCCA TTGTAAATAT TCGAATTAAT GATATGTATT TTGCACCATT	4200
50	TAGCTATGAA AAAGATGAAT ATATTTTGTT AAGCCGGTTT AAAAATAATG GGTTCAAAT	4260
	ACGAATTATC AATAAAGACA TTGCATGGTG GTCGATTGTA CAATCATATC CTTTATTGAT	4320

55

CTTAAATAAT GAAAGTATCG ATACGATTGG GCGTGTTTTA GAAATTGAAA TATACAATCA 4440
TCAAGGTGAC CCTCAACAAA GTTTATATAA CATTATGAA CAAAATGATT TGTTATTCAT 4500
5 TCGATACCCA TTAAGAGATA AAGTGCTGAA TGTTATATT GGTGTCATTA ATACATT 4557

(2) INFORMATION FOR SEQ ID NO: 342:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3931 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

TTGAGTGA CT TTATTGAAGC GCGTGTAGAA GAAATATTCT TCGAAGTATT TGATGTTTTA 60
20 CAAGATTTAG GATTAACAAA AGTAAATGGT GGGTTTATTG TAACTGGTGG ATCTGCAAAC 120
TTACTTGGCG TAAAAGAATT ATTATCAGAT ATGGTAAGTG AAAAAGTTAG AATTCACACG 180
CCATCACAAA TGGGAATTAG AAAACCTGAA TTTTCTTCAG CAATTTCTAC AATTTCTAGT 240
25 AGTATCGCTT TTGATGAGTT ATTAGATTAT GTTACAATTA ATTATCATGa TAATGAAGAA 300
ACTGAAGAAG ATGTTATTGA TGTGAAAGAC AAAGATAACG AATCTAAATT AGGCGGaTTT 360
GaTTGGTTTA AACGTAAAAC AAACAAAAAA GATACTCATG aAAATGAAGT AGAGTCAACA 420
30 GATGAAGAAA TTTATCAATC AGAAGATAAT CATCAGGAAC ATAAACAGAA TCATGaACAT 480
GTTCAAGACA AAGATAAAGA TAAAGAAGAA AGTAAATTCA AAAAATAAT GAAATCTCTA 540
TTTGAATGAT TATTGGCCAA TAAACTAGG AGGAAATTTA AATGTTAGAA TTTGAACAAG 600
35 GATTTAATCA TTTAGCGACT TTAAAGGTCA TTGGTGTAGG TGGTGGCGGT AACACGCCG 660
TAAACCGAAT GATTGACCAC GGAATGAATA ATGTTGAATT TATCGCTATC AACACAGACG 720
40 GTCAAGCTTT AAACCTATCT AAAGCTGAAT CTAAATCCA AATCGGTGAA AAATTAACAC 780
GTGTTTtagg AGCAGGAGCT AATCCTGAAA TCGGTAAAAA AGCTGCAGAG GAATCTCGTG 840
AACAAATTGA AGATGCAATC CAAGGTGCAG ACATGGTATT TGTTACTTCT GGTATGGGTG 900
45 GCGGAAGTGG TACTGGTGCA GCACCAGTCG TTGCTAAAAT TGCAAAAGAA ATGGGCGCAT 960
TAACTGTTGG TGTGTAACT CGTCCATTTA GTTTTGAAGG ACGTAAACGT CAAACTCAAG 1020
CTGCTGCTGG AGTAGAAGCT ATGAAAGCTG CAGTAGATAC ATTAATCGTT ATACCAAATG 1080
50 ACCGTTTATT AGATATCGTT GACAAATCTA CGCCAATGAT GGAAGCATTT AAAGAAGCTG 1140
ACAACGTGTT ACGCCAAGGT GTACAAGGTA TCTCAGACTT AATCGCTGTT TCTGGTGAAG 1200

	GTATTGGTGT TTCTTCTGGT GAAAATAGAG CGGTAGAAGC TGCTAAAAAA GCAATCTCTT	1320
	CTCCATTACT TGAAACATCT ATCGTTGGTG CACAAGGTGT GCTTATGAAT ATTACTGGTG	1380
5	GCGAGTCATT GTCATTATTT GAAGCACAAG AGGCTGCTGA TATTGTCCAA GATGCTGCAG	1440
	ATGAAGACGT TAATATGATT TTCGGTACAG TTATTAATCC TGAATTACAA GATGAGATTG	1500
	TTGTAACAGT TATTGCAACT GGTTTTGATG ACAAACCAAC ATCACATGGT CGTAAATCTG	1560
10	GTAGCACTGG ATTCGGAACA AGCGTAAATA CTTCTAGCAA TGCAACTTCT AAAGATGAAT	1620
	CATTCAC TTC AAATTCATCA AATGCACAAG CAACTGATAG TGTAAGTGAA AGAACACATA	1680
	CAACTAAAGA AGATGATATT CCTAGCTTCA TTAGAAATAG AGAAGAAAGA CGTTCAAGAA	1740
15	GAACAAGACG TTAATCGGTT AATATATATA CACAAATAAT TCAACACAAA TCATCAGATA	1800
	ACATATCTGA TGATTTTTTT ACTAATTTTT AGaACATGTA GAAGGACATT TAAGTTTTTC	1860
20	aAAGTTATTA AAAGTGT TTA AGTATCGTGT GAAAATTAAG TCaAAAATTA TTTGCGCAAC	1920
	ATTTTAACTT TAAACATAAA TGTTATATTA TATAATTATT AACTTTGTAC AGTTAGACGA	1980
	AGATAATTTA AATGAAATGA TGGTGACGAT CGAGTGAATG ATAATTTTAA AAAGCAACCG	2040
25	CATCATTTAA TATATGAAGA GTTATTACAA CAAGGTATTA CTCTAGGTAT TACAACTAGA	2100
	GGAGATGGTT TAAGTGACTA TCCTAAAAAT GCTTTTAATA TGGCGAGATA TATTGATGAT	2160
	CGCCCATATA ATATTACTCA ACATCAATTG CAATTAGCTG AAGAAATTGC GTTTGATAGA	2220
30	AAAAATTGGG TGTTTCCCAT TCAAACACAT GAAAATAAAG TCGCTTGAT TACAAAGGAT	2280
	GATATAGGCA CAAATATAGA CACTTTAACT GATGCGCTTC ATGGTATTGA TGCGATGTAC	2340
35	ACATATGATA GTAATGTCTT ATTAACGATG TGTTATGCAG ACTGTGTACC AGTATATTTT	2400
	TATAGTACAA AACATCATTT TATTGCATTG GCGCATGCAG GTTGGCGTGG TACCTATACT	2460
	GAAATGTAA AAGAAGTGCT AAAACATGTG AACTTTGATT TGAAAGACTT ACATGTCGTT	2520
40	ATTGGACCAT CTACATCATC AAGTTATGAA ATTAATGATG ATATTAAAAA TAAATTTGAA	2580
	ACATTGCCAA TTGATAGTGC CAACTATATT GAACTAGAG GACGAGATCG TCATGGTATT	2640
	GATTTGAAAA AAGCCAATGC TGCATTATTA ATTTATTATG GTGTTCTTAA AGAAAATATT	2700
45	TATACGACAG CGTATGCTAC ATCTGAACAT TTAGAATTAT TTTTCTCTTA TCGATTAGAA	2760
	AAAGGTCAAA CAGGACGCAT GTTAGCATTG ATTGGTCAAC AGTAAACAAG GAGGAGATAT	2820
	GTTTGCGTGT GAAAGATAAT TTACAACAAA TCTCAACACA AATTAATGAC AAAAGTGAAA	2880
50	AAAATAATTT TTCAACAAAA CCAAACGTGA TTGCAGTTAC AAAATATGTT ACAATAGAGC	2940
	GAGCTAAAGA AGCGTATGAG GCTGGAATAA GACATTTTGG TGAGAATAGA TTGGAAGGCT	3000

55

AATCTCGAAA AGTTAAGGAC GTTATAAACG ACGTAGATTA TTTCCATGCT TTAGATCGAT 3120
 TGAGCTTAGC CAAAGAAATT AACAAACGTG CAGAACATAA AATTAAATGT TTCTTGCAAG 3180
 5 TGAACGTTTC GGGAGAAGCT TCTAAACATG GTATTGCTTT AGAAGATGTT GATCAGTTTA 3240
 TAGATGATCT TAAAAAATAT GACAAAATCG AAATTGTAGG TTTAATGACG ATGGCACCAT 3300
 TGACAGATGA TGAAGCATAT ATTAGATCGT TATTTAAACA GTTACGTTTG AAAAAAGAAG 3360
 10 AAATACAACG ACTCAATTTA GAATATGCGC CTTGTGATGA ATTATCAATG GGAATGAGTA 3420
 ATGACTATCT TATTGCAGTT GAAGAAGGTG CGACGTTTGT TAGAATTGGG ACTAAACTTG 3480
 TAGGAGAAGA GGAGTGAGCC ACTTGGCTTT AAAAGATTTA TTTAGTGGAT TTTTGTAAAT 3540
 15 AGATGATGAA GAGGAAGTAG AAGTACCTGA CAAACAACAA CAGGTAAATG AAGCGCCAGC 3600
 AAAAGAGCAG TCACAACAAA CAACAAAACA AAACGCAATC AAATCAGTCC CTCAAAAATC 3660
 20 TGCATCAAGA TATACAACAA CGTCAGAAGA AAGGAATAAC CGTATGTCTA ATTATTCAAA 3720
 AAATAATTCA CGTAATGTTG TAACTATGAA CAATGCTACA CCAAACAATG CATCACAAGA 3780
 AAGTTCAAAA ATGTGTTTAT TCGAACCACG TGTTTTTTCA GATACACAAG ATATTGCTGA 3840
 25 TGAGCTTAAA AACCGCCGTG CGACACTTGT CAATTTACAA CGTATTGATA AAGTATCAGC 3900
 GAAAAGAATT ATTGATTTTT TAAGCGGTAC T 3931

(2) INFORMATION FOR SEQ ID NO: 343:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3150 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

40 AATTGTCGGG GGA CTCTTAG GTTTGTGCAT GCAAAGAACA AGATTTTGTT TAACAGGTGG 60
 CTTTCGAGAT ATGTATGTGC AAAAGAATAA TAAGATGTTT TATGCATTAT TAATCGCTAT 120
 TACTATTCAA AGTATAGGAT TATTGATTTT GACGGCAACA GATATTTTAC AAATTCCTGC 180
 45 ACATAGTTTT CCAATATTGG GAACAATTAT AGGTTCTTTT ATTTTGGAA TTGGAATAGT 240
 ATTGGCTGGA GGATGTGCAA CAGGtACTTG GTATCGCGCT GGTGAAGGGC TAATTGGTAG 300
 TTGGATTGCA TTAGTATTAT ATGCTGTTAC TGCAGCAATC ACTAAAACAG GGATTTTAAA 360
 50 GCCAGTAATG GATAAAATTA ATCAACCAAC GAATGTAAAT AGTGATATGT CTCAACAAC 420
 TGGCATTCCG TTTTGGGGAT TAGTCGTTAT ATTAACATA ATCACCATT TTTAGTTGT 480

55

	AGGTATTAGA TATTACCTTT TCGAAAAACG ATACCATCCA TTTATTGCAG CAATTGTAAT	600
	TGGACTTATC GCACTCTTAG CTTGGCCAAT GAGTGCATCA ACTGGAAGAA ATGACGGTTT	660
5	AGGTATAACA ACGCCTTCAG CAAATTTAGT ACACTTTTTG ATTACAGGTG AAACTAAAT	720
	TATTGATTGG GGTGCTTTTT TAGTTCTAGG AATTTTCATT GGTTCATATA TTGCAGCTAG	780
	AGGATCAAGA GAATTTAAAT GGCGATTGCC AGACAAGATT ACAATACGAA ACAGTGCCAT	840
10	TGGTGGCATA TGTATGGGAT TTGGTGCGTC AGTTGCTGGT GGTGTGTTCTA TCGGTAACGG	900
	TTTGGTTGAA ACGGCAACGA TGA CTGGCA AGGATGGATT GCGCTAGCAT gCGATGATAG	960
	TTGGTGTATG GACAATGAGT CATTTTATCT TTGTTGCTCC AATGAAAAAA GTACACCAAC	1020
15	AATCTGCAAA GGTTAAACAG CAAACGCAAA TAGTATAGAA GATTATTATG CAAATGATGT	1080
	TGATCAAATA AAAGTGATTG GAAAAGGAGA AATAATTATG ATACACGAAT TAGGTACAGT	1140
20	AGGAATGGTA TGTCCATTTT CGTTAATTGA AGCGCAAAAG AAAATGGCAA CATTGCAATC	1200
	TGGAGATGAA TTAAAAATTG ATTTTGATTG CACGCAAGCG ACGGAAGCCA TTCCAAATTG	1260
	GGCTGCAGAA AATGGTTATC CTGTAACAAA CTATGAACAA ATTGATAATG CTTTCATGGAC	1320
25	AATTACAATT CAAAAAGTTT AACGTTATCA TTTTAACAAT AAAATAGATA TTAGATTCTA	1380
	TGGCTACTTC CGCTAATTTA AAAGTGAGTA AGTAGTCTTT TTTTTTTTAG TTCATGAAAT	1440
	CATTTTTATA TAGTGTGGCA CATTTTATTC CAAAAGATGT AATAAACTT AACGCATTTT	1500
30	TGCTTTTTAT AAATTGTCAG ATTATTATGA AAAAAAGGGA GTGGTAAGTA TGAATCTTAA	1560
	CGATACGATA TTTATGTTTT TGTGTACATT ATTAGTTTGG TTAATGACAC CAGGATTAAG	1620
	TTTATTTTAT GGTGGGTTAG TTCAATCTAA AAATGCGCTT AATACTGTCA TGCAAAGTAT	1680
35	GGCAGCAATT GTGCTTGTTA CATTTGTATG GATAACAGTT GGTTTTACAA TTAGTTTTGG	1740
	GAATGGGAAT TTATGGTTTC GAAATTGGGA ATATACTTTT CTTAATCATG TAGGTTTTGC	1800
40	GACTCAAGAA GATATTAGCC CACATATTCC TTTCGCTTTG TTTATGTTAT TTCAAATGAT	1860
	GTTTTGTACG ATTGCAATTT CTATTTTATC TGGTTCAATC GCTGAGAAAA TGAAGTTTAT	1920
	TCCTTATTTA TTATTCGTAG TAATATGGAC TGCTCTTGTA TACAGTCCAG TAGCACATTG	1980
45	GGTTTGGGGC GGCGGTTGGA TTAACAAACT CGGTGTATTA GATTTGCTG GAGGTACGGT	2040
	TGTTTCATATT ACATCAGGTG TTTCTGGTTT AGTATTAGCT ATTATGATTG GAAAAGGAAA	2100
	CAAACATTCT GAATCAACAC CACATAATCT TATCATACG TTGAATGGCG GTATATTCGT	2160
50	GTGGATTGGT TGGTATGGAT TTAATGTAGG TAGTGCTTTT ACATTTGATA ATATTGCGAT	2220
	GCTTGCAATT ACAAACTG TCATTTGAGC CAGTGCAGGT GCTATAGGTT GGTTAATTTT	2280

55

	ATTAGTTGTC ATTACTCCTG CAGCAGGATA TGTAACATAT CTTAGTGCAA CAATAATGGC	2400
	TTTAATAGGA GGTATCTGTT GTTATATTGT CATTAAATTAC ATCAAGGTAA AACTAAAATA	2460
5	TCATGATGCA TTAGATGCAT TTGGTATTCA TGGTGTTGGT GGTATTATTG GTGCTGTTTT	2520
	AACAGCAGTT TTCCAAAGTA AAAAAGCCAA TCCTGACATT GAGAATGGCT TTATTTATAC	2580
	TGGTGACATA CATATTATAC TTGTACAAAT ATTATGTGTA ACAGCAGTTG TAATTTTTAG	2640
10	TATAGTCATG ACGTTTATTA TTGCGAAAGT AATTAAATTa ATTACACCAT TATCTGTTAC	2700
	GGAACAAGAA ACGAATATAG GATTAGACAA GATTGTTCaC GGTGAACATG CTTACTTTGA	2760
15	AGGTGAGCTA AATAGATTCA ATAAACATAT TCGATATTAG AATATATTTA CATAGAATAT	2820
	TCATTGTCCT GACATTTAAC TAAAGGTTGA TGTGTTGGACA TTTTGTTATA CAAAAGTTTT	2880
	ATTTTGAAAT CTTTTTATGA AAGAAGCAGA AATATTATTT AAAGCGGTTA CACATATGCT	2940
20	AAAATAAGGC TAAGTGTCAC AAATAATGAT AGGTGAATAA GTATGAAAAA TATATCTGAT	3000
	ATTGCCAAAT TGGCAGGCGT TTCAAAAAGT ACAGTATCTA GATTTTTTAAA TAATGGATCT	3060
	GTCAGTAAAA AAACAAGTGA AAAATTAACA AGAATTATAG CAGAACATGA CTATCAACCG	3120
25	AATCAATTTG CTCAAAGTTT AAGAGCGAGA	3150

(2) INFORMATION FOR SEQ ID NO: 344:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3719 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

	40	50	60
	120	180	240
	300	360	420
	480	540	

ATCGTTTCAA TATTACTTAT AGGGATGGCT ATCAGTAATG TTTCGAAAGG GCAATACGCA 660
 AAGAGGTTTT TCTATTTTCGC TACTAGTTGT TTAGTGTTAa CTTTAGTTGT AGTTTCAAGT 720
 5 CTAAGTAGCT CAGCAAATGC ATCACAAACA GATAATGGCG TAAATAGAAG TGGTTCTGAA 780
 GATCCAACAG TATATAGTGC AcTTnCAACT AAAAAATTAC ATAAAGAACC TCGGACATTA 840
 10 ATTAAAGCGA TTGATGGTGA TACGGTTAAA TTAATGTACA AAGGTCAACC AATGACATTTC 900
 AGACTATTAT TGGTTGATAC ACCTGAAACA AAGCATCCTA AAAAAAGTGT AGAGAAATAT 960
 GGTCTGAAG CAAGTGCATT TACGAAAAAA ATGGTAGAAA ATGCAAAGAA AATTGAAGTC 1020
 15 GAGTTTGACA AAGGTCAAAG AACTGATAAA TATGGACGTG GCTTAGCGTA TATTTATGCT 1080
 GATGGAAAAA TGGTAAACGA AGCTTTAGTT CGTCAAGGCT TGGCTAAAGT TGCTTATGTT 1140
 TATAAACCTA ACAATACACA TGAACAACTT TTAAGAAAAA GTGAAGCACA AGCAAAAAAA 1200
 20 GAGAAATTAA ATATTTGGAG CGAAGACAAC GCTGATTGAG GTCAATAATG CTCATTGTAA 1260
 AAGTGTCCT GCTGCTAGTG GCACTTTTAT AATTTTTAGA TCACGATATG ATTTATTATC 1320
 AATTCAGAAT TAAAAAGTA AATAGTATCA AAAGTAAGTG TATTTAATAT TAGAAAATAA 1380
 25 AAATTTTAAA TTTAGTATTA AAATGGAATG TTACTATATA GTTCAATGTG TATTATCACA 1440
 GAAAAATAAA TAATGCTTTA CTTCTATATT TAAAAGTGTA TAATGAAAGT TAAGTAATAA 1500
 AGAGCGTGAA GAAAAATGTG AGTTATTTAT ATAGAATATT CTCCTTTTCA TTTATGAATT 1560
 30 TGTACAAAAA TATTTAGTGC AAAAGCACGA CGGAGGTATT CAATATGaaT AACGGTACAG 1620
 TTAAATGGTT TAATGCAGAA AAAGGTTTTG GTTTCATCGa AAGAGAAGAT GGTAGCGACG 1680
 35 TATTCgTACA CTTCTCAGCA ATCGCTGAAG ATGGATACAA ATCATTAGAA GAAGGCCAAA 1740
 AAGTTGAATT CGACATCGTT GAAGGCGACC GTGGCGAGCA AGCTGCAAAC GTAGTTAAAA 1800
 TGTAATTTTA ACTTATTCAA ACAGTCCTTA CTATAGGGCT GTTTTTTTAT GCTTTAAATC 1860
 40 GATAACAGTT GGTGTGGTAA AAGCACTAGC CGTTATTTTT TTGTCCAATA AATTTAGTTG 1920
 GAGATTTAAC AATATATAAT GGTTCTAAAA TAAATCGAAC TGATGGAAAA GTTTTTTACT 1980
 TTTCATCTGT CCGACTTTTG ATTTTGAATA TAAAAAGCG CCAATACAGA ACTTTAATAA 2040
 45 TGACGAGAAT TAAAGTCTGT ATATGGCGAT AACAAGAAGT AATGTTAAAC ACTCAAATG 2100
 TTTAACAATA ATAGGATACC ACATCGCATA ATATCTTACT ACTTAATTAA TAATTTAACT 2160
 AATCAACTTT TTGTTAATTT TTTATTAAGA CTGATTAATT ATTGAGAATA TTTATTGTTT 2220
 50 TTAAATCTC ATAATAATTC AGTAATCTTG TTTTCATTTA AAAGGCGAAA CATTAAATA 2280
 ATTAAATAAA AATATTGCGT TTAATTTACA GCGTCAAATA TACTTATTC TAATGCTTTG 2340

55

ATTTCTGTGA GTATTTGGAA GCTACCATTA GGCAACGGTT TAACAATAGA CAATTGCTTT 2460
 TCCGCTTGTT GTATTAAAAA AGGTTTTGTA GATTGATTAT TAATATGCCA TTCACTCATG 2520
 5 TATGTTTTTC ACTCCTGCTT TAAAATAGGG TTAGAAAGTT TATAGTTGAG ACATTCAATG 2580
 TCAACCAAAA TTTTGTTCGA ATTCAATAAA TGTCTTGTTC AAAATAGAAA TATTGTAAAT 2640
 GTTATCGTCC AAAACTTCAC CAGTTAAGTA TTTGTTTTGA ATTAAAAATT GGCAGTTAGT 2700
 10 TAAGAAGTCT TGATAATCAC GATCGCAAAA ATAGTTTTCA CGTGCACTCT TAGCATCGCC 2760
 AAAAAAGTTA GCGACTGTTT CTGTTTCTCC TTTATTCGAA CGTTCATAT ATAATTGTGA 2820
 AAATTAGCT ATTGTATACT TTTGTTCTTT AGTTAGTTCA TTCAAAATAT TGGGCCTCCT 2880
 15 GAAATATCAT TTGTAATCTA TACCCAATTT ATTGCAAAAC AAAAATAAT TTAATATTT 2940
 GATGAACTG TGTTAATAAg CTTTAACAAG CCTTAGTTTG TATGGATCTA TAAAATTATC 3000
 20 TTTAATTGCA TAGGGTGAAA TAATATGTAG TCCATAACTT TTAAGTATT TTTCACTTAC 3060
 ACCAAATTTA TAAGCTTGGT AGATAATTTT AGTACAATAC GTAAATTTTT TGCTGTTCAA 3120
 ATTTAATGTA ACTAGATAAC GATGATTGT ATTCTCATAG TTTTCTTAA CCCATTGAGC 3180
 25 CGCTTTTTTA CCTGCACCAG GATAGCTGCA ACGATAAACT TTCATCCAAT CATTTTTGCC 3240
 ACTTGATAA TTATATTTAA AAGATTCGAA GGATTGTGTA GTTGGTTTGT CGCCAGGCCC 3300
 CTCAATTTGT AAAATCGTTT TATCATCAAT CGCGATACTA CAATGACCAA AAAATCscCA 3360
 30 CATGACAGGG CCTTTTGTA CAATAATATC ACCAGTTGT AATTGGAATT TGTCATCTTG 3420
 AATTTCTGAA TACTTATTAT CTGCAATTGT TTTTGGTGAG TTTATTGGGG ATACGACAAC 3480
 GAATAATATA AGTAAATTA TCGTTCGTTT AATATAGTTC ACTTAAAAGC TCCTTGTGTA 3540
 35 AGAAATATAT GTAAATAGTC TTAAATTAGA ATTGTAATCT TTAATAAGCT TGtAAGACTA 3600
 AAACATATCT TAAATATTAA AGTATGAGAG TGTGAAATGT CTATTAAGAA TnAAAAACAG 3660
 40 TCTGAAACAT CATTGAGACG TTCCAGACTG GATATAAAAT GAATTTTCATT TATAGCACA 3719

(2) INFORMATION FOR SEQ ID NO: 345:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1676 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

TTGCGTTGCC GCACCAAGAT ATTGAATGCC TAGCGATTCC GAGTATGCAA ACTGAACGCA 60

	CTCTTTAATA CGCGTTATCG CTTTTTGTA ATCTGCATCA TGATAACAAA TCATAACGCC	180
	ATAGCCACCT GCTGTCGGAA GATCATCTCC CGGCTTAATT ACTAACGGGA ATTCCCAATT	240
5	CTTAATCTCG TTTTCGAATT GCTCAATTTT TACAACTTTT CTTTTTGGTA AAAACTTCCC	300
	ATTTGTCCAT TCAGGTATTC TTGCTTTATT ATTTAAAGCA ACAAATAACG TTTTATCTAA	360
	TGCATAATAT TGCTGATTCA AGATTGTTTC ATCATGAATA TATTGAAAAT AAATCTTTTT	420
10	ATTTTCCTTA TGTGCCAATT GTTTGATCAA GTTTTCGTAA GATTGCTGAT TGTAAATGT	480
	ATAAATTGAG TTCGGTACTT CCTTACCAAT AACTTGAAAT AGCTGATGCA ATTTGTCTGT	540
	CGCACTAGCT TCGTGAACAA TAACAGGTAA TTGATTTGCT ATTAATAACT CCCTACCAGT	600
15	TAAAAAATTA GATTGATGTT CGTCCGGTTT CAACCATGGA TTCGATATAT ACGAAGGTCT	660
	TGACGTATAG ACAACATCTT TGTCAATATA ATCACTTAAC GTTAAGTTCG GCTCATTACC	720
20	ATTATTTGTC ATTACTTCCC ATTCCCTTTC AAATGCGCAT GCTCTTCAAT AATGTCTTGA	780
	TAAACGTCTT GATTTGTAAT TAACTCTAAC CCCATCAACG CCATTATTTT AGCGCCTTTA	840
	ATTAATGCTT CATCACCATG TACACTCGCA GCCGCTTCTC TAAATCTATG CGTATGTCCT	900
25	ACTAAATTAC GTGATCCTAT TTTAATATGA GGATGTATTG TTGGCACAAc ATgaCTTACG	960
	TTCCCTGTAT CCGTAGAGCC ATAACCAAAA TCATCATCAA TAACTGCTTC ACCAACTTCT	1020
30	TCAGCATATT TAGCAAATAA ATCATCTAAT TTCGGCGTTT TAATGAATTC ATTCACACCG	1080
	TTTTGAATTC GACCAAATTC ATAATCACAA CCAGTCTGTA TCGCAGCTCC ACGTGCATT	1140
	TGATTTACTT TTTCTGTTAA TATATCCAAT TCTTTACGCG TCATTGCTCT AGTATAAAAA	1200
35	CGAGCATGTG TATAGTCTGG AATAATATTA GCTGCTTTCC CGCCATCTAA AATCACACCA	1260
	TGCACACGTT GATCTTTTTT AATATGTTGT CGTAGTTGTG CTACACCATT AAAATAACTA	1320
	ATCATAGCGT CTAATGCATT TAACGCTTCA TCTGCATTTT CAGAGGCATG AGCACTTTTT	1380
40	CCGTAAAATT TAACATCTAA AACATCCACT GCCAAAGTAT CAATCGTTTT ATAAGTTTCA	1440
	TTTCCCGGAT GAATCATTA AATGCAATGTCT ATTTGATCAA TCACACCAGC CTTGACATAA	1500
	GAAGCTTTAG CGCTACCATT TTCCCCACCT TCTTCAGCTG GACATCCAAG AACGACTACT	1560
45	TTACCACCAA TTTGGTCAAT CACTTGCTTC AAACCAATTG CACCAAGAAC ACTTGCAGTT	1620
	CCAATGATAT TATGACCACA AGCATGACCC AATCCTGGCA AAGCATCGTA TTCTGC	1676

50 (2) INFORMATION FOR SEQ ID NO: 346:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1294 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

5 TACAGTAGGA ATCATAAAAC CTAATACAAC AAATACAAA CCATTTAAGG CATAACTAAA 60
 TGTGTTCCAA ATTTGATGGT AATTCATTG TAGTTCCGTT TGTGCTCTAA TTAAACGGTC 120
 10 GCGTTCTAAA CCATGGATTA GACCTGCGAT TACAACTGCA ATGATACCTG AAGCATGAAC 180
 TTCTTCTGCT AAAAAGTATA CGACAAAAGG AGTTAATAAT TGAATAAAAG TTAAGGTATT 240
 GTTATCTTTT AAACCTTTAT TAGCGGTAA GTCTATACGT ATTCTAACGA CAACGAATCC 300
 15 AATAATTGCA CCAATAAGTA CACCTAGTAT TGTGAAATG ATAAATTGTT CAACAGCTTG 360
 GAATAATGAA AAGGTACCAG TTACTAATGC AGTAACAGCA ATTTTAAATG AAATGATACC 420
 TGCTGCATCA TTGAGTAAAG ATTCACCTTC TAAATCGTC ATAGAACCTT TAGGTAATAA 480
 20 TTTCCGCGT GTAATAGCAG ATACTGCTAC TGCATCAGTA GGACATAAAA TTGCTGCTAT 540
 TGCAAAAGCG GCTGGCATTG GTAAGGCAGG CCAATCCAA TGTATAAAAT AGCCAACACC 600
 GACTACAGTT GCAAACACTA GTGCCATTGA CATTAAAGT ATAGGTTTAC GATATTCTAA 660
 25 TAATTTTGTG CGAGAGACGT GGGTACCTTC CACAAAAGT AGTGGCGCGA TAACGGCAAA 720
 CATAAATACT TCAGAATTGA ATTGGAAATC AACTTGTATT GGAATAATGA AAATAACGAC 780
 30 ACCTAATGCA ATTTGAATAA AGGCAGTAGG AATTGTGGG AATCGATTAT TGATAACCGA 840
 ACTAATAATC ACAGCAAAA TAAAAATTAA AAATGCTTCT AATAGTGCCA TACAATACTC 900
 CTCAAAATTT TAATAGTTAA TATTTTATCA CTTTAAAGGC ATAATGACAT AGATATATTG 960
 35 ATAAATGAA GTTATTTTCA AAAAACTCT AGTATCGGTT GAACTGATAC TAGAGCGAGA 1020
 TGTTTAAATT ATTGATTGTC ATATCTGAAA TGACCGCTGT CATTTTGTG TGTTCATAC 1080
 GCGAGCTTTT CAGCATTCTG TTTGTATTTT TtATAAAAGa AAAATAaAAA TATnAACCaG 1140
 40 AATGGCGAAA TATAAATAGC TGCTCTTGt TCGTCACTAA AGAATAATAA AATGAATACA 1200
 AAGAAGAAGA ACGCTAGAAT AATGTAAGCA ATAGGCTTAC CACCAATCAA CTAAATTTA 1260
 45 CTGTTTTTAT GTGCCTCAGG ATGCTTTTTC AAAT 1294

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1935 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	ACATGATAAT GATGACGCTA TTAAACACG TTTTATTATTT TTCATTGTTA TAACCTTCTT	60
	TCGTATGATT GATATTTGTT GATATGTATC GACATGTGAA TAATATCACA AAAACAGAGA	120
5	ATATATAATTT AACTATTTAT TAAATGATTT TGTTAATATT ATTAAATACT TTATCCTCTT	180
	TAAAAATAAT GTGTGTACAA AGTCATTAAT TTAGCAAATA TTTTATTTA GTAGTTAATA	240
	ACCATCGATT TGAAATTTAT ATATAATTAT TAGCTAAATA ATATCCTGCA TCTTTCTCAT	300
10	ACAATTTACT ATAAATagC ATATCCGATA TCAGCGTTAA TAAGATCGTT GATACTAGmC	360
	AGTTAATTTT ATAGAACGAA ATCAAATAAC ACACTACTTT CTGCATTTTA AATTATGTTT	420
15	AAGAATCaNA ATTATGTTTA nATAAATATA TATACTACTT TGAAAGGTGT GAGCTTAATG	480
	ACAACTTTTA GTGAAAAGA AAAAATTCAA TTACTAGCAG ATATTGTTGA ACTACAACT	540
	GAAAATAATA ATGAAATAGA CGTTTGTAAT TATTTAACAG ATTTATTCGA CAAGTACGAT	600
20	ATTAAATCTG AAATTTTGAA AGTTAATGAA CACCGCGCCA ATATCGTTGC AGAAATCGGT	660
	AACGGCTCAC CTATACTCGC ATTGAGTGGT CATATGGATG TTGTTGATGC AGGAAATCAA	720
	GATAATTGGT CATATCCCC TTTTCAACTG ACAGAAAAAG ATGGCAAATT ATATGGCCGA	780
25	GGCACTACAG ATATGAAAG CGGTTTAATG GCTTTGGTCG TATCTCTAAT CGAATTAAAA	840
	GAACAAAATG AATTGCCTCA TGAACGATT AGATTACTGG CTA CTGCTGG CGAAGAGAAA	900
30	GAACAAGAAG GTGCCAAATT ATTAGCTGAT AAAGGCTATT TAGACGATGT CGATGGCTTA	960
	ATTATTGCTG AACCAACTGG ATCTGGAATT TATTATGCAC ATAAGGGGTC TATGTCATGT	1020
	AAAGTAACTG CAACTGGTAA AGCTGTCCAT AGCTCAGTTC CATTATTGG TGACAATGCA	1080
35	ATTGATACAC TGCTTGAATT TTATAATCTA TTTAAAGAAA AATATTCAGA GCTTAAACAA	1140
	CAAGATACTA AACATGAATT AGATGTTGCG CCTATGTTCA AATCATTGAT TGGAAAAGAA	1200
	ATTTCTGAAG AGGATGCAAA TTATGCATCT GGTCTTACAG CTGTATGTTT GATTATAAAT	1260
40	GGCGGcAAAC AATTTAACTC TGTACCAGAT GAAGCTTCAC TTGAATTTAA CGTAAGACCA	1320
	GTTCTGAGT ATGATAACGA CTTTATAGAA TCGTTTTTCC AAAATATCAT TAATGATGTG	1380
45	GATAGCAATA AGCTTTCACT CGATATTCCA AGCAATCACC GACCTGTAAC AAGCGATAAA	1440
	AATAGCAAAT TAATTACTAC GATTAAAGAT GTAGCTTCTA GTTATGTAGA ACAAGACGAA	1500
	ATATTTGTTT CAGCGCTTGT AGGCGCAACA GATGCCTCTA GTTTCTTAGG AGATAATAAG	1560
50	GACAATGTTG ATTTAGCCAT TTTTGGACCA GGTAATCCAT TAATGGCACA TCAAATCGAT	1620
	GAATATATTG AAAAAGATAT GTATCTGAAA TATATTGATA TTTTAAAGA GGCTTCCATT	1680
55	CAATATTTAA AAGAAAAATA AGAACGATGC TGTCAGCTGC CCTATTCGCG TGCTGGCAGT	1740

TCAAATATCA ACAAGCACAT TTTCATTGAT TAAGTGATGT AAAACTGAAA TTATTGTGCT 1860
 GATTGTGCAT ACATATATTG ACTAATGGGC ATATAAAAAG ATAGCCTCTA ATAGTnACAT 1920
 5 AAACTCGTAA AAnCC 1935

(2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1351 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

CCTTnCCTA AACAATTTTT AGATTTAGAC AACAAACCGA TTTTAATCCA TACATTAGAA 60
 20 AnATTTATTT TAATTAATGA TTTTGAAAAA ATTATTATCG CGACGCCACn ACAATGGATG 120
 ACGCATACGA AAGATACACT TAGAAAATTC AAAATTTCTG ATGAAAGAAT TGAAGTCATT 180
 CAAGGTGGTA GCGATCGTAA CGATACAATT ATGAATATCG TTAAACATAT TGAATCAACA 240
 25 AATGGTATTA ACGATGACGA TGTCATTGTG ACACATGATG CAGTTAGACC ATTTTAAACG 300
 CATCGTATTA TTAAAGAAAA TATTCAAGCT GCTTTAGAGT ACGGTGCAGT AGATACAGTG 360
 ATTGATGCTA TAGATACGAT TGTTACATCT AAAGATAATC AAACGATTGA TGCAATTCCA 420
 30 GtGCCGTAATG AAATGTACCA AGGTCAAACA CCTCAATCGT TTAATATTaA TTTATTAAAA 480
 GAaAGCTATG CACAGTTGAG TGATGAGCAA AAGAGTATTT TATCTGATGC TTGTAAGATT 540
 35 ATTGTAGAAA CAAACAAACC GGTTCGACTT GTAAAAGGTG AGTTATATAA CATTAAAGTA 600
 ACAACACCTT ACGATTTAAA AGTAGCGAAT GCTATTATTC GAGGTGGTAT TGCCGATGAT 660
 TAATCAAGTA TATCAATTAG TTGCACCTAG ACAATTTGAA GTTACGTATA ACAACGTAGA 720
 40 TATTACAGT GACTATGTCA TTGTACGTCC TTTATATATG TCAATTTGTG CTGCCGATCA 780
 AAGATATTAT ACTGGTAGCC GTGATGAGAA TGTCTTATCT CAGAAATTGC CAATGTCTTT 840
 AATTCATGAA GGTGTTGGTG AGGTCGTATT TGACAGTAAA GGTGTGTTTA ATAAAGGTAC 900
 45 AAAAGTAGTT ATGGTACCGA ATACGCCGAC AGAAAAAGAC GATGTCATTG CTGAAAAC TA 960
 TTTAAAATCG AGCTACTTCA GATCAAGTGG ACATGATGGG TTTATGCAAG ATTTTGTGTT 1020
 50 GCTAAATCAT GATAGAGCTG TACCACTACC TGATGATATT GATTTAAGTA TTATTTTCATA 1080
 TACAGAGCTT GTAACAGTAA GTTTGCATGC TATTCGTCGT TTTGAAAAGA AATCTATTTT 1140
 AAATAaAAAT ACATTTGGTA TTTGGGGTGA TGGTAACTTA ggTTACATTA CAGCCATTTT 1200

GAGTCACTTC TCATTGTGTTG ATGATGTCTT CTTTATTAAT AAAATACCTG AAGGCTTAAC 1320

ATTTGATCAT GCATTGAGT GTGTGGGTGG T 1351

5

(2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 411 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

TCATCAAGTC TACGATAAAT TAAGTCCATA TCTAAAGGCT CGGGGTCGAC AGTTTGTAAC 60

GTATAACCAA CTGCACAGTG GCTACAACGC ATATTACAAA GATTTGTAGT TGTAATTCG 120

20

ATGTTACTTA AAGTTAATTG GCCATGTTCT TTAACATCGT TATATGCTTC CCATGGGTCG 180

TTTTGAATAC TTATTTTAGG CTTGTTATTA CGCATTTTAT AAACCTCTTA ATTGTTATTT 240

25

GATACCAATT TGATACCGTT TAATCAAATA TGCTCATAGC TTGATGTTTT TTATCAGTAT 300

ATAAATGAGA GTACGTTTGA ATTGTTTCTG TAATGTTAGA ATGCCTCATT AATTCCATTA 360

ATAAATACAT ATCTACACCA TTATTAATTA AaTAGCTAGC GTACGAGTGT G 411

30

(2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1639 base pairs

(B) TYPE: nucleic acid

35

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

40

TCATTTTCAT AGGTTATTAC GCAGATCAGC ATAATAATCC ATTCCATATG AGTcCTTATT 60

TTGGTTATGC AGCACGTCTA TTGGCAACAA GTGGCATTGA CTATACGTAT GTAAGAATGG 120

45

CAATGTACAT GGATCCACTT AAACCATATT TACCAGAATT GATGnATATG CATAAACTGA 180

TTTATCCnGC TGGCGATGGT CGTATTAATT ATATTACTAG AAATGATATT GCTAGAGGTG 240

TCATTGCTAT TATTAAAAAT CCAGATACTT GGGGCAAACG CTACTTATTA TCAGGCTACA 300

50

GTTATGATAT GAAAGAACTT GCTGCAATTT TATCTGAGGC ATCaGGCACA GAAATTAAAT 360

ATGAGCCCGT TTCATTAGAG ACATTTGCAG AAATGTATGA TGAACCTAAA GGCTTTGGTG 420

55

CATTATTGGC ATCAATGTAC GACGCAGGAG CAAGAGGACT ATTAGACCAA GAaTCCAATk 480

TTAATAATAA AGGAGCGTTA TAGTGAATAT CATCTCAACA ATTtTAATCA TATTTGTGGC 600
 ATTAGAGTTT TTCTATATTA TGTACCTTGA AACGATTGCT ACAACTTCCA AAAAGACTAG 660
 5 CGAGACATTT AATATAAGCG TCGATAAATT GAAAGACAAA AATATTAACC TACTTTTGAA 720
 GAACCAAGGC GTATATAACG GTTTAATCGG AGTTTTGCTA ATATACGGTT TGTtTATCAG 780
 CAGTAATCCA AAAGAAATAT GCGCAGCTAT TTTAGTGTAT ATCATTGGCG TTGCTATTTA 840
 10 TGGTGGCCTT TCAAGCAATA TTAGTATCTT TTTCAAACAA GGCACATTGC CAGTATTGGC 900
 ACTCATATCA ATGCTTTGGT AAGTATTGGT GTTTGGGGGG GTGGAGATGT AGTCGGAGGT 960
 15 TTGGAGGATT TGAGCGAATT GTGTGTGGAC TTTAGACTCA GAGTATTTCA TCCTAATTAT 1020
 TTCAAGCAGA GGTGACAGTA GCGTTGCCCTC TGTTTCCTTA TAAAAAAATT ATTTAATGAA 1080
 GAAAACCCAT ATCTGATTTA ATTTTCAGCT GATAAATACT CCATATATTA GAATGGCTAC 1140
 20 TTTATCTATT GCATCAATCC TTTAAACAA AAAACCCATG ATTTCGAAAT TCCCGTATGA 1200
 TGGGGTTCCT ACTCTCATGG ATCAGTTAAA TAAATATTAT CACTATCAGT TTATTATTTT 1260
 AATATTATTA ACAATATATG TAGTCGTAAA AGGAAAGAGG ACATGAGAAC TTCGGTGTtG 1320
 25 ATTGGCATTa CATAACGCTT CCAAACATAT TATTTGGTAA CAATAAGAAA CTATTTACAC 1380
 AATATATTTT GTATAGTAAA ATTATTTTAT AATATTTTAA TCCAATTGCA CAAGGACTGA 1440
 30 TTATCaTGGT ACCAGAAGAA AAAGGTtCTA TTTACTTTGTC AAAAGAAGCA GCTATCATAT 1500
 TTGCAATCGC AAAATTCAAA CCATTTAAGA ACAGAATTAA AAATAACCCA CAAAAACAA 1560
 ATCCATTTCT TAAATTACAT GAAACaAAA AATCTTAATC ACTTTtATTT ATAGcATTTT 1620
 35 TAATCTCAGA AATGCTATA 1639

(2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1816 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

AAAATCGCAT ATAGTAATAT GAATAACCAG ATTGTATCTA CAAAAAAGTA TATTGAAAAA 60
 50 CCAAGCGCAC CCATTAATAA TGCGAGAATA ATAATAATTT TTCTATTAAA GTGATGCGTA 120
 TCCGAAAATC TAGCAATAAT TGAATTTACT GTAACTGGC TAATCGCTGC AGATGCTAGA 180
 55 AGTAATCCAT ACTGATTTGT TGTCATACCT AAATCTTTAG TTGCAAAAAG AACAAGATAT 240

	TTCTyTATTT GTAaTAACGC TGCAAACATA TCCATAACCC CGCTTCTTAG AGCCCCTTTT	360
	AATTnATnAA TTAGGGGCTC TTATGCAGTT GGTGCATTAG CAACCAACTG TATTCCTTTG	420
5	TCCCCTTTTA ATTTATTaAT TAGGGGCTCT TTTGCTGTTG GTGCATTAGC AACCAACTAC	480
	GTTCAATTTA ACCGAATAGT TTAAAATTAA ATACAAACCT TAAATTAGTC TAAAACTACG	540
10	CCTTTGGTTG TTCAACAAAG CTCGCCATGA GATTTACAAA AGAATCAACT TGTGGCAATT	600
	GCAACATGCT CGGATCATAA CTCATAAATG TCGAACGAAT CAGCGGTTCA TTATCAATTT	660
	CTACTTTTTTC AAACCTCAAAT TGTTCCTTGC TGATATTTTT CATCATAATT TCTGGCAAGA	720
15	TTGTAACACC TACACCACTA ATCAACATTT CTTTGCAAGT TGCTACTTGA TCCACTGTAA	780
	TAGTTGCATG GTAATCTTGT TCTAAATTAT CGTTATACCA TTCCTTTTATT TGATTTATAT	840
	AAATCGGATC AGCTTGAAAC TCTATAAATG GTAACCTTGT AACATCATCT CGTCTATTTT	900
20	TTGGAAAAAT AAAATAATGA TCATCATTAA ATAAATGTGT GTTAGCTAAA TTCATTACCT	960
	TATTTCCACG AGTTATCATA ACATGATAAT CTCTATGATT TGCTTTAATT TGTTCAAGTTG	1020
	AACCAACTTG CACTTGTATT TCAACATTAG GAAATTGGGC ATTATATAGG CTCAAAACCTT	1080
25	CAGGAAGTAA GGTTTGTCCA ATCAAAGAAG AACACCCGAT TGATATTGTT CCATTCACTT	1140
	CACCAATATG TGCCTGCATT TTGTCAAAAA ATAATCGCTC TCTTTTCAAC ATGTCACGAG	1200
30	CATGCTCAAT AATCATTGTT CCTTCAGTTG TTGTAATCAA TTGTTTTTTT GTTCTGATAA	1260
	AAATATCTAC TCCAAAAGCA TTTTCAATAG CTTTTAGTCT TTGTGTAACA GCAGGTTGAG	1320
	ATATATATAA AATTTACGCC GCTTTACGTA ACGTTTTCGT TTCGTCTAAT GTTATTAGTA	1380
35	AACGATAGTC TTCAATCTTC ATAATTTCCC CCCATAAATT ATTCAATTAT TGAACCTTCA	1440
	TGGCTACAAG CATTCATGAG TTCATTACTA ACGAATAATT TCACCAATTT TATTGGTATG	1500
	GCTGCAGCTT GAATTACTTA GTTTTTCTTT TGTGTGTTGGT GATTTTTAGT TTGATTATAT	1560
40	TGCTTAGGCT TTATTGTTT GCTTTTTTCA ATATTAGTTT TATTTTGTGG CTTTGTATGA	1620
	TTTTTTTGAG CCTTGCATT AATTTTATTA AAGCAGTACA TGATTTTCTT TTGGAATCCT	1680
45	TTAAAATCAT TTTCTAACTC TGCCATAATT TGATGTGCAA TCATATATGC TTCATGAAAT	1740
	TGCTTTTTTG TAATTGCTC ACTTTCTAAT GCAAACATTA AATCATCTTC ATCTACCAAC	1800
	TCATaTcACC ACTTGG	1816

50 (2) INFORMATION FOR SEQ ID NO: 352:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9956 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

5	GCGTTGTCGT CGATGATTAA TAAAGTATGG GTATACCATT AAGAATAACG CTACCCAAAT	60
	gAkTGCTAGT GACGTGCCGC CAATGACATC TGAAAAGTAA TGTGCATGAA AATAAAGGCG	120
	ACAAAATAAT ATGCTAAGCC ATAATATTCC CATAACCAAC GCACTCAACA CTTTTGTTAT	180
10	TGTCTTAGCA GCAAGTGAAA TAATAATGAT CATTAAAGCG AAATATAATA ATGTGCTGGC	240
	GTTGGAATGT CCACTCGGAA ATGAAAAGCC TGTATCAACG GCTAAATGAT TATATGGTCT	300
	TGGACGTAAT ACAGTATCTT TAATTAATTT GTTCATGATG ACACCTGAAA CCAAATATGT	360
15	CACAAACCAA ACCGCTAAAT GCCTCTGTTT AATAAACAGT ATGATTGTGA CGATAATGGA	420
	AATCAACACG ACACCTTTGA CATCTCCAAT TTCCGCACTA AACGTCATAT AGTAATTAAA	480
20	CAAATTGTTA ACATACTGAC GTTGTGGCTC ACCGAAATAA TCTGTAAACC ATGTTAATGA	540
	TCCCATATCT ATATTTTTAA GCCATTCTTG ATTTGTCACT ACACTGTAAA ACATACCTAT	600
	AAATACAATC AGCGCGATTA AAAATAAAGG CACTGTcATT TTCGGTGATG TTAATTTTTT	660
25	ATCTATCATC TTACAATCTC CTCGTATCAT CATTTTCATT TTACAAATGT TATCCATAAT	720
	ATCAATGTGC CACAAATTTT ACTTTACCGA CAATATCAAA ATTATAAAGT TCATATTGTT	780
	ATGTATATTG CAAATAAAAC ATTGTATAAT TGAAATAACA ATATTTTGCT ATTTTCAATT	840
30	TAGTACGATT TATATTTATT ATACAGAGGG GGTAAGGCGT ATCAATAGAG TTATTTTTGT	900
	CTATATAGCG TTAATCATTa CATTAGTTAG ATTCTTTACC CCTATTcATC CATCATTTTT	960
	AAATTTGATT TACTGGATAT TTGTATTATA TTTTATTcCT ATTATACTAT GCGTTATCGG	1020
35	TTTCAAGGCC GAAAAACTTA TTGCAACAAT GGTCATTATA CCTAATTTTT TAGGAATACT	1080
	TTATCGATTA TATGCCTACG TCACACATAT ACTCTTTATG TAAAAGGATA GTGGGCATGT	1140
40	CTCGCAACAA ATATGCTTGC GTCGACTTGT CACCATTcGC AAATTTTATG ACATACGCCT	1200
	TTCACGGGCT ATATTCAGAC CCACGCATTc ATCCACGTAA TAAACACATC ATGTAATAGA	1260
	AAAACAGCAC ACCCAAATAT ATGGCGTTGC GCTGTTTAAc CAAGCATACT TCTATAGCTT	1320
45	TAATAAGCCA GCAGAAGCAT ACCTAACCTT CTTAAATATG CTTTTCCAAA TTATCCTCAA	1380
	GTTTGAATAC GATAATACGT TCACCTGTAA CTGTACTTAA ATCACTATGG AAGCTCATCA	1440
	CTTTGATACC TGTAATTTTA AAAATGATAT CATTCAAATC TTGCTCACCG GATTCAACTA	1500
50	ATTCAGAACG TGTTCGTTTA ATATTTAATA ATCCTTCATT CGTACTACAT ACACGATATT	1560
	CAGCTGGCGT TAAGATACCT TGTAAACTAA TAATCACCAT ATCTCTTAAa ATGTCTGATT	1620

55

	TTTCAGCTTC	GATTTACCT	TTCGTTCTTT	TCATATCACT	AACTCCAATA	ATATTAAAAT	1740
	TGATTACTTC	ATCTTTGTAT	CGTTATCAG	ACATCAACTA	TTACATTAAG	TTTATCATTT	1800
5	TTAGTATATT	TTAAGAAGCT	AGAACATTGT	AGATATGATG	ATATATTAGT	TACTTAGCAT	1860
	CGCAACATAT	CATCGTTAAA	TCCAACTTTT	AAAACGCCCT	TCCTCATTAA	CGCTCATTAA	1920
	ACGCAGCCAA	TGATTAGACA	CCTTCCTAGC	GAAATGCTCA	TTATTGCGGA	GTAGTCTTGC	1980
10	TACAACATAG	TCGGGTGCCT	GAATAACGAC	AAGTAAACGA	ATTGGCGAAT	GATACATCGT	2040
	CCGATCAGCA	GCCATAACAG	ATTGCCATGA	TAAGCCATAC	ATCAGATCAC	TCGCATTACC	2100
	TTGCATGACA	CCAACACCTG	ACGTGACGGT	TTGTGTCGCT	TTATTTCCAC	TTCCGTAAAA	2160
15	ATGCGGCGCA	ACTGTCGACG	CATAATATTG	TAAATTAATC	CATTGTGCCA	CAAGTGCCGG	2220
	ACCAGAAATG	ATGGTATTTA	ATAATGTGCC	ATCTTTATCT	TTACGCCAAT	CATAATTGTG	2280
20	TAAAAATGTC	CGCCCTTCTA	AATCAATGCC	TTTTGTTAAT	TGGCGTCGTC	CAATTATAAA	2340
	TGATGCATTT	TTAGCCAATC	CCCATTTCTG	ACGTACCTCA	CTCCAATCAC	TCGCAAACCG	2400
	CTGCGCTTCT	TCCACTGGAT	GATTCACACG	ACCAATCGTT	GGCAGTTTGT	CCAAACGTTT	2460
25	GCGATTGCGg	TGtTCAGAAA	TCATCGGCAT	CGCGTCATT	AATGATTCAT	ATGCATCTAA	2520
	AGCAATAGAA	GATAATGTGT	CTGGCACATA	TACCCATGCC	AACGTATCAG	TAGACGTATG	2580
30	ATGTTCTGCT	ACCGCAAAAA	CAGTTGTCTC	TGGAATATAC	ACACCTGATT	GTTTTAATCC	2640
	TTGTCTGACA	TTTGACGAT	TACATATCAT	CGCTAATAAC	TTAGCATTAA	AACCGCTTGA	2700
	TGCGCCACCA	CAAGCCCCAC	ATTCAAGTGA	TGCATGATGT	GGATTATTGT	GAGAATGACT	2760
35	AGCATGACCT	GCTAACACAA	CGAACGGCGC	AAATGCTTCG	GTTAAATCCA	TCAATTTCAA	2820
	CGCTTGTAAC	GCGAAATCAA	TTTGCTCTTG	CTCAGTAAAT	CCAACAGGTA	AGTCTGATGT	2880
	TCGGTCAAAC	TCACGATCAA	TCGTCAACTT	TGTTTCAGGC	TTTTTCAACC	ACTTTTGT	2940
40	TATTTTTTGT	AAAGACGCGC	GACTTTTTCT	AGGCATAATC	GAATTGACAA	TGGTACTTAA	3000
	GCTTAAAAAT	GGCCCACTTA	ATTCAAGCAA	TAACAGACTA	GGCATGACAT	TATTTTTCAT	3060
45	CAATTTAAAT	GTGTAAAACA	TCGATGACAT	TGTCTGTTGC	TGTTGTCGAT	AAACATTCAT	3120
	ATCGTAGCGG	TCTGCAAATT	CTTTAATGCG	ATATGCCGGC	GGTACCATGA	CAGGTAATGA	3180
	ATCATGTTTG	AATTGTTTCGT	CTACGGCATC	TTTTTTGAATA	GGTAATCCAA	AGAAGCCTGC	3240
50	AATACCAATC	GTTTCAAAGG	GCCCTGCTGC	TTCGATATGT	CTACGAAATG	GTTCTGAACG	3300
	AACATCTATA	CAAAATGCAA	TTTGGCGCTT	CGTTGATGTG	CCCATCTGAT	TTAGCTCGCT	3360
55	ATTATTTTCA	TCAACTGCTT	GTGTGTCATT	TAACAATACT	GAATGTGGCT	GATTAGCGTT	3420

	TGCTTTAATT	TTTTGTTTTA	ACTGAGATTG	GTATGTCATT	TCCCAGGCAA	TTAGCCATAA	3540
	ATTTTTAAAT	ACATTTTTAT	TCATAGTTGC	TGCAAAATGA	ATAAACGTTT	GAATTTTCATT	3600
5	GACGTCATGT	TGTAGTAATA	CATCGCTAGG	CATATCACTG	TAGTAACACC	ATGATGCAAC	3660
	AGTTTGCTTA	AACCAATTTT	CCGATCTACT	TTCACAATCT	TTAGCGACTG	ACTTAAACTC	3720
	ATCACCAACT	AGCAATTGTT	CGACAATAA	CCGAATTGCC	AAATAATCCG	TTAACAAATG	3780
10	TTGTTCAAAG	TGATGCTGTT	GTGAACGGTA	ATACAACATA	CCTGCCCAAC	CCGTAACGC	3840
	CAAAAGATGT	CCTTCAACAT	AAGCTTGGTA	GTCTTCCTGA	TCTATTGAAA	AATGAGTTAA	3900
15	TACTGACTCT	ATCGTCATTT	CAGGATCATT	GGGTAAGCCT	TTAATCACTT	GGCGCTGTGC	3960
	TTTAGTAAAA	CTATGGTCAT	GTTGCGCTAA	ATGCAACCAT	GCATGGTAAA	AACTTTGCTC	4020
	ACGCTTCGGC	ATTGTCCAAC	TCGATAGAAA	TTGATCGATA	TAAAGTTTCG	TCCATTTAAT	4080
20	CATTTGACGA	TTCACTTGTT	CGCTAAGTGG	CTCACCTTGT	TCATCTATTA	TTGCATCACT	4140
	CATCGGACGT	ACATCATAGT	GATGATATGA	TTCAGCCATA	TCACGTTTGT	ATTTTTCTAA	4200
	TAGTAGATCA	GCAACAACAT	CAACATTTGA	ATGATTCATA	TATGATGCAG	GTACGTCTTT	4260
25	TAATGTTTTA	ATGTTATCAA	TATAAAGATT	GATGTAGTGT	TGCGGGATAT	TGTAGTGATG	4320
	TTCAAGTAAC	ATATCAGTAA	CAAGTTGATT	AAAGACACTT	TCATCTAATT	CACCACGTGC	4380
30	CACAGCGCTT	TCTATTAATG	CTTTATTTGG	GAAAATATCC	ACATCTCGAA	CATCACGTAA	4440
	CCATTTTGCG	ACATCTTCAA	ACGTATCCGC	TTCTAATCCT	TCCCATGGAT	TTCGTGCTGC	4500
	AAAAATCGAA	ATTGGTGATA	ATGGTGTAAT	AACACGTTTC	GCATTTTCAA	TGACTGAATT	4560
35	GATATTTAAC	TGTGTTGTCA	TACCTTTCAC	CTCCTATAAA	TACTTCTTCA	AATAATTCGG	4620
	ATGACTTTCT	ATCGCTTTTCG	AGCGTGCTTC	ACCTAGATTG	ACTAACCACA	CGTACAATAC	4680
	CGCAAAGCC	TTAGAGTATC	GATGCCGCGC	CACCCAAATA	CTTAATAAAC	TGCCAAAGAT	4740
40	TAAAATAACA	ACACTAATGA	TGACACTCAC	TGTAGGCGGC	GTTGTCGCAT	GTGTTGTTAT	4800
	ATTTTGTAAT	ACAGCGTAAA	AATAATTATG	TGTGATGACG	TAGATAAATG	TCACGATTGC	4860
	AATCAAAATC	ATACCAACAA	GACGTGCCAT	GCGTCCTTTA	CTAAAGGCTA	CCATTTGATT	4920
45	CCAAGATACA	AGTAATGACC	ATCCTAGAAT	GAGTGCACTT	AACACTTCAT	ATGCACTTCT	4980
	GTCACTACTC	ATCCAAAATA	GAAATGCCAC	GATAATAGCT	AATACACGTC	CCATGACAAT	5040
50	CCAGCCATAA	GCGTCTTTAG	CAGATGCTTG	TTTTGGAATA	TTGAATCGCT	TCACGATAGA	5100
	ACCTGATTGT	AAAAATAATG	TTGCTTTAAA	AATACCGTGC	AATATTAAAT	GAATAATCGC	5160
	TGCTGAATAT	ACACCCAATG	CACATTGAAC	TAACATAAAG	CCCATTTGAC	TCATCGTAGA	5220

55

	AGAAATACTA GAAAGGATAA GTAATAATGA TAACGCAAAT CCATTATCAA ATATCGGCGC	5340
	AAAACGAGTT AGAATAACAC CACCTGCATT CACAATTCCT GCATGCATAA TTGCCGATAC	5400
5	TGGCGTTGGT GCCGTTACAG ATTCAATCAA CCATCGATGA AAAGGAAATT GTGCTGCCGG	5460
	TATCATGACA GCTAATACAA GTAGTACATT CGTCAACAAT GACCATGTCTG GATGAACTAT	5520
	ATGTTGTGGT ACCCGCCACT CGCCAGTCGC AATATAAATA GTTACAATTG CTCCAACGAA	5580
10	TGCAAGCCAA CCACATAAAA ATGTCATGCT TGATAATTTT GCAGACTCAC GTGGCACTTT	5640
	CCAAAAACGA TTAACGTTCA TCAGCAATGT TAAACATAAT AATGTAATAC CCCAGCAGAG	5700
15	TGCCATCAGT CTTAAGTCTT CAGACATCCA TGCTAAAGAT GCAAACGACG TAATCGCAGT	5760
	GAACAATGGA AAGTAATGTC TATAATGATG ATCACCTAGT AAATATCGCA TTGAAAACCTT	5820
	TTGAATAATA AAGCCAAGCG CCATTACAAA GCCAGCTAAT AACCAAGATA AACGATCTAT	5880
20	TTTAAATGGA CCTAAGACAT GTTGACCATG AATACCGAAA AAGCCAATGA CTGCAAATAA	5940
	TACTGGCATG ACTAGTATGT ATAAATGTAA TTAAATATAT CTCATTGGCA TAACTGGTGC	6000
	TAAAAACAAC AAGCCACTTA TCAATGCAAT GATAAGCGCA ATAACAAACA GTGAAAATAG	6060
25	CAATTGAAAA CTTAACACTG CATAACCTCC TTATTTCTAA TCTCTCGCAT AATTGCTTAT	6120
	GTATAAAAAAT AAAAACCTAC AATAGTAGAT TCTGTACATA ATGGCAGAAA ATTTACTATT	6180
	GCAGGTTTCA GTTTAACTAG ACACTGCATC ACGGTACGTT GATATACCTT GTTGCACTGT	6240
30	TCTCTTTAAG CGTGCTCCCA TGCACATATG TATATAAAAT GTTACTTCTG TCTGTTCAAT	6300
	TCATCTTCAT AAATATGCTT TGCCTAGACG AGACCTAACG TGTATTTCGT TTTAACTTA	6360
35	TAACATAAAA TATAATTAAA TTTCTGCTTC ATGTCAAATT CATGAGCTTA ACCTCTATTA	6420
	AACCAATGAT TGTAAGATT TTGTAAATGC ACCTGTACAG TTAGGCAGTA TTTCCCGTCC	6480
	TTTTAAATA AAAAATTCGC AGTTATGATC ATAACAATTC AAGTTAGGAA AAAATCAAT	6540
40	TACGCACAAG ATAACTATGT ACAATGAAGT TAACTCATAA GCAAAGGAGG TAATCTTAAT	6600
	GGGTATCATC GCTGGCATCA TTAAAGTTAT CAAAAGCTTA ATCGAACAAT TCACTGGTAA	6660
	ATAAGATTTT ATAACAAACA AAGGAGGTCT TTCACATGGG TATCATTGCA GGAATCATT	6720
45	AATTCATTAA AGGATTAAAT GAGAAATTCA CTGCTAAGTA AGTTATAAAA ATCTCATAGA	6780
	TATGAACATC TTATTTGAAG GGGGCCATTC ACATGGAATT CGTAGCAAAA TTATTCAAAT	6840
50	TCTTTAAAGA TTTACTTGGT AAATTTTGTAG GTAACAATA ATCTCAAACA TTAACGATCA	6900
	ACAACTCATC ACTATGTTAA ATCAACATAC AGGAGGACAA AACGATGGCT ATTGTAGGTA	6960
	CTATCATTAA AATCATCAAA GCAATTATCG ACATTTTCGC AAAATAATTT AAGCGAATTG	7020

55

	TTATTGATGT GAGGTGAGTC TTGTTAGTTT GTTGCAAATA AATGGTCTTG GTGTTTTTTG	7140
	TATAGGACGT TCTTAGTGGG ACATACGGAA TATTCGTGAT CTTTGTAGTC TGACGCGTTA	7200
5	TATTTTTGTG GCGTGTTTTA TGTTTGATAC TCGAGTTCTG AGACATTCAT GATTTGGCAT	7260
	GCGAAATCTT AATGATTTTC ATGATCTAGC GCAAGATATA TTGGCCACGT GCGGAATTGC	7320
10	GTTGCACGTT TAGACTGAAA CACTCGTGTG ACCGTAAGTG TTAATAGTAC ATTGATAGCT	7380
	GCATTTACTT CACTCATTTT TATGACTGTT AAACAATGAT TGTACCTTCA ATTAACAGTT	7440
	GGTACGATGG TTTTGCCATT TTTCATCAAC GTAAATATAA AAAGGACTAA GACACATACA	7500
15	TGTCCTAGCC CTATGGATAA AATGCAAATT TCTGCTTTAT CAAAACATC ACACTTTAGA	7560
	TAGATTGAAA ACAAAAAGAT CCTAAGAACA CCTTAACTTT TTATTAATTG TCATAAATTG	7620
	CAAACAATTA AGCCACAATT CAAAATGAT TATACTTCAT TCAACTTATC GTGCTGGTCT	7680
20	AATTTGCCAT TGATATGGAT CTTCAAATTG TTGCCAATCT GCATCAATTT CTTGCGCATT	7740
	GACTAAGCAT GCGTCGAGTT CTTTTGTAA TTTTCTTCA TCTAATTCTG TACCAATAAT	7800
	GACAAATTGT GTATGACGAT CGCCATATTC TGGATCCCAT TCAGCTGCGA CATCTTGACG	7860
25	TTCTGCTAAT ATTTGTGTTT GTTGCGCTTC AGACATACTA GCCACCCAAT ATGTAAGTGG	7920
	ATGAATATTG CAAGATGACC CTGCTTGAGA TAATAAACAT GCTACGTGAT TGTATTGTGC	7980
30	TAGCCATACG ATACCTTTTG ATCGAACGAC ATTATTTGGC ATGCTTTCTA ACCAATCATT	8040
	GAACCTTTTA GCATGGAAAG GTAGACGACG TTTATATACA AACGATGATA TACCATATTC	8100
	TTCTGTTTCA GGTGTATGCG ATGCATGCCC ACCAGACTCA AGTTCTTTGA TCCATCCTGC	8160
35	TGACTCGCTC GCTTTTTCOA AATCAAAACG CTGCGTATTC AAGACTTCTT TTAAATCTAC	8220
	TTCAGAATTT GTTGTCTTAA TAATTTTAGC AGTCGGTTGC AATGCGCTTA ACATTTTTTC	8280
	TAACTTCGCT AGTTCTTCTT CACTAATTAA ATCAATTTTA TTAATAATCA ATACATCACA	8340
40	AAATTCAACT TGGTCAATTA ATAAATCAGC AATCGAACGC TCATCTGTTT CGTCAACGCT	8400
	TTGATCACGA TCCATCAATA AATCTTCTGA GTTGATGTCA TGTACGAAGC GGTAGCATC	8460
45	CACAACTGTA ACCATTGTAT CTAAACGGCA AATCGCTGTA AGATCAATGC CAAGTTCATC	8520
	ATCAATATAT GAGAAAGTTT GTGCAACAGG TACTGGCTCT GAAATCCCTG TTGACTCAAT	8580
	AACAATTTGA TCGATGCCAC CTTTTTTCAC TAAACGCTCA ACTTCTTTTA ATAAATCGTC	8640
50	TCTAAGTGTA CAACAGATAC AACCATTAGA AAGTTCGACT AATTTTTTCAT CTGTACGCGA	8700
	TAGTCCCCCA CCATCTGCGA CAAGATCTTT ATCGATATTT ACTTCACTCA TATCATTTAC	8760
55	AATTACCGCG ATACGTCGAC CTTCTCGATT TTGTAAATA TGATTTAACA ACGTTGTCTT	8820

ACTTCAATTT ATTTGTAAAT AGGAATAATT CTGTTTTACA TTATATAGGA GCGTTTCCTC 8940
 TTTTCGCAATC TTCGATAATA AAAAAATAGT ATACTTAATT AAATTATTGA GCGCTTTACT 9000
 5 TTATAATGGA GACAAAGATA TATCTCACGA AAGAGAATCG AGGTGTATAA ACATGTTATT 9060
 TGTCAATTTA GTTTTATATG TTACTGGTAT TGCATTTATT CTACTCAGTG TTTTGGTTC 9120
 AAAGACTGAA GGATTATCTA CGAAACATAC TTTATATACC ATTGGCAGTG CTATTATAAC 9180
 10 GATTGCTATT TTCATTTCAA TTGGCTATGC CATTCAATAC TTAAGTGCAG CGCTTTATGG 9240
 TTTGTAAGGT GAAGGTGATG AGTAACGGGT AGTTCGGGAG AGGTAACTT GCGTTGATTT 9300
 15 TGATAAAGTG ATCATAGCTT TTAGTACTTG AGGATTTTTA TTGTGCTGT TACGAATGTG 9360
 GTCATGTTTA ATGCGGGACA GTAATTTAAG TTGTTTTTTT ACAATTGAGA GTGTGATATT 9420
 TCGATTCGGT TCGAATTACT TTACATGGGA ATAATATAAA TTAAAAAGAA GCGGCCTAGT 9480
 20 GTCAGTTGTG AATATACTGA ACATTGGTCG CTTTATTTAG TAGTATGATA TGTAGTTTAG 9540
 CTATTAATTT TTTTCAGGTC ATCCTTAATG CTGTCTATCT CAGACATGGC ACTTTTAACC 9600
 CAATCTCCTT GAGCTGCACC TTAAAAATTA GCTTTAAAAG ctTCGCAATG TTGCGCCATT 9660
 25 TGTTCAAATTA ATACTTTTTT TTCACCTTTT AATCCGTTTT CAATATCTTT GTATTTATGC 9720
 TTATGTTTCA GTGCAATAAC TGTGCGAATA TTTCTTTTTT GCGCTTCCAT TTTAGATATG 9780
 AGATTAAGTG TTTCTACTGT AGTACTTATA TCTGGCATTG TTAAGGTCAT ATCTGGTTCT 9840
 30 ATTAGAGTCA TTTAATCTCC TCCAAATTAT CAGTCACTTA GCTTATCTAA CTGCTTTTCA 9900
 TAAGACTTTT TTAAGTCTTC TTTATATTCT TCTAATTTCC CATTCTTGCT TTCTGA 9956

35 (2) INFORMATION FOR SEQ ID NO: 353:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2411 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

TTTCTTTTAA CAGaTTTATC cCATTAAATTG TTCTACAAGC CCACGaTGAG CAATATCATT 60
 TTTAGCAACC ATTAATAAAC CAGAAGTATC CATATCTATA CCGTGAACAA TACCTGGACG 120
 50 AATTTCTCCA TTAATACCTG ACAAATTTTT AATTTGATAC ATTAAACCAT TAACTAATGT 180
 ATTGGTATAA TGCCCTGGTG ATGGATGAAC TACCATGCCT TTCGGTTTAT ATACAACTGC 240
 AACATCGTCA TCTTCATAAT AAATATCTAA ATTTAAATTT TCAGGTAGAA TATCAGCTTC 300

	AACAACCTTTA TCGTTTGCAA CGACTAAACC TGCTTTAATC CAATCTTGTA TCTGGTTACG	420
	AGACCAATCA TTATTTAATT CAGGCAGCAA CTTATCTACA CGCATACCTG TTTGTTCTTT	480
5	ATCTGTAATG TTAAATTCAT AAGTCTCCAT TACTTAACCT CCTTCTCCTT TTTATTGGAA	540
	GTATCCTTTA ATAAGGCAAT AATAATTAAT ATTACACCAA TTGTTAAACT TGAATCTGCG	600
10	ATAITAAATA TTGGAAAATC ATAACCAAAA ATATTTGTAT CAATAAAGTC AACAACCTTCT	660
	CCTGTTAAAA TTCTATCAAT AAAGTTTCCA AGTGCACCTG CAAAAAGTAA ACTAATAGCA	720
	ACTTGACATA ACAAATTATA TTGAGCATCT TTAATAAAGA AATATACTAA GGCTATTAAAT	780
15	ATAATAATGG TAATAATAAA GAAAAATGTC ATTTTCCAC TCAATATTCC CCATGCAGCA	840
	CCATTATTTC GATGTGATGT TATGTTTAAA AAGTGCAGTA TCACTTCAA TGAATCTCCA	900
	ATTTTCATTG TAGTAGCTAT AATATATTTA GTAACCTGGT CAAATATAAC GACAAATACT	960
20	GCTATTAAAA TGGAAGTGCC AATAAAATAT TTTTGTGCA TTTTCGTTCC TCCAATCAAT	1020
	CGTCCATGAG ACAACTCTTT ATATTATAGC TTACACCTGC TAATAAAAAA AGTAAGCATA	1080
	TTACATTAAA TCTAATGTTA CTAACCTCAAT ACTTGATAAA CTACTATGTT TTGACATTAA	1140
25	ATATGAACTT AATTATTCAT TTATCATATT TAAGATGACA TTAAAAATTA GGAAAGCAGG	1200
	CTGGAACATA AATCCCTAAA AAGACAGTAG TAAGATATTT TCTAATTAAA AATTATCTTA	1260
30	CTGCTGTTCT CTATTTATAC AATACTTCGT ATTGAATGGC TTCGCTATGC CCATCTGGCA	1320
	CATTACTGTA AAATTCTATA AATAGAATTT TTGATGATGG GTCCCTTCCT AGGGTGCCGT	1380
	CTCAGCCTCG GtCTTCGACT GGCACGTCT CCTCAGGAGT CTCGCCATTA ATACTACGTA	1440
35	TTAACATGTA ATTTTACTTT TAAATACTTT AAAAAAATAA GACATGAATC GTCTACACTT	1500
	AATTGGACAA ATTCTATGAG AATAGATATT GTTAATTTAA GAAAGTAGGC TATTTTGAGT	1560
	TtCACTCGAA TGTCAGTTCG AGGAATAAAT AAAGTTAAAC GAGAGCTAGG TTTGTATTA	1620
40	ATGGCAATTA ATATAAGGAA AATAGCAGCT CAACGAGCTG TACATTATAA AATACATATC	1680
	AAAAAAGCTG ATTTCTATCA AATAATTAAT AGAAATCAGC TTTTTTACAT TGCCTAAGAA	1740
	CTTAATGTCC CAAGCCCTAA AACTTGTTGT TATTTATTTG ATTTAGCAGC GATACGTTTA	1800
45	TATCTTAAGT ACATAAATGC TAAAAGTATA AACCAAATCG GAATAAAATA AATTGCACGT	1860
	CTTGATCAAA CATTAAATAA TAATAACCCG AACACAAAAA TGAAGAATAC AAATATTAAG	1920
50	TAGCCCATAT ATTTGCCACC TAATAGTTTG TACGTAGCAT TTTTATGTAG ATCTGGGTTT	1980
	TTACGACTAT AATTGATATA TGCAATGATA ATCAGACCCC ATACAACTAA AAATAACACT	2040
55	GTAGAGATGG TAGTCACATA CGTAAATACT TTTGTGCGAT CTGGGAAAAT ATAGTTTAGT	2100

TTATTCGTCT TAGAAAAGTT CGGAGGTGCT TGTYGTTGAC TTGATAAACC GaAAAGCATA 2220
 CGGCTATTTG AGAATATACC ACTGTTACAT GATGAAGCAG CAGCGGTAA TACTACAAA 2280
 5 TTAATCAAGC CCGCAGCAA CGGAATTCG ATCAATGCGA ATrATTTTnc GAATGGACTG 2340
 TTATCAGGAT CAACTTGCTG CCAAGGGGTA ATAGACATGA TAACCGCTAA CGCCCCAACG 2400
 TTnnATATTA A 2411

10

(2) INFORMATION FOR SEQ ID NO: 354:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 605 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:

GnGAATTATT TTTAATAATG AAAGGATTAC TtnCATGGGT TTTTACTAGG AntACCCAGA 60
 AGGTCAAAT ATTTTGTATG CGCTAAGTCA ATATGAAGTT AAGCGACGCG GCGATATGGA 120
 25 AGAGGATCCA TCATATAAAC AACTCATTTT TTATTGTTTA CTTGAAAATG AGCATGGCGA 180
 GATATTAGTG TATGAACGAT TATCTGGCGG TGGAGAAGCT CGATTGCATG GACAATCTTC 240
 AATAGGTGTA GGCGGTCATA TGAATGATGT TCCAGGAGCA GAATCTATTA ACGAAGTATT 300
 30 GAGAGTTAAT GCACAGAGAG AATTAGAAGA AGAAGTAGGT TTAAGTGAGC AAGATTCACA 360
 AAATATGGAA TATATCGGTT TTATTAATGA CGATAATAAT GAAGTGGGCA AGGTACATAT 420
 TGGTGTTGTA TTTAAAATCA CTGTAAGTAC GAATGATGTA GAAGCTAAAG AAACAGATAC 480
 35 TTTACGAATA AAATGGGTTG AAAAAGGCAA CATAGAGTCA TATGATGATT TCGAAACGTG 540
 GAGTGCATTA ATCCTTCAAG ATTTATAATC AAACGAGGTG ACATATATGT CAGATATTAT 600
 40 TCCAG 605

(2) INFORMATION FOR SEQ ID NO: 355:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 668 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

TTTATTAGCA CATCCAAACT ATTCATATGT TGGACAATTT TTAAACGAAC TAGGATTTAA 60
 55

ACAATTAGAC ACTGAACATT TAGCTGATTT AAATCCAGAG CGTATGATCA TTATGACAGA 180
 TCATGCTAAA AAAGATTCTG CTGAATTCAA GAAGTTACAA GAAGATGCAA CATGGAAAAA 240
 5 GTTGAATGCA GTTAAAAATA ATCGCGTGGA TATTGTTGAC CGTGATGTTT GGGCAAGATC 300
 TCGTGGCTTA ATTTCTTCTG AAGAAATGGC TAAAGAACTT GTTGAATTAT CAAAAAAGA 360
 AAAAAAGTAA GGTGGAAGTA AATGGCTATA AAAGAAATAA GTAGCCAATC TGCCATAGAT 420
 10 CATAAAGAA AAAGACGCAC AACACTCACG TATATAGTGA GTTTGTGCTT TCTTTTATT 480
 TGTATATATT TAAATATGGC GATTGGTTCT TCGAAAATTA ATTTTAGCGA TATCATTCAC 540
 TATGTTACTG GTCATACAGA TACGAAAGCA ACGTTTTTAT TGCATAATGT ACGTATGCCA 600
 15 AGGATGATTG CAGGGTTATT TATTGGCGGT GCATTAGCGG TATCTGTTT GTTAATGCAA 660
 GCAATGAC 668

20 (2) INFORMATION FOR SEQ ID NO: 356:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 787 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:

30 ATACAAAAA ACATATCGAA AATAAAGCTA AAAGAACTA TCAAGTTCCA TATTCAATTA 60
 ATTTAAATGG TACATCTACA AACATTTTAT CGAATCTTTC ATTTTCAAAT AAACCTTGGA 120
 35 CAAATTACAA AAATTTAACT AGTCAAATAA AATCAGTACT GAAGCATGAT AGAGGTATTA 180
 GTGAACAAGA TTAAAAATAT GCTAAGAAAG CTTATTATAC TGTTTATTTT AAAAATGGTG 240
 GTAAAAGAAT CTTACAGTTG AATTCAAAAA ATTACACAGC AAACCTTAGTT CATGCGAAAG 300
 40 ATGTTAAGAG AATTGAAATT ACTGTTAAAA CAGGAACTAA AGCGAAAGCA GACAGATATG 360
 TACCATACAC AATTGCAGTA AATGGCACAT CAACACCAAT TTTATCAAAA CTTAAAATTT 420
 CGAATAAACA ATTAATTAGT TACAAATATT TAAATGACAA AGTGAAATCT GTATTAAAAA 480
 45 GTGAAAGAGG CATCAGTGAT CTTGACTTAA AATTGCGGAA ACAAGCAAAA TATACAGTAT 540
 ATTTCAAAAA TGGAAAGAAA CAAGTAGTGA ATTTAAAATC AGACATCTTT ACACCTAATT 600
 TATTTAGTGC CAAAGATATT AAAAAGATTG ATATTGATGT AAAACAATAC ACTAAATCAA 660
 50 AAAAAAATAA ATAAATCTAA TAATGTGAAA TTCCCAGTAA CAATAAATAA ATTTGAAAAC 720
 ATAGTTTCAA ATGAATTGT GTTCTATAAT GCAAGCAAAA TTACmATTAA TGaTTTAAGT 780

55

(2) INFORMATION FOR SEQ ID NO: 357:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 534 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

```

AAAGTAAAAA TAAATCTCCC TTTTAACTT TCGTTTCTGC CATAGCCATT GCTTCTTCTG      60
TGATAGTTGC TACAATATCT TTTCTTTCAC GGTAAAAATG TTCAACTTGT TCTGCTAAAA      120
ATGCAGCTTC TTCTTCGACG TCAGTCATCA ACAATTTCGCa AGCTAATGAT GCGTCATCTA      180
AACGACCTAC AGCATTAAAGT CTAGGTCCAA TAATAAAACC AATTGTTTCT TCATCAATAT      240
TGTCATTGTA TCCCGCTTCT TTTAGCAATG CTTTAAACAGA GGTCGGACAT TGATCATTTA      300
AGACTTTTAA TCCTTGTtTC ACTAATGATC GATTTTTCATC AGTTAAGGAT ACTAAATCCG      360
CAATGGTACC TATCGCAACT AATGCTTTAA AATAATCAGG TACATTTtCA ATCAATGCTT      420
GTGCTAATTT GTATGCAACA CCTGCACCAC ACAATTGTTG GAACGGATAA TTAAACGATG      480
GATGCATTGG ATGTACGATT GCATATGCTT CTGGTAATGT ACTACCAATT TCAT          534

```

(2) INFORMATION FOR SEQ ID NO: 358:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3621 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

```

GGTGAGTCAA ATTAAATGAA TCTAATAAGT CATAACTATC TATTTGTAAT GTGCAACGCT      60
TAACGCATAT ACAAATGAA TGTGCTGATA ATGATTTACT CAAATTAAAA GGTGATTTTT      120
ATTCAATGAT GAATGAAAGT TGCCTTTTAA TTTTGGTAA AAGTTAATGC GTCAGTGAAT      180
TGTGTAAGTT TTTCAAAAAG TAAAAAGAAA TAATAAAGGT GAATTATTAG AATTCCArAA      240
ATAATTCATT ACATTCATAA AGCATTTTAC AAATGGTAAG AAAATGAGTG TTACAAATCT      300
AAATATTGCA AAAGAAGCTG ATTTAGTCAC AAAAAATGTC CTATGTAATA ATTCGAGAAA      360
GATGCACTAT ATACGGTCTT CTTACTATTC AAATGTAAAA GTTGCTTATT TCGGTGGCTT      420
TTTGTTTTAT AAAAGTATAA AATTTTACTA TAATATATCT TGTAGAGAAC AATGAAATGA      480

```

	ATAGCAAAC	GTATTACTTT	GATACAAAA	TGGTTGTAAT	AAATATTTAT	CGATATGACG	600
	ACTTGAATAT	GATAAAGTGA	CATATTTATG	TATATGACTA	TTTCGCAAAA	TGTAATCGAG	660
5	GTAGAATTC	TTGACAATTC	TGTCAGTTTA	TAAGATGTTA	TAAATATGTA	GTGTATAAGG	720
	AGGCAAACAA	GATGACTGAA	GAATTCAATG	AATCAATGAT	TAACGATATT	AAAGAAGGTG	780
10	ACAAAGTCAC	TGGCGAGGTA	CAACAAGTTG	AAGACAAGCA	AGTTGTTGTT	CATATCAACG	840
	GTGGTAAATT	TAATGGGATT	ATTCCTATTA	GTCAACTATC	TACGCATCAT	ATTGATAGCC	900
	CAAGTGAAGT	TGTAAGAGAG	GGCGACGAAG	TTGAAGCATA	TGTCACTAAA	GTTGAGTTTG	960
15	ATGAAGAAAA	TGAAACTGGA	GCTTACATCT	TATCTAGAAG	ACAACTTGAA	ACTGAGAAGT	1020
	CTTATAGTTA	TTTACAAGAA	AAATTAGATA	ATAATGAAAT	CATCGAAGCG	AAAGTAACAG	1080
	AAGTAGTTAA	AGGTGGTTTG	GTTGTTGATG	TAGGACAAAG	AGGTTTTGTT	CCGGCTTCAC	1140
20	TAATTTCAAC	AGACTTCATT	GAGGATTTCT	CTGTGTTTGA	TGGACAAACA	ATTCGTATTA	1200
	AAGTTGAAGA	ATTGGATCCT	GAAAATAATA	GAGTCATTTT	AAGCCGTAAA	GCAGTTGAAC	1260
	AAGAAGAAAA	CGATGCTAAA	AAAGATCAAT	TATTACAATC	TTTAAATGAA	GGCGATGTTA	1320
25	TTGATGGTAA	AGTAGCGCGT	TTAACTCAAT	TTGGTGCATT	TATAGACATT	GGCGGTGTTG	1380
	ATGGTTTAGT	GCATGTATCT	GAACTTTCTC	ACGAACATGT	TCAAACACCA	GAAGAAGTAG	1440
30	TTTCAATTGG	TCAAGATGTT	AAAGTTAAAA	TTAAATCTAT	TGATAGAGAT	ACAGAACGTA	1500
	TTTCATTATC	AATCAAAGAT	ACGTTACCAA	CACCTTTCGA	AAATATTAAA	GGTCAATTCC	1560
	ACGAAAATGA	TGTCATTGAA	GGTGTCTAG	TAAGATTGGC	AACTTTGGT	GCATTTGTTG	1620
35	AAATTGCACC	AGGTGTACAA	GGACTGTAC	ATATTTCTGA	AATTGCACAC	AAACACATTG	1680
	GTACGCCAGG	TGAAGTGTTA	GAACCTGGTC	AACAAGTAAA	TGTTAAAATA	TTAGGTATTG	1740
	ATGAAGAGAA	TGAAAGAGTA	TCACTATCTA	TTAAAGCAAC	ATTACCAAAC	GAAGATGTTG	1800
40	TTGAAAGTGA	TCCTTCTACG	ACTAAGGCGT	ACTTAGAAAA	CGAAGAAGAA	GATAATCCAA	1860
	CAATTGGCGA	TATGATTGGT	GATAAACTTA	AAAATCTTAA	ACTATAATTT	AATATTTAAT	1920
	AGTCAACTCC	ACATGTTTAT	GATTGCATGT	GGAGTATTTT	TATGTAACAA	AATATACTCG	1980
45	GAATGATAAC	GTGGgACAAA	TTTAACTAAG	TGTTTTAAAA	GATAgAGTTT	TAAGTGCtGa	2040
	tTTTTATCAT	TACAGTAATA	AACTCATTTT	GAATACACAG	TCTCATGTGA	TATTATTAAA	2100
50	AAGATATaAG	AAAGAGAGGA	AGTTAGCTTA	TGACTAAACC	TATAGTAGCT	ATTGTAGGTA	2160
	GGCCTAATGT	AGGTAAATCT	ACAATTTTTA	ATAGAATAGT	TGGAGAACGT	GTTTCGATTG	2220
55	TGGAAGACAC	GCCAGGTGTA	ACACGAGATC	GTATTTATTC	TTCAGGTGAA	TGGTTAACAC	2280

AAATTAGAGC GCAGGCAGAA ATCGCCATAG ATGAAGCGGA TGTATTATT TTTATGGTTA 2400
 ACGTGCCTGA AGGATTGACA CAAAGCGATG AAATGGTCGC TCAAATTTTA TACAAATCTA 2460
 5 AAAAACCAGT CGTATTAGCG GTTAACAAAG TAGATAATAT GGAAATGCGT ACAGACGTGT 2520
 ATGATTTCTA TTCATTAGGA TTTGGTGAAC CGTATCCGAT ATCAGGGTCA CATGGTTTAG 2580
 GTCTTGGTGA CTTGTTAGAT GCAGTTGTTT CTCATTTTGG TGAAGAGGAA GAAGATCCTT 2640
 10 ATGATGAAGA TACAATTCGA CTATCCATTA TTGGACGACC AAACGTAGGT AAATCAAGTT 2700
 TAGTAAATGC TATTTTAGGT GAAGATCGCG TTATCGTTTC TAATGTTGCA GGGACAACGA 2760
 GAGACGCTAT TGATACAGAG TATAGTTATG ATGGACAAGA TTATGTTTTA ATCGATACTG 2820
 15 CTGGTATGCG TAAAAAGGA AAAGTATATG AATCAACTGA GAAATATTCA GTATTAAAGAG 2880
 CTTTAAAAGC GATTGAACGT TCAAATGTTG TTTTAGTGGT CATAGATGCA GAACAAGGCA 2940
 20 TCATTGAACA AGATAAACGT GTTGCCAGGAT ATGCACATGA ACAAGGTAAA GCAGTCGTGA 3000
 TTGTCGTAAA TAAATGGGAT ACTGTGGAAA AAGATAGTAA AACGATGAAG AAATTTGAAG 3060
 ATGAAGTACG TAAAGAATTC CaATTTTATG ATTATGCACA AATTGCTTTT GTGTCTGCTA 3120
 25 AAGAACGCAC AAGATTACGT ACATTATTCC CTTACATCAA TGAAGCAAGT GAAAACCATA 3180
 AAAAACCAGT TCAAAGTTCA ACTTTAAATG AAGTTGTTAC TGATGCAATT TCCATGAACC 3240
 CTACACCAAC AGACAAAGGT AGACGTTTGA ATGTCCTTTA TGCAACACAA GTTGCTATAG 3300
 30 AACCACCGAC ATTTGTTGTA TTTGTTAATG ATGTAGAATT AATGcATTTT TCTTATAAAC 3360
 GCTATTTAGA GAATCAAATC CGTGCCGCTT TTGGTTTTGA AGkTACACCA ATTCATATTA 3420
 TAGCTCGAAA GAGAAATTAA CGATTGGGGG ATAACAATGA CTAAATTAC CGTTTTTGGT 3480
 35 ATGGGAAGTT TTGGGACAGC CCTTGCCAAT GTTCTTGCAAG AAAATGGACA TGATGTTTTG 3540
 ATGTGGGGTA AAAATCAAGA TGCTGTTGAT GAATTAAATA CATGTCATAC AAATAAAAAG 3600
 40 TATTTAAAAT ACGCGAAATT A 3621

(2) INFORMATION FOR SEQ ID NO: 359:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 643 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

CTTTCGGAAA TTAGGATCnG nGCTATCTTG GCCCCAATTA CCAAGGGAAC TAnTGGCACC 60

AATGCTCTT TCATCTCCAT GCCCTGTTGC TCATTATTAA TAACACGGTC TATTAACACA 180
 ATGGCATTG TTAACGAT TCCAATTAAC ATTAGCATA CAATTAACT TGGTACTGAT 240
 5 ATTGTTTCTC CTGTGATTAA TAGTGCAATA ATTACACCGA TAACTGTAAA TGGTAAAGAG 300
 AATAAAATTG TAAATGGTGC TAGGCCACCT TTAAATGTAA TAACTAGGAT TAAATATACG 360
 ATAATGATTG CAGCTAACAT TGCAAAGGCT AATTGTGTCA TTGCATTGTT AATATCATCT 420
 10 GATGCACCAC CGATATTAAC CTTTACATTA TTCGGTTTAT CCAAATTATT TATTTTAGAC 480
 ATCACTTGTG GTGTTGTGCC ACCCACATCT TTATTTGTGA CTTTAGCAGA TACCGTCGTT 540
 15 GCATAATCTC CTTGTTCTTG CGTCAATTTA CTTGGTGTG TTGTTTTAAC TAACGTAGCG 600
 ATATCTCCCA ATTTAATCGT ACCACCAGTC GGCTTTTTCA AAG 643

(2) INFORMATION FOR SEQ ID NO: 360:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2524 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:

TTTCAAGGCG TGCAAGGGCT TTTTCTTTTG CTTTAAAGAT TATGATTAT CGTGCAAAGT 60
 30 TAAGTGGTCG TATATAGTTT TAGTTTTTAA AAGGTAATTA AATAAAATAG TTTGCCGAGG 120
 GAGATGTCAA AATGATTAAA ATACCTAGAG GGACGCAGGA TATTTTACCT GAAGATTCAA 180
 35 AGAAATGGCG TTACATTGAA AATCAATTAG ATGAATTAAT GACATTTTAT AATTATAAAG 240
 AAATAAGAAC ACCAATTTTT GAAAGTACAG ATCTTTTTGC AAGAGGTGTT GGTGATTCAA 300
 CCGATGTCGT ACAAAAAGAA ATGTATACAT TTAAAGATAA AGGCGATAGA AGTATTACAT 360
 40 TAAGACCTGA GGGAACAGCT GCAGTTGTGC GTTCATATAT TGAACATAAA ATGCAAGGTA 420
 ATCCAAACCA ACCAATTAAA CTTTATTACA ATGGACCGAT GTTTAGATAT GAACGTAAGC 480
 AAAAAGGACG CTATCGTCAA TTTAATCAAT TTGGTGTAGA AGCTATTGGT GCTGAAAATC 540
 45 CTAGCGTAGA TGCAGAAGTA TTAGCTATGG TTATGCATAT TTATCAATCA TTTGGATTAA 600
 AACATTTAAA GCTTGTTATT AATAGTGTAG GGGATATGGC GTCTCGAAAA GAATATAACG 660
 50 AAGCGTTAGT GAACACCTTT GAACCAGTAA TTCATGAATT TTGTTTCAGAT TGTCAATCAC 720
 GTTTCATAC AAATCCGATG CGAATTTTGG ATTGTAAAGT AGACCGTGAT AAAGAAGCGA 780
 TTAAGACTGC ACCTAGAATC ACTGATTCT TAAATGAGGA ATCTAAGGCA TATTATGAAC 840

	GTGGATTGGA TTATTATACA CATAACGCAT TTGAATTAAT GATGGATAAC CCTAACTATG	960
	ATGGTGcCAT TACAACGCTT TGTGGTGGTG GCCGTTATAA TGGTTTATTA GAATTGCTAG	1020
5	ATGGTCCAAG TGAAACAGGT ATTGGTTTTG CGCTAAGTAT AGAACGATTA TTGCTTGCAC	1080
	TTGAAGAAGA AGGTATCGAA TTAGATATTG AAGAAAACCT AGATTTATTC ATTGTTACAA	1140
	TGGGTGATCA AGCAGATCGA TATGCTGTGA AGCTATTAAA TCATTTGAGA CATAATGGTA	1200
10	TTAAAGCAGA TAAAGACTAT TTACAGCGTA AAATTAAAGG ACAAATGAAA CAAGCAGACC	1260
	GTTTAGGTGC CAAGTTTACA ATCGTTATTG GTGATCAAGA ATTAGAAAAT AATAAAATCG	1320
	ATGTTAAAAA TATGACAACT GGTGAATCTG AAACAATTGA ATTAGACGCA TTAGTCGAAT	1380
15	ATTTTAAGAA GTAGAGAGGG CGTTAAAATA TGAGTAAGAG AACAACTTAT TGTGGATTAG	1440
	TTACTGAGGC ATTTTtagGA CAAGAAATTA CATTAAAGG ATGGGTTAAC AATCGTCGTG	1500
20	ACCTTGGTGG ATtGATTTtC GTTGATTtAA GAGATAGAGA AGGAATTGTA CmAGTCGTGT	1560
	TTAATCCTGC ATTTTCAGAA GAGGCaTTGA AAATTGCTGA AACAGTACGT TCTGAATATG	1620
	TTGTAGAAGT TCAAGGTACA GTTACGAAGC GTGACCcTGA AACAGTTAAT CCTAAAATTA	1680
25	AAACTGGCCA AGTTGAAGTA CAAGTTACAA ATATTAAAGT GATTAAATAA TCTGAGACAC	1740
	CACCATTTTC TATAAATGAA GAAAATGTTA ACGTTGATGA AAATATTCTGA TTAAAATACC	1800
	GTTATTTAGA TTTACGTCGT CAAGAGTTAG CGCAAACATT TAAAATGAGA CATCAAATTA	1860
30	CACGTTCTAT TCGTCAATAT TTGGATGATG AAGGGTTCTT TGACATCGAA ACACCAGTAC	1920
	TAACGAAGTC AACACCTGAG GGTGCACGTG ACTATTAGT ACCATCTCGT GTTCATGATG	1980
	GTGAATTTTA TGCATTACCA CAATCACCAC AATTATTTAA GCAATTATTG ATGATTAGTG	2040
35	GATTTGACAA ATACTACCAA ATCGTAAAAT GCTTCCGTGA CGAAGATTTA CGTGCAGATC	2100
	GTCAACCTGA ATTTACACAA GTCGATATTG AAATGAGTTT TGTAGACCAA GAAGATGTGA	2160
40	TGCAAATGGG TGAAGAAATG CTTAAAAAAG TTGTAAAGA AGTTAAAGGC GTTGAAATTA	2220
	ATGGCGCTTT CCCACGCATG ACATATAAAG AAGCGATGCG TCGCTATGGT TCTGATAAAC	2280
	CAGATACACG TTTTGAAATG GAATTAATTG ACGTTTCTCA ATTAGGACGT GATATGGACT	2340
45	TTAAAGTATT TAAAGATACT GTTGAAAATG ATGGTGAAAT TAAAGCAATT GTCGCTAAAG	2400
	GTGCAGCTGa ACAATATACT CGTAAAGaTA tGGGaTGCTT TAACAGAATT TGtaAACaTC	2460
	ymTGGtGCTA AgGtTAGCGT GGGGTAAAG TTGTGGGAAG GTGGTTTTGA CAAGGTCCCA	2520
50	ATGG	2524

(2) INFORMATION FOR SEQ ID NO: 361:

55

(A) LENGTH: 1507 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

10	TCGTTGAGTA AAAGTCCAGA AAATTGGATG AGTAAACTTG ATGATGGAAA ACATTTAACT	60
	GAGATTAATA TACCGGGTTC ACATGATAGT GGCTCATTCA CTTTAAAGGA TCCAGTAAAA	120
	TCAGTTTGGG CAAAGACTCA AGATAAAGAT TACCTTACCC AAATGAAGTC GGGAGTCAGG	180
15	TTTTTTGATA TTAGAGGTAG AGCAAGTGCT GATAATATGA TTTCAGTTCA TCACGGCATG	240
	GTTTATTTGC ATCATGAATT AGGAAAATTT CTCGATGATG CTAAATATTA CTTGAGTGCT	300
	TATCCAAACG AAACAATTGT GATGTCTATG AAAAAGGACT ACGATAGCGA TTCTAAAGTT	360
20	ACGAAGACAT TTGAAGAAaT TTTTAGAGAA TATTATTATA ATAACCCGCA ATATCAGAAT	420
	CTTTTTtACA CAGGAAGTAA TGCGAATCCT ACTTTAAAAG AAACGAAAGG TAAAATTGTC	480
25	CTATTCAATA GAATGGGGGG TACGTACATA AAAAGTGGTT ATGGTGCTGA CACGTCAGGT	540
	ATTCAATGGG CAGACAATGC GACATTTGAA ACGAAAATTA ATAATGGTAG CTTAAATTTA	600
	AAAGTACAAG ATGAGTATAA AGATTACTAT GATAAAAAAG TTGAAGCTGT TAAAAATTTA	660
30	TTGGCTAAAG CTAAAACGGA TAGTAACAAA GACAATGTAT ATGTGAATTT CTTGAGTGTA	720
	GCGTCTGGAG GCAGCGCATT TAATAGTACT TATAACTATG CATCACATAT AAATCCTGAA	780
	ATTGCAAAAA CGATTAAAGC AAATGGGAAA GCTAGAACGG GTTGGCTGAT TGTTGACTAT	840
35	GCAGGATATA CGTGGCCTGG ATATGATGaT ATCGTAAGTG AAATTATAGA TAGTAATAAA	900
	TAAGGATTCA ATAATGATAT TAAGACGAGT ATGAAAATAG TTAGATTCTA ATTATTTTCA	960
40	CTACTCGTTT TTATTTTGAA AATAAGTAAT AATTCAACAA TATTATAAAT TGAACAGATT	1020
	GTTTGTGAAA TTTTGTGATAA TATTAAAGTG AAAAAGTGTT ATAAATTGAT AAATATATGT	1080
	AATTAACAAA AACAAATCAT TTTAAAAAGA AGAGAGTTGT AAGATGATGa AACGATTAAA	1140
45	CAAATTAGTG TTAGGCATTA TTTTCTGTT TTTAGTCATT AGTATCACTG CTGTTGTGG	1200
	CATAGGTAAA GAAGCGGAAG TTAAGAAAAG CTTTGAAAAA ACATTGAGTA TGTACCCTAT	1260
	TAAAAATCTA GAGGATTTAT ACGATAAGGA AGGCTATCGT GATGATCAGT TTGATAAAAA	1320
50	TGATAAAGGT ACATGGATTA TAAATTCTGA AATGGTTATT CAACCTAATA ATGAAGATAT	1380
	GGTAGCTAAA GGCATGGTTC TATATATGAA TAGAAATACC AAAACAACAA ATGGTTACTA	1440
55	mTATGTCGAT GTGACTAAGG ACGAGGATGA AGGAAAACCG CACGACAATG AAAAAAGATA	1500

(2) INFORMATION FOR SEQ ID NO: 362:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1216 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

TACATGTTTC GGATGCTACT TTATTTAGTT TGAAGGGTGC ATTATGGACG TTAGCGCAAG 60
 AAGTTTATCA AGAATGGTAT TTAGGATCGA AGTTGTATGA AGATGTTGAA AAGAAAATAG 120
 CACGAACTAC TTTAAGACA GGTATATTT ATCAAGAAAT TATTTTGAGA CCAGTAGATG 180
 AAGTTAAGGT ACTTCTGAAT GATTTAAAAG GTGCTGGTTT CGAATTAGGT ATTGCAACAG 240
 GTCGTCCTTA TACTGAGACT GTTGTGCCAT TTGAAAATTT AGGATTGTTA CCATATTTTG 300
 AAGCTGATTT TATTGCAACA GCAAGTGATG TTTTAGAAGC AGAGAATATG TATCCGCAAG 360
 CACGACCATT AGGAAAGCCG AATCCTTTTA GTTATATCGC AGCTTTATAT GGTAAATAATC 420
 GCGATAAATA TGAATCTTAT ATCAATAAGC AAGATAACAT TGTAATAAAA GATGACGTAT 480
 TTATAGTAGG CGATTCGTTA GCTGACTTAT TAAGTGCTCA AAAAATAGGT GCAACGTTTA 540
 TTGGAACATT AACAGGTTTA AAAGGTAAGG ATGCTGCAGG TGAGTTAGAA GCGCATCATG 600
 CCGACTATGT TATTAATCAT TTAGGTGAAC TTAGAGGTGT ACTAGATAAT TTGTAATTTG 660
 ATTGTTGTTT GACAGCATAA CTTGTAGTGA ATGATTGAAC CAAAGGTTTC ATATTGAGTT 720
 ACAATGAAAT TAATAATGAA AAAATGCCAA GAAGCAATGG AAGTAATCCA ATGCTTCTCT 780
 GGCATTTTGA ATTTACATAA ATTGTTTATG ACTGTACCGT CAATTCAGTT GTGAAAATTT 840
 GATTGTATTC ACCAACTTGT TTAAGTTCAT CAATTATATT GTTTGAAACA GGTGATCAA 900
 CGGATAAAAT CATTAGCGCA TCTCCGCCCG CTTCACTTCT ACCTAAAGTC ATAGATGCAA 960
 TGTTGATATT GTATTTACCT AACAAATGCGC CAGTTTTTCC TACCATACCT GGAGTATCAT 1020
 TATGATATGA CACAATTGGA TATTGATTTG GCTTTAAGTC TACAGAAAAA TTATTAATTC 1080
 TAACAATTCT TGGACCGAAA CCTGTAAAGA CAGAAGCGCC AACTTTAACG GAATCGChAT 1140
 CGnTTGATAG TTCTACCTCT AAGTAGTTAC TAAACCTGT CTCTGCTTTA TTATTTTCAA 1200
 TATTTAATGT CACTTG 1216

(2) INFORMATION FOR SEQ ID NO: 363:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 867 base pairs

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

ATCAAAATAT TTAAATAAT ATTGATGGTC ACATTGTAAA TTTAATAGAA AATAAATTG 60
 10 ATCAAAATATT ACAAGAACCA TTAAATCCAT TAAATTATGA TACTGTCACT GGATTAGCTG 120
 GGATAGGGAG ATATTTGCTA AATAGAGTAG ATGAGAATGA ATTTAATGTT AAAGCATTAA 180
 AAAGCATATT AGTATACTTT AAAGATATTC AATATTCTAA AAATAGCTGG GTAGTCCCAC 240
 15 AAGAAAGTCA ATTTTGTAGAG TCTGATAAAA ATTATTTTAC TGAAGGTAAT ATCAATCTTG 300
 GCCTTGACCA TGGaGTGCTA GGACCGATGT CTTTATTTGC ATTATGCGTG ATTAAAGGAA 360
 TTACGATTGA AAATCATCAG CACATATTAA AAGACATGTA CAAATTTATC ATGGACGAAA 420
 20 AATTTTGTA AACCAGAAAG TGGTTGCAGC GTTACGATTT AATTTCTGAA CGTAATCATT 480
 TCAATTTTAT TCGGAATGGT TGGTGTATG GCAATACGGG TGTAATGACG ACGTTGTTTT 540
 25 TAATCGGCCA AGCATTACAA GATGATGAAA TAATTAAAAT GTCTAAAAAA GTGATGCTAC 600
 AAGTAGTAAA TGATAAAGAT GAAAATTTAA TAAGTCCAAC TATTTGTCAT GGATTGTCAT 660
 CACAAATATT AATGTTAACA ATTATGAATT TGAATTTTGA ATTAAATGAA GTGTCTGATT 720
 30 ATATCACTGT ATTAATAAAT AAATGATTT CTCATTATAA GGAAGATTAT CTGGTGAATT 780
 TTATAGACAT TAATGAAAAT AAGCAAGATG TATTTAAAAG TAGGAAAGTT GGCCTTTTAG 840
 AAGGTGAATT AGGGGTCATC TAACATT 867

35

(2) INFORMATION FOR SEQ ID NO: 364:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10813 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

45

TACCTTTTCT TTTAAATCAT TTTATATTTT CCCACTAATA TCCGcTGtTA ATCaATCctG 60
 ACATCctTGT ATCaCTATGA CAATTAATTG TTAAATACAT GAATTTCTAC ATTTTATGAA 120
 50 AAAATCCATT TTTATTACAA TTCAACACTT TATATGACAA CTTCAATTACA GTTACTTTTA 180
 TTGTGATTG CTTACATTGT TTTCTAAAAA AAATTTGTGA TCATAATTAA CGTTGAATAA 240
 AGAAAAAAT TAACTTGGGA GATAAAAATG GAATATAAAA AGATACTAAT TCGTTTATTA 300

55

	CAATCGCATG CCGCAGTTAA TTATTATAGT AAAACCAAT GTACATGGTG GGCATTTAAA	420
	CGTCGCGCAC AAGTCGGTAA ACCTGTTTCT AATAGATGGG GCAATGCTAA AAATTGGTAT	480
5	TACAATGCAC GTAAATCAAA ATATCGGACT GGTCTGACAC CAAGAAAATT TGCTGTCATG	540
	CAATCAACTG CAGGATATTA TGGACATGTC GCAGTTGTTG AACAGTATA TAAAAACGGT	600
	AGTATTAAAG TTTCAGAATA CAACTTTTAT CGCCCATTA AATACAATAC ACGTGTTACTA	660
10	AGCAAAAAGG CAGCACGTAA CTTTAACTAT ATTTACTAAT CAAAAAACTT CTATCACGAA	720
	CGCTTCAATT TCCTGTATGC GTGTGATAGA AGTTTTTATT TTATGAAATT ATATTATTAC	780
15	TTCTACAAAT TTCAAATTGC CGTAATTGAA CGTATATTTT TTCTTCAACT ATTATTTTAT	840
	CTTTAGCATA ATCTATATAT AAAATTTTAT GCTATTATTT AAATAATTCTG CTATAACTTA	900
	ACATACGTTT TCGATATAAA CCTTGTCTA AATCTCAATA ATTTTTTGCT GTTTTCATCG	960
20	TCATTAGTTA AAAAAATAAT TTAAGTGTG TTTTGAGACC TGTTTAAATT GATGACTCAA	1020
	AACCTTTATC CCCTTTTCAC TCGGTTTAAT TGCTTTAATA TTTAACACAG TCTCATTGTA	1080
	ATTTTGCTAC TAAGTTTGAA ATATTTGAT TCGAATGTGA ACATCATTTA TTATTACTTT	1140
25	CGACAGCTAG AAAAATATTG TTAAACCAAA ACTATAATTA CCACTTTAT ATAGAACTAT	1200
	ATATAAATT TAACTTTGAG ACAGAACTGC TAGTCAGGTT TATGAATATA TTTCTTTAGT	1260
	TTACTTGaTA TACTTATTGG TAAATCATTA TTTATTGAA GAGCATTTAT AACAAAAAGT	1320
30	ATAATCCCAG TTATAGCGAT TCCATGACCA TTTTATCAG TCAAAGATTC ATCAACAAAA	1380
	ACTTTTTGAT AAATTTATAT ATTTGTATAA TTTATTATGG TAGATATCTA CACCCGTATC	1440
35	TATAACAGCT ACAGTTATCA TTGCGTCTTA TCCAAATAGT TTTAAGAAA TAAATATAAT	1500
	TCATTAATAT CATGGCTATT TGTAACGTCA GtGCTTaAGA CCTGTCAAT TATGTTTTTA	1560
	TCTGTTTCTA AGAGCCCTG ATTTTCTGCA TACATATTTT GTATATCGCT GTTAATCTCA	1620
40	TTTAATTTTT TAAATTTTT TCTAACCTTT TTCAAAATAA AATGCTCTTT AATTCCATCT	1680
	CTTAAATACA TATAAATCTC TATTATAATC TCCATAAATG TCATATCTCC TCTCTCAAAA	1740
	AACATTATAC TCAATATATT AAAAATAAGA AAGGTTTTTT CAGAATTTTT AGTATTTTCA	1800
45	GATATTTATT CTTAAGTATC TTATTATTAC TTCTTTTTAA CGCTAGTGGa ATAATTAATA	1860
	ATAATTGTAT TTAAATGCAA TGCTGTTGTT CTTTCATACT TACAAGCAAG TCATAAGAAA	1920
	TGAGAAATTA AATTCTTATA TGATGAATAA ATGATGACTG CAGAAGAAAT TGCTAAAAAA	1980
50	GGAGTGGAGT GAGGCATTCT ACTGTTTATA TTGTTAACAA ATATCAAAAT GATAGGAACT	2040
	TGAAGAATCT AAGACTGTTA AAATCGAGAC AAGAAGAAAC TCGTTCCAAA TTTGAATCCT	2100

55

	AAACTTCTA TCACGAACGC TTCAATTTCC TGTATGCGTG TGATAGAAGT TTTTATTTTT	2220
5	GTTTCATATTA ATTTATCTAA GCGCTACGAT GGAaCTGACTT AAACTTTTCT TTAAAATCGC	2280
	TATTGCGCCAT TTACTATTGT TGTCTAATTT CTTGTAAAAT ATGTTCCGCT GCTTGTGTAT	2340
	TTGCACGGGG TTCTTTTTTC AAAGCTTCAG CTACTTTAGC AATTTTCATCA CCTTTTGCCC	2400
10	CTACAACGAT AGCTAATGAT TTATATTGTA AACTCATATG ACCTTGTTGA ATACCTTCTG	2460
	ACACAAGCGC GCGACATGCT GCAAAGTTTT GCGCTAAACC AACGGCAGCA ACTACATGAC	2520
	CTAATTCTTG TGCTGACTCT ACATTTAGTA GCTCTAATGA AGCTTTAGCA ATTGGTAATA	2580
15	CTTTTGTAAC ACCGCCAACG ATTGCCAATG TCATAGGCAC TTCAATTGTA CCAATCAATC	2640
	GTTGACGATC TTGATCGTAA CGCCATGTAG CAATACCACG ATACTGTCCG TCACGACTCG	2700
	CGTATGCATG CGCACTTGCT TCTGCACCAC GCGTATCATT TCCTGTTGCT AAAACAACAG	2760
20	CATGTATGCC ATTCATAACA CCTTTATTAT GTGTTGCTGC ACGATGAATA TCTACTTGGG	2820
	CCAATACAGA AGCAGGTTCC ATTCGTTTGG CAACCTCTTC TCCAGTTCTC TCGCCCCTTG	2880
	CTAAATCTTT AACATCAATT TCGCCTTGAA CTTTAAACAAC GGACGCTGTT GCATGATTGG	2940
25	ATAAAATACT CATTAAAATG TCGCTTTGCG GAAATTCATT TTTTAAAAAT GCAGTTATGG	3000
	CCTCTAAAAT CGTATTAAAGC ATATTAGCGC CCATAGCATC TTTCGTATCA ACAAATACTT	3060
30	TTAAAGATAG TAACTGTTGC TCAGGAAATG TATCAATCGC TATACGTTGG TAACCACCAC	3120
	CACGCGCTTT AATAGAAGGA TATGCCTCAT CCGCAATTTT ATGAATTTGC TTTTCTAAAG	3180
	CTTTAATGTC TGCTGATAAT TTTTCAGTAT CGTCAACGCC ATCAAAGACG ATTTGACCTA	3240
35	TCATAATACG TTCAGAAGAT ACCGTTTTAA ATCCGCCAGT CTGATTCACT AGCTTTGCAC	3300
	CATAACTAGC TGCAGCGACA ACTGAAGGCT CTTCCACCAT CATAGGTACA ACATATGCCT	3360
	TATCGTCCAC AATGATATTC GGTAATAATC CAACGGGTAA TGCACCTTGC GCGATGACAT	3420
40	TTTCAATTAA ACTATTGGCT ACTTCTTCAT CGATTAATGG ATGATTCAGT AAAATGTCGA	3480
	ATTGTTCTTC TGATAACCAT TGCTTATCAA CCAATTGTTG TAACTTTTCT TTACGAGATA	3540
	AATGTCGAAA ATTCTTATCT AAATTTTGCA TGGACGTACT CCTTTTACTT CACATAATTT	3600
45	TTAACATTTT AATCACTACT ATTTTACCA CAAAATAACG TCATTTCGTCT TAAAATTCAA	3660
	TTGAATAATT GTCGTTTTGA CTTTAAAATA AAACAAGGTA AATTAAAACG CTTACAAGAA	3720
50	ACGACAAATc ATTTTAAAT TTAGTATATT TCTTTGTATA AAATTAGCAT ATTCTGATAT	3780
	GATACAAGTG TTGCTTTTAT AAATTTGAAA GGATGTAAAA CCTTATGACA ATAGGTATCG	3840
	ATAAAATAAA CTTTACGTT CCAAAGTACT ATGTAGACAT GGCTAAATTA GCAGAAGCAC	3900

55

	CTGTAAACCA AGACATCGTT TCAATGGGCG CTAACGCTGC TAAGGACATT ATAACAGACG	4020
	AAGACAAAAA GAAAATTGGT ATGGTAATTG TGGCAACTGA ATCAGCAGTT GATGCTGCTA	4080
5	AAGCAGCCGC TGTTCAAATT CACAACCTAT TAGGTATTCA ACCTTTTGCA CGCTGCTTTG	4140
	AAATGAAAGA AGCTTGTTAT GCTGCAACAC CAGCAATTCA ATTAGCTAAA GATTATTTAG	4200
10	CAACTAGACC GAATGAAAAA GTATTAGTTA TTGCTACAGA TACAGCACGT TATGGATTGA	4260
	ATTCAGGCGG CGAGCCAACA CAAGGTGcTG GCGCAGTTGC GATGGTTATT GCACATAATC	4320
	CAAGCATTIT GGCATTAAAT GAAGATGCTG TTGCTTACAC TGAAGACGTT TATGATTTCT	4380
15	GGCGTCCAAC TGGACATAAA TATCCATTAG TTGATGGTGC ATTATCTAAA GATGCTTATA	4440
	TCCGCTCATT CCAACAAAGC TGAATGAAT ACGCAAAACG TCAAGGTAAG TCGCTAGCTG	4500
	ACTTCGCATC TCTATGCTTC CATGTTCCAT TTACAAAAAT GGGTAAAAAG GCATTAGAGT	4560
20	CAATCATTGA TAACGCTGAT GAAACAATC AAGAGCGTTT ACGTTCAGGA TATGAAGATG	4620
	CTGTAGATTA TAACCGTTAT GTCGGTAATA TTTATACTGG ATCATTATAT TTAAGCCTAA	4680
	TATCATTACT TGAAAATCGA GATTTACAAG CTGGTGAAAC AATCGGTTTA TTCAGTTATG	4740
25	GCTCAGGTTT AGTTGGTGAA TTTTATAGTG CGACATTAGT TGAAGGCTAC AAAGATCATT	4800
	TAGATCAAGC TGCACATAAA GCATTATTAA ATAACCGTAC TGAAGTATCT GTTGATGCAT	4860
30	ATGAAACATT CTTCAAACGT TTTGATGACG TTGAATTTGA CGAAGAACAA GATGCTGTTC	4920
	ATGAAGATCG TCAATTTTTT TACTTATCAA ATATTGAAAA TAACGTTTCG GAATATCACA	4980
	GACCAGAGTA GTCGGTGTAT TTAAACACA TATAATAAAA CCTAAAAGCA GCAGTAAGAC	5040
35	CACTTCTAAT TGAAATCGTC TTACTGCTGT TCTCTATTTA TAACACTTCG TATTGAATGA	5100
	ATTCATTATG CCTATTTGAC ACATTATTGA AGTTTTCTTA ATGCCTGGAT CCTTTATACG	5160
	TTACGGCTTC GTGCTATGTT TTGGTACATA AAGCTTTGAC ATATCGATAT TCTCCAATC	5220
40	TAACAGCTTA ATTTTATTAT TAATCGTTCC ACCGAACCCT GTTAAGCTAC CCGTTTTACC	5280
	GACAACACGA TGACATGGCA CGATAATAGA TAATGGATTA CTTCCGACTG CACCTCCAAC	5340
	CGCTTGGGCT GACATTTTTG GCTTGTTAAG CAGCTTGCCT ACTTTTTTGG CAATAGCACC	5400
45	ATACGTTGTT AGAGTCCCAT AAGGAACCTG TCTTAATTCA TTCCAAACAC ACTGTTGAAA	5460
	ATGACTACCT GTTGGCTTTA AAGGTATTGT GATTTTCAGGA TTGTCACCTT TAAAATACGC	5520
	GTCTAACCAC TGTGTCGCCT CTCTAAATAT CGCTAAAGAC GTATTTTCTT CCCTAGTACC	5580
50	ATCACCTTGT TGATTTTCAA ACAAACAGC GGTCAACTT ACCCCATCAC TCAAAAGTTC	5640
	CAATCGTCCT ACAGGCGAAT CATAGTAACT CTTTACTTCC ATAAAAATTC CCCCTTTTTC	5700

55

	ATAAGTCGTC AATTACGTAT ATAAACACGT AATACCAGCT ATCACTTTGC TGCAATATAC	5820
	AGTTACATAT CTTACTACAC GTGCTAACCT CTTACTTTGT AAACCAAATC TTAAATTAAA	5880
5	ATATTGAAAA TGCAATGAAT CCTTAATATT TTATTAAACC TATAATTACT TATTAAAAAT	5940
	AACACACAAT ATTCATAAAG TTTTAAAAAT ATTCTGTTTT ATCACCTACT ATTAGTGGA	6000
10	AAGTACAATT GCAATTGTAT ATAGTTTGCA TAACGCTTCA AAAGTAATTT CTTTTTTGTT	6060
	TAGTTCAAAA AAATTTAGAG GTGATGTTAT ATGAATAACG GTTTTTTCAA TAGCGACTTT	6120
	GATTCAATTT TTCGAAGAAT GATGAAAGAT ATGCAAGGTT CAAATCAAGT CGGAAACAAA	6180
15	AAGTACTATA TTAATGGTAA AGAAGTTTCA CCTGAAGAAC TAGCGCAACT CACACAACAA	6240
	GGTGGCAATC ACTCTGCTGA ACAAAGTGCG CAAGcTTTT CACAAGCAGC ACAAAGACAA	6300
	CAAGGGCAAC AAGGTGGCAA CGGCAATTAT TTAGAACAAA TTGGTCGTAA CCTTACGCAA	6360
20	GAAGCACGTG ACGGTTTATT AGATCCAGTC ATTGGTCGTG ATAAAGAAAT TCAAGAACT	6420
	GCTGAAGTTT TAAGTAGACG AACTAAAAAC AATCCTATAT TAGTTGGAGA AGCTGGTGTT	6480
	GGTAAACTG CGATTGTTGA AGGTTTAGCA CAGGCAATCG TTGAAGGAAA TGTACCAGCA	6540
25	GCAATCAAAG ACAAAGAAAT TATTTCTGTA GACATTTTCAT CATTAGAAGC TGAACGCAA	6600
	TATCGTGGTG CTTTTGAAGA AAATATTCAA AAATTAATCG AAGGTGTTAA ATCTTCACAA	6660
	AATGCCGTAC TATTCCTTGA TGAAATCCAT CAAATTATCG GTTCAGGTGC CACAGGAAGT	6720
30	GATTCAAGTA GCAAAGGGTT ATCTGATATT TTGAAACCTG CATTAGTCTG TGGTGAGATT	6780
	TCTATTATTG GTGCAACAAC ACAAGATGAA TATCGAAACA ATATTCTTAA AGATGCTGCA	6840
35	TTAACGCGCA GATTTAATGA AGTGCTTGTT AATGAACCAA GCGCTAAAGA TACTGTTGAA	6900
	ATTTTAAAAG GTATTCGCGA AAAATTCGAA GAACACCATC AAGTAAAATT ACCAGATGAC	6960
	GTATTAAAAG CATGTGTTGA CTTATCAATT CAATATATTC CACAACGATT ATTACCAGAT	7020
40	AAAGCAATCG ATGTGTTAGA TATTACAGCA GCACATTTAT CTGCGCAAAG TCCAGCTGTC	7080
	GATAAAGTTG AAACTGAAAA ACGAATTTCT GAATTAGAAA ATGATAAACG TAAAGCAGTA	7140
	AGTGCTGAAG AATATAAAAA AGCTGACGAC ATTCAAAATG AAATCAAATC ATTACAAGAT	7200
45	AAATTAGAAA ATAGTAATGG TGAACATACT GCTGTTGCTA CAGTTCATGA TATTTAGAT	7260
	ACTATTCAAC GATTAAGTGG TATTCCAGTT TCTCAAATGG ATGATAACGA TATTGAACGT	7320
	TTAAAAATA TTTCTAATCG TTAAAGAAGT AAAATCATAG GTCAAGATCA AGCTGTAGAA	7380
50	ATGGTTTCAC GTGCAATTCG CCGTAATCGT GCTGGGTTTG ATGACGGCAA CCGTCCAATT	7440
	GGCAGTTTCC TATTTGTTGG CCCTACTGGT GTTGGTAAAA CAGAGCTTGC TAAACAATTA	7500

	GACACAACAG CTGTTTCAAA AATGATTGGT ACAACTGCTG GTTATGTTGG TTATGATGAC	7620
	AATTCAAATA CGTTAACTGA AAAAGTACGC CGTAATCCAT ACTCAGTCAT TCTATTTGAT	7680
5	GAAATCGAAA AAGCAAATCC ACAAATTTTA ACATTGTTAT TACAAGTAAT GGATGATGGT	7740
	AATTTGACTG ATGGTCAAGG TAATGTCATC AACTTTAAAA ATACAATTAT TATTTGTACA	7800
10	TCAAATGCTG GCTTTGGCAA TGGCAATGAC GCTGAAGAAA AAGATATTAT GCACGAAATG	7860
	AAAAAATTCT TCCGCCCTGA ATTCCTTAAC CGCTTCAACG GCATCGTTGA ATTCTTACAT	7920
	TTAGATAAAG ATGCATTGCA AGATATCGTC AACTTATTAT TAGACGATGT ACAAGTTACA	7980
15	TTAGACAAAA AAGGTATTAC GATGGACGTT TCTCAAGATG CGAAAGATTG GTTAATTGAA	8040
	GAAGGCTATG ATGAAGAATT AGGTGCACGT CCATTAAAGAC GTATTGTTGA ACAGCAAGTA	8100
	CGTGACAAAA TTACAGATTA CTATTTAGAT CATAACAGACG TTAAACATGT GGATATAGAT	8160
20	GTTGAGGATA ACGAATTAGT CGTAAAAGGT AAATAACGAC ACTTTAACAT ATCGCGCATC	8220
	AAAAATGAGC ATCAGGTGCG CCTTGCCTGT GCTCATTTTT TTAATTATTT CCCTGGAAAA	8280
	TGATTGCTG TGTGCTGTTT TGTTCACAA CAATCACGAT TAATGTCACA TGTACCACAT	8340
25	TTTCCTTGTT TTGAACGCTT GAAAAATTTT ACTAGTGTAT ATAAGGCATA TCCGAAAATT	8400
	GCTAAAAAAA TTAAATGTT AATAATGACT GACACTTTAA CCACTCCTTA AACAAATAAA	8460
	TGTCCGACTT GATAAAAAAT GAATGTTAAG ACATATGCAG TGAAGTAGAGG ATAGGCAACT	8520
30	GCAAGTGCCG TCCATTTCCA TGAATAAGTC TCTTTACGGA TTGCTGCTAC TGTAGAAACA	8580
	CAAGGAATAT ACAATAGTAT AAATATCATA AATGCATACG CAGATAGCGG TGTGAATTGA	8640
35	TTTTGAATCA CATTACAAG GCCTGCATCA CCTGATGAAT AGATAATCGC CATCGAACTT	8700
	ACGATAACTT CTTTTGCTAA AAATCCTGGC ACTAACGTAG CACCTGCTTG CCATGTTCCA	8760
	AATCCGAGCG GTTGCACTAA CATACCAAAG AAATACCAA CCATATGTAA AAACTTTGA	8820
40	TTGATATTCA CATTGATACC ATGTGGTCCT ACATAACTTA ATAGCCAAAT GACTACTGAG	8880
	CCGCCAAAAA TAAATGTACC TGCTTTACGA ACAAAGCCCT TAGCCTTTTC CCAAGTACTA	8940
	CGCCACAACG TTTTAATGGA AGGCACACGG TATGTTGGCA ATTCCACAAT AAAGATTGCA	9000
45	TTATCATTTT TTAAATCGT CTTAGTAAGT ACTGTACTGA CTAAAAATGC CATAATAATA	9060
	CCTAAAACAT ACAGGCTTAA TACTACTAAA GATTGATTCT CTTTGAAAAA GATACCTACG	9120
	AACAACGCAT AACTGGCAG TCTAGCAGAG CATGACATGA ATGGTGCAAT TAATATCGTT	9180
50	GTAAACGCT CTTTTTCATT TTCAATACTG CGCGCAGCCA TAATACTCGG TACATTACAA	9240
	CCAAATCCGA TAATCATTGG TATAAAGAC TTCCCGCTTA AACCGAACGA TTCCATAATA	9300
55		

	AAAAAGAGCA CAACAATTTG TGGTACAAAG ACTAATACTG ATCCTACACC AGCAATAATG	9420
	CCATCTGTAA TTAAATCTTG TAAAAATGGT ATAACACCAA GATAATTCAT AATCGTCTTC	9480
5	ACACTATCTG TAAATGTACC ACCTATAAAT GCATCGAGTT GATCCGACAA AGGTGTGCCA	9540
	ATCCATGTAA ATGTAGTTTG AAAGATCAAC CACATAATTG CTAGAAAGAT AGGCATCCCT	9600
	ATATATTTAT GTGTTAATAT CTTGTCTATT CTAGAGCTGA AATATTGCTT ATCTTCATCT	9660
10	GGATACGTTA CCACGTCTTG CAATAACGTC TCAATATAAT GATTGCGTAT ACGCTCCATC	9720
	TCTCGACGAA CAGATACAGC CCCTACTTGT TCAGCAACTT GATCACGTAA ACTCGACAAT	9780
	TTATTTACAA CCTCTGAATT AAGTTCGTTT GCAATTTGCA TGTTATTTAA TAAGAATTGA	9840
15	ATCGCAATAA ACCTAGCTTG ATACTTATCA TGAGATGTCT CTGTCAATTAT TATTTGACAC	9900
	ATATTTTTAA TTGTCTCTTC AATCTTCTCA CCATAATTGA TTTTAAATG CGGTTGATAC	9960
20	CCITCCCCTA GATGCTTTAT TTCGCCAAGT AAATATTTTG TTCCTTTGCC TGTACGTGCC	10020
	ACAACTGGAA AAATAGGTGT TTTTAACTTT TTCATCAATT TATGATAATC GATTTTTATC	10080
	CCGCGCTTTG TAGCTACATC AATCATATTT AATCCGATGT ATATTGGTTG ATTAAGTTCT	10140
25	AACAATTGTA CTGTTAATTG CATATTTCTT TTTAGTTGAC TCGCATCAAC AATGTTAATG	10200
	ATTCTGAAA ATGAATCGTT TAATAAATAG TCTGTCACTA CAGTTTCATC TTTAGAAATC	10260
	GGCGATAAAT CATATGTACC TGGTAAATCA ATTAATTGTC CTACATTTTC TTTAAGTTTC	10320
30	CCTACTTTTT TCTCTACCGT TACGCCACTC CAGTTGCCTA TATATTCATA CGAACCAGTT	10380
	AAAGCGTTAA ACAAAGATGT TTTACCAACA TTAGGATTTT CTAAATACA ATAATTTTCC	10440
	ATTCGTCCGG CTCCTATTCT TCTAATGCAA TAGAACAAGC ATCGCAATGT CTAATACCTA	10500
35	ACTGTTGTCC GTTTACTTCA ATAATACATG GCCCTTTAAA TAAACATTTT TGTTTAATCG	10560
	TTATGATAGC GTCATCTGTT AACCCTAAGG CACTTAGACG ATACAACATA TTCTCATTAG	10620
40	CAATATCCAT TCGCTTTATT TTATAAGCCT TATTCATTTT ACCATTTTTA ATGTTTAACA	10680
	TACTATTTTG CTCTCCTATT AGAAATAATA ATCATTATCA CTTAAAAATC ATAACCCTTA	10740
	AAATGTAGC TCGCAATACT TTATTTAAAT AATTTTCATT TTTTCATGTAA AATTTGTGAC	10800
45	ATTGCAAAAA TGT	10813

(2) INFORMATION FOR SEQ ID NO: 365:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6804 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

	TGATGGATTA GCAGACATTT TACGAGCGAA TGGTTTCAAA GTGTTTGGTC CAAATAAGCA	60
5	AGCAGCTCAA ATCGAAGGCT CAAAATTATT TGCTAAAAAG ATAATGGAAA AATATAATAT	120
	TCCAACCTGCT GATTATAAAG AAGTTGAGCG AAAAAAGGAT GCTTTAACAT ATATTGAAAA	180
	CTGTGAATTG CCCGTTGTTG TCAAGAAAGA TGGGTTAGCT GCTGGGAAAG GCGTTATTAT	240
10	TGCAGATACT ATTGAAGCAG CCAGAAGTGC TATTGAGATT ATGTATGGTG ATGAAGAAGA	300
	AGGTACTGTT GTATTTGAAA CGTTTTTAGA AGGTGAAGAG TTCTCGCTAA TGACATTTGT	360
	TAATGGTGAT TTAGCAGTAC CTTTCGACTG TATTGCACAA GATCATAAAC GCGCATTTGA	420
15	TCATGATGAA GGACCAAATA CTGGTGGTAT GGGGGCTTAT TGTCCmgTAC CACATAATTAG	480
	TGACGATGTT TTAaMACTTA CAAATGAAAC AATTGCACAW CCCATTGCAA AGGCAATGCT	540
20	TAATGAAGGT TATCAATTCT TCGGTGTATT ATACATTGGT GCTATTTTAA CTAAAGATGG	600
	TCCAAAAGTA ATAGAATTTA ATGCCCGTTT TGGTGATCCT GAAGCTCAAG TATTATTAAG	660
	TCGCATGGAA AGTGATTTAA TGCAGCATAT TATTGATTTA GATGAAGGAA AACGTACTGA	720
25	ATTCAAATGG AAAAATGAAT CTATTGTAGG GGTCAATGTTG GCATCAAAAG GATATCCTGA	780
	TGCATATGAA AAAGGGCATA AAGTAAGTGG CTTTGATTTA AATGAAAAC TTTTGTGTTAG	840
	TGGATTAAAG AAGCAAGGTG ATACCTTTGT TACTTCAGGT GGTAGAGTTA TACTTGCCAT	900
30	CGGAAAAGGT GACAAATGTAC AAGATGCACA GCGAGACGCA TACAAAAAG TATCACAAAT	960
	ACAAAGTGAC CATTTATTCT ATCGTCATGA CATTGCGAAT AAAGCACTAC AACTTAAATA	1020
	AGTAAATTTA AAATACTAAG aTTAGCTATG AACGAATCTA TAACGATAGA TTTTTTCATA	1080
35	GCTTTTTTAG TTGTAGAGTC TAGGACATTG ATTTCTGTAC CAAATTTGTG ATTATGCATA	1140
	TGTAATACAA AAGAGGCGCC ACAACATGTT TGGATGAACA AAATAACATG TTTGTGGCAC	1200
40	CTCTTTTGTT TAGTATGGAA TAAATGGTTT TCTTTTTCTA TACAATGAAT TTCTAATTTA	1260
	GTATCTATAC AATTATGGAT AAAATTTAAC CTACACGACC AAGACGAACA TCATCTATGC	1320
	CCGTGATGGG TAAGGTGATT GAACAATAAT ATGCCATAGT AATAATGGCA ATTAAACTA	1380
45	TAATAAAGAT TATATCTTTA TATGAGAAAG GTACGTTGTA ATAGTAAGTA CGAGGACCAT	1440
	CTCTAAATCC TTTCGACTCC ATCGCAACTG ATAATTGATG TGCCTTTCTA ATATTTTGGC	1500
	TTAATAGAGG TATAATTAAA TGCTTAAATC GCTTTAACCC TCTATAATTT GCCGCGTCTA	1560
50	TCATCTGATA GCGCATTTTT AAAGATCTGC GAAGcTGTA TAAAGAACTA ATCATTAAAG	1620
	GTATCATACG AATGGCAGCC ATGAATGCAT AAGCAACTTT TGATTTAACC TTTAAATGTT	1680

55

	ATGAAATAGC AATGGTTCTT AATGATACAT GTAAACCACG AACTAAACTT TCTGTTGTAA	1800
	TATGGATAAA TCCGAATTC AAAATTGTAT GGCTACCATT CCCGTATAAA ATCATGAACA	1860
5	GGGAAGAGAG TAATGCAAAG CCAATACTTA TAGTTATAAA AATTGCTGTA ATTTTAAACT	1920
	GAGTACCATT AAACATCAAT AAGAAAATA ACATTAGAT AGTGATATA AGCATAAAAT	1980
	CGAAATTATG CACAAATATA ATAAAGAAAA ATAGTATAAT TCCAAGAAAT AGTTTCGTTA	2040
10	TAATGTTGAC ATCATCAACA AATGATTGCC GAACTTTCCA TTGCTCATAC ATTCGTATCA	2100
	CCATCACAAT CTAGTAACGC ACCATCTGAA ATTTTAAGTC TTCTTGATGG ATAACGTTCA	2160
	ATTATTTTCAT CGTCATGTGT AaCCATGACA ATACTTTGTC CCAAATTAAT TCGCTTTTGG	2220
15	AAAAGTTTGA TCAACTGGAA TGTATTATGG CTATCAAGTC CAAATGTCGG TTCATCTAAA	2280
	AAGATAATAT CAGCTTTAGA ACTTAGTGCG GTAGcTACGC TAAGGCGTCG TTTTGTACCA	2340
20	ATAGACAACT CATAAGGATG TTGATCTTTT ACATTTTGTA AATCTAAAAG TTTTAAAAGT	2400
	TGTATCGTTT CATCATCACT TTGATCTTTA GAAAGGTGAT TAAATGAAT GTTAATTTCA	2460
	TCATAAACCG AATTTGTTAT AAATTGTAAT TCTGGGTTTT GATAAACTAG GTACATGTGT	2520
25	TTTGCTGCAT GTTTAATTTT TGTTAAACGC TGATTTTCAA AATAAACATC ACCTTGATAT	2580
	TTAATCAATT GCATAATTGA TTCAAGCAAG GTTGTMTTAC CACTACCATT TGCCCCTGTA	2640
	ATTGTAATCC ACTCACCTAG ACCAATTTCT AAATCTGAGA ATGAGAGCAA TGTTGATTTA	2700
30	CCGCGAATAA TACGTCCATT TTTAAATTGT AATAAGTGTG AGTTTGTGTG TGGAAAGTCA	2760
	ACACGACTTG GTGCGAATTC CCATGCACGT GGATGCCACA CACCATATTC ACTGAGTAAA	2820
	TGAACATACT TCTGTAATAT GATTCAGGA CATTATCGG CAATGATATT TCCGTTATAA	2880
35	TCCATCAAAA TGACGCGGTC GACATGATTC CAGATGTGTT TAACTTTATG TTCAACGATT	2940
	ACAACCGTTT GATCTTCCCA AAGTTCAATT AGTTTAGTCC ATAAATCTTC TGTGCTTGA	3000
	ACATCTAACA TTGCTGTCGG TTCATCTAAA AACAATGTTT TTGATTGTTG AAGAATGGTT	3060
40	TCAACAATTG CCAATTTCTG TTTTATCCCG CCACTTAAAT CTTTGATATA CGTTTCAGGG	3120
	GTAACATTTA AATTGACCAT ATTTAAAGCA TTGATAATTA ACGCATCCAT GTCTTCACGT	3180
45	GGTAATTGTC TATTTTCTAA AACGAATGCA AGTTCTTCGT ATACTTTTGG CATACAAAAC	3240
	TGGCTATCAG GGTCTTGGAA AATAACGCCA CTTAATGGGT CAACGATTAG TTCATCATAT	3300
	TTCATAGGTA ATTCAATTAA ATTAGGAACA ATACCACTTA ATACATTCAG AAGTGTAATT	3360
50	TTACCGCAAC CAGAAGGACC GAGTAAAAGT ACTTTTTCTT TGTCTTGAAT AGTGATATTT	3420
	AAATGATCGA AAATTTTACG TTGACCACTT GGATATTTTA ATCGTAAATC ACTTACTTTT	3480

55

ATTTTGTTAC GCCTGTCTTA TCTAAAGCTT TTAATAAAG GTAAGATAGG ACGCCGCGGA 3600
 CTAAGTGCACC ACTAATTAAT CTAAATACGA TGAATAATGT TAAGTTCCAA CCTGCAACTT 3660
 5 CATTTAAATA ACCATAGAAA TAATCTATCG GGAAAGCCGC GATTGCTGTA CAAAAACCTG 3720
 CTAACATAGC TACCATAACT GAACGTGATT GATATTTAAA AATTGCAAAG ACAAGTTCAC 3780
 ACGCTAAACC TTGTATAAAA GCGTAAACGA TTGTGCGAAT ATCGAAACGA CCCATAATGA 3840
 10 TAGTTTCGCC GGCACCTGCA GCAAATTCAG CCAGTAAAGC AATACCTGGT TTTGGAATAA 3900
 TTAGATAGCA GACAATCGCT GCCATGAACC AAACCCCGTT TGTTAATTGT TCGAGGTGAA 3960
 GGCCTGTAGC TTGCACACCA TTGTAAACAA ACCACCATAA ATTGTAAATA ACTGCGAATA 4020
 15 CTAAGTGAAT AAGTACGGTT ACTAGTATTT CAGATAGCTT TAAACCTTTT GACATTTTTA 4080
 CATCTCCTA ATAAAAAAC GCACAACCAT CCATAGGAAA GTTATGCGTT CACAATATAT 4140
 ATTAGTAAAA CATATGTATA GTAACACTTT CCTACGCTAG TTCAAGCTAG ATCAGGTTCA 4200
 AAGGGTTTGA GGGCAAGCCT CATCTCAGTA TAAACACCC CTAGTGTGTG CGATTTATTT 4260
 AATTAATTAT ACTGTAAGAC GTTTGTAAAC TTATGTCAAT AGGTTGTCTT CATGAAATTT 4320
 25 CGTTTAATTC GATTTAAAT TTATAATATT AGCATTGGAT TTAAATTGAA GATGTAGTAG 4380
 GAATGTTAGT AATTAAAGAT ATAAAAATAT GTGACATGTA ATAATATTGA GCTGATAAAT 4440
 GAAGAGGGAT ACTTATCAAT CATACCTCTT TAACAACAGT GAAGAACCCG TGCATAATGG 4500
 30 CTTACGAATT ATAGTTTATA AGGAAGAAGA GGGATACATG CGCCGAGCAC ATGCATAAAA 4560
 GCCCCTAACA ACTAAAGTT GTAAGGAAGG AGAGGGATAC ATGCGCCGAG CACATGCATA 4620
 AAATCCCCTA ACAACTAAAA GTTGTAAAGG GATTTAAATT AATTTAGTGT ATCTTGATA 4680
 35 TCTTGTTTTG kTTGaTTAAT ATCTTCTGTT TTTTCTTCTT TTTTATCTTT TAATTTTTCT 4740
 TCAACTTCTT TAGCTTTTTT TGCTGCTTTT TTATTTTGAT TTTTATTAGA CATGATTAAT 4800
 TCCTCCCAA TTGGATAATT ATTTATATAT AAATCTTACC CGGTTGTACT TTCGTTAAAC 4860
 40 TTTTCTAAGT CTATAGCACT ATTTATTCAT TTATCTAAAG ACAACAACAT TAGATTAATA 4920
 TATAATGATT TTGAGGTGAA CATAATGTCT TTTCTTAGGA AACACGCCGA AATTATTTTT 4980
 AGCTATTTAA TCGGTmwcGT TcACTCTTCA CTGGkcTCAT TATTTTAATT AACTTGCCAT 5040
 TAATTAAACA ATTAAATGGT GGTAAAAAAG TTGATACACA TGTTTATAAT GTGTGGGAAT 5100
 TTCTGAATGC ATTTTTCAGT GAAATTATTA AAGTAATGAG TCGATTTATA GGTAATTTCC 5160
 50 CtATAGTTAG TGCAATTGTG ATAATTATAT TCGGTATTTT AGTTATGTTG ATTGGTCATA 5220
 CATTACTTAG AACTATTAAG TATGACTATG ATATTTCTAT CTTTTTCTTA GTTATCGGTA 5280

55

TTTTCATTAT TCCATTTACA ATTCAATATAG GATATATCGT CTATAAAGAT GAATTGAATC 5400
 AGGAAAATGT AAAAAATCAT TTCATGTGGA TAATTGTGAG TTATGGTATA AGTTACTTAA 5460
 5 TTACACAAAT TGCATTGTAT GGCAGAATTG ATGCTAATGA AATAGAGTCA ATTGATATCT 5520
 TAAGTGTCAA TGCTTTCTTT ATAATTATGT GGTACTTGG TCAAATGGCT ATTTGGAATT 5580
 TCTTGTTCTT GCGCCGAGCT TTACCTTTAA CAAAGCAAGA ATTAGGTGAA GAGGAGCCAG 5640
 10 AATTATCAAG AACAAGTAAA GGAATGTCA CGAATCAAAC TAAATTCAC TTGAAACAAC 5700
 TCCAAGATAA GACTACAGAA TATGCACGTA AGACAAGAAG AAGTGTCGAT TTAGATAAAA 5760
 TTAGAGCTAA AAGAGATAAA TTCAAAAAGA AAGTTAATGA TATTATCGAT ATTCAAGAAG 5820
 15 ACGATATTCC TGATTGGATG AGAAAACCGA AATGGGTAA ACCAATGTAT GTCGAACTAT 5880
 TTTGTGGTGT CGTCATCTTT TTATTCACAT TTTTAGAATT TAATAATCGT AATGCATTAT 5940
 20 TTGTATCTGG TGATTGGAAA TTATCACAGA CACAATATGT TATTGAATGG GTTACATTAT 6000
 TAATTCTGTT ATTCATTATT ATCGCATATA TCGCTACAAC GTTAACTTTC CACTTGAAAG 6060
 GTAAGTTTTA TTATTTACAA TTATTTATGG GGAGCATTTT ATTCTTTAAA TTGTTAACGG 6120
 25 AATTTATAAA TATAATGATT CATGGACTAT TACTTTCAGT GTTCATTACG CCAACATTAC 6180
 TATTAATGTT ATTGGCAATC ATCATTCTTT ATTCGTTACA ATTACGAGAG CGACCATAAT 6240
 TAAAAGCATT ATAAAAGTAC TATCTATTAA ACATTTTGAT GTGTACGCTA TAAGTTAGAT 6300
 30 ATATCTCTAA CTTACTTAGA TACAGGTCAA TGAAGTTTAT GGATAGTACT TTTTTTGTA 6360
 CTAGATTGGA TTGATTCAGG TGATGTGAAT TAAGTATTGA TAATTGTATA CAAAGTTTTA 6420
 GTGCAAATAA AATAGTTGAA AAGTTATCCA TTTGTAAAAT CAAGAAAAC AGTAAATAGT 6480
 35 TGAAGCGACT TATGGaATTT GCGAAACGAT ATATAGTATT TCCTTTGTAG AAATTTmACA 6540
 TATATCATTC AAATTACTAA TTTGTTAAAA TCAACAGTAA GATTAGAAGT AGATGATATT 6600
 40 GAAATTTGGC AAACAaTtTA ATCTATATAA AACTACAACG AAACACAGAA AGGAAGTTGT 6660
 CAGATGAAAA TAGCAACTCT GAACAAAGGC AAAGAAACAA AATATTTTAA TGGATATCCT 6720
 TTAATTGAAG AAGAGGATAT CTATTCACAA GATCATTTAA AAGAAGGAGA TATTTTTCAA 6780
 45 ATGTGACTG ATAAATCACA ATAT 6804

(2) INFORMATION FOR SEQ ID NO: 366:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1717 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:

	aaaAGAGACG CATTaAAACA AGCTATACAA ATTATCGATA AATTAACATG GGGTGTITAG	60
5	TGGTGGTTAA AGAAATTTTG AGACTATTAT TCTTACTAGC GATGTATGAG CTAGGTAAGT	120
	ATGTAAGTGA GCAAGTATAT ATTATGATGA CGGCTAATGA TGATGTAGAG GCGCCAAGTG	180
	ACTTTGAAAA AATCAGAGCT GAAGTTTCAT GGTAATAGCT ATTATCATT TTGAATTAAT	240
10	TATATTAATG TGTGTAGCAA TAGCACTGGA GGTGTTGTAA ATATGTGGAT TGTCAITTCa	300
	ATTGTTTTAT CTATATnTTT ATTGATCTTG TTAAGTAGCA TTTCTCATAA GATGAAAACC	360
	ATAGAAGCAT TGGAGTATAT GAATGCTTAT CTTTTCAAGC AGTTAGTAAA AAATAATGGT	420
15	GTTGAAGGTT TAGAAGATTA TGAAAATGAA GTTGAACGAA TTAGAAAAAG ATTCAAAAGC	480
	TAAAGAGAGG CGTTGGCTTC TCTGCTCTAT CyAAAATAAT GAAAGGAGCC saACATGTTA	540
20	GaCmaAGtCA CTCAAATAGA AACAATTAAA TATGATCGTG ATGTCTCATA TTCTTATGCT	600
	GCTAGTCGTT TATCTACACA TTGGACTAAT CACAATATGG CTTGGTCTGA CTTTATGCAG	660
	AAGCTAGCAC AAACAGTTAG AACTAAAGAA GATTTAACTG AGTACAATAA AATGTCTAAG	720
25	TCTGAACAAG CCGATATAAA AGATGTTGGC GGATTTGTCTG GTGGATATTT AAAAGAAGGC	780
	AAACGGCGTG CTGGTCAAGT CATGAATCGT TCAATGCTAA CACTTGATAT CGATTATGCA	840
	GCCCAAGATA TGAFTGACAT ATTATCTATG TTTTATGATT TTGCATATTG TTTATATTCA	900
30	ACACATAAGC ATAGAGAGAT AAGTCCAAGA CTGCGTTTAG TGATTCCTTT AAAACGAAAT	960
	GTAAATGCAG ATGAGTATGA AGCTATTGGG CGTAAAGTCG CAGATATCGT TGGCATGGAT	1020
	TACTTCGATG ATACAACCTA TCAACCACAT AGGTTAATGT ATTGGCCTTC AACTAGTAAC	1080
35	GATGCGGAAT TTTTCTTTAC CTATGAAGAT TTACCTTTGT TAGACCCAGA TAAATATTA	1140
	AATGAATATG TTGATTGGAC TGACACATTA GAATGGCCAA CGTCTTCAAG GGAAGAGAGT	1200
40	AAGACTAAAA GATTAGCAGA TAAGCAAGGC GACCCAGAAG AAAAGCCGGG AATTGTTGGT	1260
	GCATTTTGTA GAGCCTATAC GATAGAAGAA GCTATAGAAA CTTTTATTCC TGATTATAC	1320
	GAAAAACATT CTACTAACCG TTATACCTAT CATGAAGGTT CAACTGCAGG TGGATTGGTG	1380
45	TTATACGAAA ATAACAAGTT TGCCTATTCT CATCATAATA CGGATCCCGT AAGCGGTATG	1440
	CTTGTGAACA GTTTTGATTT AGTACGCATA CACTTATATG GTGCTCAAGA TGAAGAACT	1500
	AAAACAGATA CTCCGGTTAA TCGACTACCT AGTTATAAAG CAATGCAGCa AAGAGCGCAA	1560
50	AATGATGAGG TTGTTAAAAA GCAATTAATT AATGACAAAA TGTCTGATGC AATGCAGGaT	1620
	TTCGATGAAn GAGAAAAATAG CGATGATGCA TGGTCTGAGA CGTnGAAAT TACTTCGAAA	1680

55

(2) INFORMATION FOR SEQ ID NO: 367:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1847 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

	AAGATTAGGC ATCAAGACAG GGTCCGATT GTTTGAAATC CCACATAGnA ATGATATTTa	60
5	CaTTATCAAT CCaAGTATGC GTAAATATCT TAATGTTTCA GTTGCTATTT CTAAGATTGC	120
10	ATTGCGTTAT ATTCCACCTG AAGATTTACA CCAATATAGT ATTGACGAAT TTTTATGGA	180
	TGTTACTGAT AGCTATCATA GATTAGTTC TACAGTACAT GCATTTTGCG AAAGACTTAA	240
15	ACGTGAAATT TATGAAGAAA CAGGCATTTA TTGTACTGTG GGCATTGGTT CTAATATGTT	300
20	ATTAAGTAAA ATTGCTATGG ATGTTGAAGC GAacATAGTc AAAATGGTAT AGCTGAATGG	360
	CGATATCaAG ATGTACCAAC GAAATTATGG CCAATTcmGC CCtTGCAGAG TTTTGGGGT	420
25	ATTAATCGTC GAACAGAAGC CAAATTGAAT AAAAGAGGAA TTTTACTAT AGGAGATTTA	480
	GCGAAATATC CATATAAATT TTTAAAAAA GAGTTCGGTA TTTTAGGTGT TGATATGCAT	540
30	CTACATGCGA ATGGGATAGA TCAGAGTAAA GTACGTGAAA AGCACAAGAT CAGCAATCCA	600
	TCGATATGCA AAAGTCAAAT ATTAATGAGA GATTATCATT TTGATGAAGC AAAAGTAGTA	660
35	ATGCAAGAGT TAATTGAAGA TGTGCTAGC AGAGTTCGAG CAAGAAAAAA AGTGGCAAGA	720
	ACGATACATT TTGCCTTTGG CTATAGTGAT GAAGGCGGTG TACATAAGCA ATATACTTTG	780
40	AAAGATCCAA CAACTTAGA AAAAGATATT TATAAAGTAG TAATGCATTT CGCAGATAAA	840
	TTATGTAATA AACCAAGCACT ATATCGTACG CTAAGTATAT CTTTGAGTCA ATTTATTAAT	900
45	GAGGATGAGC GACgTTAAGT CTGTTTGAAG ATGAATACCA ACGCAAACGT GACGAATGTC	960
	TAGCTAAAC GATAGACCA TACATTTGA AATACGGCAA AGGTATTGTG TCCAAAGCAG	1020
50	TATCGTTTAC AGAAGCAGGT ACAAACACG GCAGATTAGG TTTAATGGCT GGACATAAAA	1080
	TGTAATGACT ATACGGTTTA AGTAATATAT AACTGTGATT CGTATAAAAT AAGTCTCTAA	1140
55	AGATAAATAT TTCATATATC ACAATAGATT TTCACAATA TATCTAAGAA TACATGGAAT	1200
	TTATCAAAAG AGACTTAATA ATTATTGGAT ATAACAATCA AAATCACTCA ATGCTTGCAT	1260
	ACCGCGTTCT CGGTCAGTAG GGTTTTTGAA ACTAATTTTT AAAGCACCGT ATATATCTTC	1320
	GCGTACTTCT AAGATTCTTA AGTTGCTTAT AGATATGTTA TGTAAACTCA GGATATAAGT	1380

TAGTCCACCT AGTTGTTTAG CGGGTAGTGC GTCGCGATAC GATTTAGCTT GGGCAAAAAA 1500
 TGATAACAAT TTTTCAGAAT CATTGCTTTC AATTAGTCTT TCTAAATCTT GAAACTGACT 1560
 5 TTTTAGCTGT CGAATCATT CTAAAATATA CGTTTTATTA CTCAAGGTGA TATCTTTCCA 1620
 CATTGTGCA TTACTIONACTAG CTATACGAGT GATATCACGA AAACCACCAG CTGCAAGTTT 1680
 10 ATTAATAAAA TGATGTTCTT GACCGTTCTT TTGACTAACA TGAATAAAC TAGATGCAAC 1740
 GATATGAGGT AAATGACTTA CGACGCTTGT TACGTAGTCG TGTTCTTCAG CAGTAGTTAC 1800
 AATAAATTTA GCAAGAGTAG GTGATAACAG TTCTTTTAAC GTGTTTG 1847

(2) INFORMATION FOR SEQ ID NO: 368:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 494 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

AACAAAAGGC ATAAGTTCGT GAATTAATGC GTATACAAGG ATAAAGCTTA TAACAGTAGT 60
 AATTGTTGCT ATCAAACGAA CAACATATAT TCTATTTTCA GATAGCAATC GATTCATAA 120
 TCGATAAATTT ACGTATACAA GAATTAACAA CAGCACAATA TAAACAATAA TCATATTTGG 180
 30 CCCCATTTAA TTTTAAATTT TGTTTTCACA TCATTTTCTA CTTTATTATC ATACTAATTT 240
 TAAAGGCAAA GGTGGACATC GGCAACCTCT CGTAAACTAT TTATCAAAAA TAAACGTATC 300
 TCATTGTTAT GATATTTATA AATCAATTCG TTTTATTAA AGTCTTTTTC AACAGTTTG 360
 35 TCACTATCTA TTAAATAATC ACGCATGCAT CCTTTTAAGA AATCATCTTT ATAACCTGGT 420
 GTGTACCATT TTCCATCCTC TTCAATGACA ATGTTGCCAA TATCAAATTC AAGGACCTTG 480
 40 CCGTCCTCTG AAGT 494

(2) INFORMATION FOR SEQ ID NO: 369:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2518 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

GTCCATATTT CCCGTCCAAC CAACTAAAAA TGGGTAAATC CCTGGATTTa AATCTACTTC 60

	GATAATTTCT	AATGTAGCGA	CCATCATTAC	GAGATGATTT	GATATAAGCA	CAGTTTGGAT	180
	GTTGACCAAT	ACTATCGCCT	TCTTCTTCGA	TGATATCTAT	TTTAATACCA	TCATCAGCTG	240
5	CAATTTCTAA	TGAAGATTTA	ATTGGGTTAT	CAAATGTTGA	ATATCCCATT	GCTCCACCCA	300
	CAATAGCGAC	ATCTGTACCA	TGTCCTTGGT	GTGTTTGAGC	AAATGATTCA	TAATAATGTA	360
	TTTCAATATT	TTTAGGAATA	TCTCCCAATA	TTGCGCGTGC	TGAATTCCCA	ATCTTTACTG	420
10	CACCAGCCGT	ATGAGAACTT	GAAGGGCCCA	TCATAACTGG	TCCGATAATA	TCGAAAGCAC	480
	TTTGATAATC	ATAGCTCTTT	GCCATAATTA	AACACTCTCC	TTAATATGAT	TCTTTTTGCT	540
	CGGCATTTTA	AAGTTGATAT	TCATTAAATT	AAACTTATTA	ATTAGTGTTC	CAATAATATA	600
15	GGCTAAGACA	ATGCTGACAA	TAATCACTGT	TGCGATTGTT	ACAATTGATG	TCACTGCATT	660
	ATTAAAGCCA	AACAATACGA	TGGCGCCTGC	AATTGGTGTG	GCCATACCTT	TGACACCTAT	720
20	TACTAGTCCG	CTAAATGTCA	CGATACATGC	GTTGACGACC	CCAATCAGTG	CATTTGTACC	780
	ATATAGTTGT	ACTGGATATT	GCGCTATTAA	ATCAATTTGC	GTCAATGGCT	CAATACAAAC	840
	TGCAAATGCT	TTTGACGGTC	CACCAATGTT	TAATTTTTCG	AATAAAATAA	GGTTAACAAA	900
25	TGAGCTACCT	GTACATGTTA	GTGCTCCAAT	AGCCATAGGA	ACACCTGTCA	GTCTTAATAA	960
	ACTTGTTAAT	ACCATTGAAC	TTAGCGGTGT	CATACCTGTA	ACAGGAATCA	CTAGTCCTAA	1020
	AATGACCGCT	AATGCATATG	GATTGTTATC	ACCTACCGCA	GTGACAGCAC	TACCTATTTG	1080
30	TTTTAATGTT	GCTAGCACAC	CAGGTGTAAT	GATTGATGCA	AGTCCGAAAG	CAATTGCTGG	1140
	TGCAAATAAG	ATCACCACAA	TTAAGTCCAA	GCCTTCTGGA	ACTTTCCTTT	CAATCCATTT	1200
	AAATAAAAAA	GCTACGCCAT	AAGCTGCGAT	GAATGCTGGT	AATAATTTAA	AGTCATGTAA	1260
35	TACTAAACCA	ACAATGACCG	CAAATACTGG	TGCAACGCCT	AAGTTTAAGC	ACGTTAGAAT	1320
	ACCTACTGCG	ATACCGCTTA	AACTTCCTGC	TAAATCCCCA	ATATCTTGTA	GAAATTTAAT	1380
40	ATCAAATACG	CCACCAATAG	CATAACTTAA	GAATGCTTGT	GGTAGAAATG	TCGCACAAGC	1440
	TGCACCGGAT	AATGCTTGTA	GTCTTGTTTT	ACCGTACGGT	GCATACTTTA	AAAATAGCGT	1500
	CATGATCACT	AAAACCAAGA	CTAATGTGCC	TACACCTAAC	AGAATATCCA	TTTCCCAAAA	1560
45	ACCTCTCTCT	ATGTTTATTT	TATTTTCAGA	CCATAAACAT	CGTACACCCT	TAAGAAAACG	1620
	TTTTCAACTT	TTATCTGTTA	TCAAATCAAA	TATTTAAGTG	AAATATTTCA	TATTTGTGAA	1680
	AGATTTTTTA	AATGGATTGT	TTCAAAAAC	ACTTATGTTG	TCGTTAATAT	TTACTAATTA	1740
50	ACTTTTTACT	CTATATTTCA	AACAGTTGTG	TGACAGTTTT	TTGATAACTT	TTTTACATCT	1800
	GAAAGTAAGT	AATATTTCTA	AAAACCTTTA	ATATTTATAC	ACTTTATCTT	TCGAGCTATT	1860

55

ACACATATAT TTGCAATAAG ATAATTAAAG TAGGATATTA TTTTtagTTT TCTGATAGGA 1980
 ATGATGATAG TCTATAGGTT GAATCTTTAC TTTTTTTTAA AGCTAAATTT ACATCAACTT 2040
 5 AACAAATGGTT GGTtATACTG AAGATGAAAA TATTTAGTAT AACTTAGTGG AGGCGATAAA 2100
 GGTGCAATTG AGTCATTCCG TTAAAGTTGC AATTTCTATC TATTTAGCAC TTATCTTTAT 2160
 AACGTTCACT TCTTATTTAG TCATTATTTT ATATACGAGT ATGACTGGAC ATGATGTATC 2220
 10 ACATTTCTGTG TTAGATAGTC AGCATTCTCA TCATGGATCT CTTACGCAAA AACATTTGAG 2280
 TCTTCCTGAA ATCTCATTTA AATAGTTTAT CCTCTGTGTT TCAACATTCA TTTCCCATAT 2340
 CGATTCATT ATCTATCATC TAGACCACTA CATCTTAGAT GATTTTTTTA TTTTCTCATT 2400
 15 TCACTCTTTC TTAAAGTCG ATATAATGAA TTAAATCATT ATCATACACC GACATATTTT 2460
 ATGTTGTTGG TGTTAAGTTT aaAGGGGTGA GATACTTGGC GAATaATCaT TCAGCTTT 2518

20 (2) INFORMATION FOR SEQ ID NO: 370:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 790 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

30 AATACTACTGG ACCGTCnACC AATTGACAAA ACCTGAAAAC TGGATTTTTA TTGCAGGATT 60
 ACATCGAACC TGGAAGAACA ATCCCCAGGA GTTGATGATA TTCGAGATAT GGCATACAAT 120
 CAAGGTAGTT TAGATAAGAC AATTTATGAA ATTTCTAAAC GCACAGTACT ATTTTAAATA 180
 35 CAGAAAGATA TTACGGTATA TAATAAGACG ATTGACTGTT TAAATTATTA TAACTATAGT 240
 GACGAAAGAA TAAAGGATGA TTAAATGAAT TCACAAGAAT TATTAGCAAT TGCTGTGGAT 300
 GCAATTGACA ATACCCcAGG CGAAGATACG ATTTCTTTAG AAATGAAAGG TATCAGCGAT 360
 40 ATGACAGATT ATTTTGTGTG AACGCACGGA AATAATGAAC GACAAGTTCA AGCGATTGCT 420
 AGAGCGGTGA AAGAAGTAGC CAATGAACAA AATATAGAAG TAAAACGTAT GGAAGGATAC 480
 AATGAAGCGC GTTGGATATT AATTGACTTA GCTGATGTTG TGGTACATGT TTTCCATAAA 540
 45 GACGAAAGAA ATTATTATAA TATTGAAAAG TTATATCAAG ATGCACCATT AGAATCATAT 600
 AGTCAGGTTG CGTATTAATT ATGTCGCAAT ATGCAGAAAT GAGCCTAGTG TACGATCAAT 660
 50 TGACTCAAGA TCAACCATAT GAAAAATGGT TTGAAATTGT AAAAAATCAC TGCAAAGATG 720
 AATCAAATAT TTTAGATATT GGATGCGGTA CTGGTAGTTa ACAGTTCAAT TAGAAGCTTT 780

55

(2) INFORMATION FOR SEQ ID NO: 371:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1823 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

	ATAGATGAAG GTGCAAATAT TGAAaTAGGT TATTTACCTG GACGCTTGAA ATGGTTAGTT	60
	GCTGATTTAT TAACTAAACA AGGATTAAAA GTAGTTAACG ACGATATGAC AGGAAGAACG	120
15	TTAAAAGATC GTAAATTATT AACAGGTGAC AGTCCTTTAG CTTCAAATGA GTTAGGAAAA	180
	TTAGCAGTTA ATGAAATGTT AAATGCAATA CAAAATAAAT AATTAAATAT TAATTAGAGG	240
20	AGCCTCATAT GTAAATGTAT GAGGGCTCTT TTTTGGCA AAATTTAAGT GATACTTGTA	300
	AAATAGAACC TATTATGAGT ATGATTTAAG AAAACGCTTG CAAACTAAT AACCGCAACT	360
	AGCGATATGG AGGAAACATG ATGTCTTATA GCATTGGAAT TGATTATGGA ACTGCTTCAG	420
25	GCCGTGTGTT TTTAATTAAT ACAACTAACG GTCAAGTAGT ATCAAAATTT GTGAAACCAT	480
	ATACACATGG TGTCAATTGAG AGTGAATTAA ATGGTTTGAA AATACCACAT ACATATGCAC	540
	TTCAAAATAG TAATGATTAT tTAGAAATTA TGGAAGAAGG AATATCATAT ATAGTACGTG	600
30	AATCAAAAAT AGATCCAGAC AATATAGTAG GTATTGGTAT AGACTTTACT TCATCTACTA	660
	TTATTTTTAC TGACGAAAAC CTTAACCCGG TACATAACTT AAAACAATTT AAAACAATC	720
	CACATGCGTA TGTGAAACTT TGGAAACATC ATGGTGCATA TAAAGAGGCA GAGAAATTAT	780
35	ATCAAACTGC TATTGAAAAT AATAATAAGT GGTTAGGCCA TTATGGATAT AATGTTAGTA	840
	GTGAATGGAT GATTCCCAA ATAATGGAGG TCATGAATCG AGCACCAGAA ATTATGGAAA	900
40	AAACGGCTTA TATTATGGAA GCGGGCGATT GGATTGTAAA TAAATTAACT AATAAAAATG	960
	TACGCTCGAA TTGTGGATTA GGTTCCAAAG CATTTTGGGA AGAAGAAACA GGGTTTCATT	1020
	ATGATTTTATT TGATAAAATA GACCCCAAAT TATCAAAAGT AATTCAAGAT AAAGTATCTG	1080
45	CACCGGTTGT TAATATTGGT GAAGCAGTAG GGAACTGGA TGATAAAATG GCACAGAAAT	1140
	TAGGATTATC AAAAGAACT ATGGTAAGTC CTTTTATTAT TGATGCCCCAT GCTAGTTTAT	1200
	TAGGTATTGG GTCTGAAAAA GATAAAGAAA TGAATATGGT GATGGGAACA AGCACATGCC	1260
50	ATCTTATGTT AAATGAAAAG CAACATCAAG TGCCAGGTAT ATCAGGTTCT GTAAAAGGAG	1320
	CAATTATTCC AGAATTATTT GCTTATGAAG CGGGGCAATC AGCAGTAGGT GATTTGTTTG	1380

CTGTATTGGA ATTAATGAAT GAAAAGATAA AACATCAAAT GCCAGGTGAA AGTGGGCTCA 1500
 TTGCTCTTGA TTGGCATAAT GGAAATCGAA GTGTATTAAG TGATAGCAAT TTAACAGGTT 1560
 5 GTATCTTTGG ATTAACCTTA CAACTAAGC ATGAGGATAT TTATAGAGCm TATTTaGaAG 1620
 CTACaGCATT TGGTACTAAG ATGATTATGC mACAGTATCA AGATTGGCAT ATGGaAGTAG 1680
 aAAAGGLATT TGCaTGTTGg gGTATACcTA AAAAGAATGC TGTTATGATG GATATCTATG 1740
 10 CGAATGTACT GAATAAAAAA CTAATTGTTA TGGATAGTGA GTATGCACCA GCAATAGGCC 1800
 CAGCAATATT AGGTGCAGTC AGT 1823

(2) INFORMATION FOR SEQ ID NO: 372:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1600 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

25 ACGATCATCT GCATCAGCGA ATTCCGATGc AaTtCATATG tCcTaATAAC AAAAGAATAG 60
 GTTTTaAAAG ATATGCATAC CGTAATGaTA GATATAGTTT TAAACGTGAC TTCAAGCTAT 120
 ATGaATGTGA TGA CTGTTCA TCATGTTCTT TGAGACATCA ATGCATGAAG CCAAATTCGA 180
 30 AATCCAATAA GAAAATTATG AAGAATTATA ATTGGGAATA CTTTAAAGCC CAAATTAATC 240
 AAAAGCTTTC TGAACCAGAA ACGAAAAAAT CTATAGTCAA AGAAAAATTG ATGTAGAGCC 300
 TGTTTTTGA TTTATGAAGG CTATTTTGGG TtTCACTCGA ATGTCAGTTC GGAATAAAT 360
 35 AAAGTTAAAC GAGAGCTAGG TTTGTATTA ATGGCACTTA ATATAAGGAA AATAgcaGCT 420
 CAACGAGCTG TACATTATAA AATACATATC AAAAAAGCTG ATTTCTATCA AATAAATAAT 480
 AGAAATCAGC TTTTTTACAT TGCCTAAGAA CTTTAAGGAA CTTAATGTCC CAAGCTCTTT 540
 40 TTTGTATAT CTAATTCGTA ATTTATGATT GTTTATTCCG TCCTTTGATG TTTACTAAAT 600
 GTGACTTTAA ATCTTGTTCT AATTGTTGTA ATTCTTTTTC AGCTAATTGT CGTTCTTCGC 660
 45 GACCGTGTTG TTGAATAATT AATGTTTCTT CAATTGTCTC AATAATGTTA CGCTGTGTAC 720
 GTTTCaATGT ATCAAGATCA ACAATGCCAC GTCATTTTTT TGTTGCAGTT TCAATCGCAT 780
 TTTGTTTCAA CATTTcAGCA TTTGCTGTTA ATAAATCATT AGTTGTATCA GTGACAGCTC 840
 50 GTTGTGCAGC AACTGCATTA CGTGCTCTCA TTAATGTAAG CGCAATGGCC ATTTGATTTT 900
 TCCATAGTGG AATACTTGTC AAAATTGAAC TTTGTATCTT CTCGGCAAGT GCTTGATTAA 960

5 AGTCATATAT GCGTTTATCT AGTCTATCTA TAAATTGCTG CATATCTGCA ACTTGTTGAA 1080
 TATCCATTTC ATTAGTGGAT TgtGCGCTTG CTGTTGCAAT TGTGGTAGCT TTTCATTTC 1140
 10 TAATTGCAAC TTTTTTTGct GTGCAGCAAT GATATGCAAT GATAAGTCAT CAAAGTATTG 1200
 TTTGTTTTTA TCATATAGCG TATCTAATAA TTCAATATCT CTTGTTAAAT GTGTTTGATG 1260
 TTTCTGCAGT TGAATCGTTA TGCGATCGAC TTGAGCACTA ACTGATTGCA TTCTTGAAAA 1320
 15 GATTTCATTG ATAGACGACT TTGCTCTGCT AAAAATTCTT TTTAACATAG ATGGTTTATC 1380
 AGTATTTAAC TCATTGGAT TAACTGACTT TAGTTTTGAC ATTAGATCTG ACAAAGTATC 1440
 TCCAATAGGA CCAACATCTT TACTTTGTAC TTCATCCAAC ATTTGATGTG AAAATTGAGA 1500
 20 CATTTGTTTC TGKAAATCAG AACCAAACGC TAATAAACCT TCATTGTCTA AAGGGTTAAT 1560
 TTGTTTACTG ATTGTGTCTA CCTGTTTTTG TTGTTCAATT 1600

(2) INFORMATION FOR SEQ ID NO: 373:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1227 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

30 ATGTTGATAA TGGTAGCTTC TACAAGAATA AAGACCAACA AGTTGGTGCA ACAATTCCTG 60
 aTAGTAAAAC TGGTGGTTTA GTTGCTATAT CTGGTGACG TGATTTCAAA GACGTCGTTA 120
 ACAGAAACCA AGCAACAGAT CCTCACCCTA CTGGTTCATC TTTAAAACCT TTCTTAGCGT 180
 35 ATGGACCTGC CATTGAAAAT ATGAAATGGG CAACAAACCA TGCGATTCAA GATGAATCTT 240
 CATATCAAGT TGATGGTTCT ACATTTAGAA ACTATGATAC GAAGAGTCAC GGTACTGTAT 300
 CTATTTATGA TGCTTTACGA CAAAGTTTCA ATATCCCAGC TTTAAAAGCT TGGCAATCAG 360
 40 TTAAGCAAAA TGCTGGTAAT GATGCACCTA AGAAATTCGC TGCCAAACTT GGCTTAACT 420
 ACGAAGGCGA TATTGGTCCA TCTGAAGTAC TTGGTGGTTC TGCTTCAGAA TTCTCACCAA 480
 45 CACAATTAGC ATCAGCATTT GCTGCAATCG CTAACGGTGG TACTTATAAC AACGCGCATT 540
 CAATTCAAAA AGTAGTTACT CGTGATGGTG AAACAATCGA ATACGATCAT ACTAGCCATA 600
 AAGCGATGAG TGATTACACT GCATACATGT TAGCTGAGAT GCTAAAAGGT ACATTTAAAC 660
 50 CATATGGTTC TGCATATGGC CATGGTGTAT CTGGAGTAAA TATGGGTGCT AAGACAGGTA 720
 CTGGTACTTA CGGTGCTGAA ACTTATTCAC AATATAATTT ACCTGATAAT GCAGCGAAAG 780

AAGTTAAACA ATATGGTGAA AACTCATTTG TrGGACATAG CCAACAAGAA TATCCACAGT 900
 TCTTATATGA AAATGTGATG TCAAAAATTT CATCTAGAGA TGGCGAAGAC TTTAAACGTC 960
 5 CTAGCTCAGT AAGTGGTAGT ATCCCATCAA TCAATGTTTC TGGTAGTCAA GATAACAACA 1020
 CTACAAATCG TAGTACACAC GGTGGTAGTG ACACATCAGC AAACAGCAGT GGTACTGCAC 1080
 AATCAAATAA CAATACTAGA TCTCAACAAT CTAGAAACAG CGGTGGATTA ACAGGTATAT 1140
 10 TCAACTAATC CACTCAACAT AAAATCCTCA GTTATACCAT ATTTATGGTG TAGCCGAGGA 1200
 TTTTnTTAGG TTCTTCATCT TTTATGG 1227

(2) INFORMATION FOR SEQ ID NO: 374:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1953 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

25 CCATATGGtG CAACATTCTT CGTATTTAGT GATTATTTAA AACCAGCGTT ACGTTTATCA 60
 TCAATTATGG GATTAAATgC aACGTTcATC TTCACACATG AyTCaATTGC AGTAGGTGAA 120
 GATGGTCCTA CTCATGAACC AATTGAGCAA TTAGCTGGAT TAAGAGCCAT TCCAAATATG 180
 30 AATGTTATCC GTCCTGCTGA TGGTAATGAA ACAAGAGTAG CATGGGAAGT TGCCTTAGAA 240
 TCTGAATCTA CACCTACTTC ATTAGTATTG ACACGTCAAA ACTTACCGGT ATTAGATGTA 300
 CCAGAAGATG TAGTTGAAGA AGGCGTTCGA AAAGTGcCT ATACAGTTTA TGGCTCTGAA 360
 35 GAGACACCAG AATTCCTATT ATTAGCTTCA GGTTCAGAAG TTAGTCTTGC AGTTGAAGCT 420
 GCTAAAGATC TTGAAAAACA AGGTAAATCA GTACGTGTTG TTTCAATGCC TAACTGGAAT 480
 GCATTTGAAC AACAATCTGA AGAATATAAA GAATCAGTTA TTCCATCAAG CGTAACAAAA 540
 40 CGTGTTCGCA TTGAAATGGC TTCACCGCTT GGATGGCATA AATATGTAGG TACTGCAGGT 600
 AAAGTTATTG CTATTGACGG CTTTGGCGCA AGTGcACCTG GCGATTTAGT AGTTGAAAAA 660
 45 TATGGATTTA CAAAAGAAAA TATCTTAAAC CAAGTTATGA GCTTATAAGA ATAATTTATA 720
 AAGCGAGTAT GTTTAGAAGT CTAGGATGCA TAATCTTAGG CTTCTTTTAA AGTGTGAAA 780
 TTTAGAGTAT AGCACTTAAA CTACATCATA AGTGATAAGT TATGAAAGTA TACTATTTCA 840
 50 GATTAATCTT TAAAAGCTCT GTTATAACAG CATGATTTT GATATTATTT TTAGTATCGA 900
 TATTAAATA CTTGAATAAA CTAGTCTTG AAATAATGTG ATGAATTTAG TAAAATTCAG 960

TGCATTAATT TTAGGTTTAA TTGGAGGTTT CCTTTTAGCT AGAAAATATA TGATGGACTA 1080
 CTTGAAGAAA AACCACCAAA TCAACGAAGA AATGCTTCGT ATGATGATGA TGCAAATGGG 1140
 5 TCAAAAACCT TCTCAGAAGA AAATTAATCA AATGATGACG ATGATGAATA AAAATATGGA 1200
 TCAAAATATG AAGAGTGCGA AAAAGTAAAT TCGCAATTGA TAGAGGCTAT TTTCCAGATA 1260
 TGGAAATGGC CTCTTTTAT AATCAAATTA ATAAGAATAA ATATGTTTAT TAAAATTAAA 1320
 10 GTTAACAAAA TGACGAATAG ACTGAGAAAT GCTATAATTC ATTTTGTATG ATTTACAGAG 1380
 AGTTTATTTA ACGAGAAGGT GTCyGCGTGC TCTATTTAAT ATTTTCAATC ATTGTAGCTT 1440
 TATTTATGGG AACTATAGTT ATAGTTATTC GTATGAAAGC TCAAAATTAT CCGGTAAAKG 1500
 15 AGAAAAAAT AGTTTTGCCa CCgTTTTTTA tGGCgACCGG TGCATTGATG TACGTCGTTT 1560
 CaTATTTTAG GCTAACAGGA TCGGAAATGC TAGAAGCCTT TATAATTGGT TTGCTTTTTT 1620
 20 CtACAGTTCT AATTTGGACT TCTCGATTG AAGTCAAAGG TACAGAAATT TATATGAAAC 1680
 GATCTAAAGC ATTTCCAGTT ATTTTGATT CATTACTTAT CATTCTACT GTGATGAAAA 1740
 TATTCATTAG TAATGAAATA GATCCTGGAG AATTAGGCGG CATGTTCTTT TTATTAGCAT 1800
 25 TCTGTATGAT TGTTCTTGG AGAGCAGCAA TGCTATATAA ATACAAAAA CTAAAGAAAA 1860
 CATTAAATCAA TTAATTACTT TTaAAACCAC TTGTGATCGA CTTCTAAATC AGTCAATGAG 1920
 TGGGTTTAAT nTTACTTGGA AAAGGnGGAA AGG 1953

30 (2) INFORMATION FOR SEQ ID NO: 375:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3787 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

ACATTTGATC AAntTATCGA CATTAAAGAT GAATTCAntT GATCGTnTCA ATGATTATCC 60
 TGTTGAAGTA GCACGTTTGC TTGATATAGT GGAmATaAAA GTACACGCAT TACATTCAGG 120
 45 TATCCACGTT GATTAAAGAT AAAGGGAAAA TAATTGATAT TCATTTATCT GTAAAAGCCA 180
 CTGAAAATAT TGATGGCGAA GTGCTGTTCA AAGCAACACA ACCTTTAGGT AGAACAATGA 240
 AGGTTGGTGT TCAAAATAAT GCAATG:CAA TTACTTTAAC GAAACAAAAT CAATGGCTTG 300
 50 ATAGTTTGAA GTTTTtagTT AAGTGCATTG AAGAAAGTAT GAGAATCAGT GATGAAGCAT 360
 AAAGAAGCAT TTAATGGCGT TGTCGTGTTA ACTGCTGCAT TAATTGTCAT TAAAATTCTG 420

55

	CAACAAGTGT ATCCAATTGT AGCATTAGGA ATGATATTAT CGATGAATGC CATTCTTAGT	540
	GCAATTACAC AAAATATAGG GAAGTATCAT AGTGACGAAG CATATGCAA AGCAGTCGCT	600
5	TATATACAAT TAGTTGGTAT ATTATTATTT ATTGCTATTT TTGTGTTTGC GAACAATATT	660
	GCACATATGA TGGGTGATGG CCATTTAACA CCAATGATTC AAGCTGCAAG TTTAAGCTTT	720
	ATATTTATAG GTATGCTTGG CGTGTTAAGA GGTATTATC AATCTGCAA TAATATGACA	780
10	GTTCCGGCTA TTTCCAGGT TATAGAACAA GTTATACGAG TAGGTATTAT CATTGTTACT	840
	ATTGTTATTT TTGTAGACAG AGGTTGGACG ATATATGAAG CGGGAACAAT TGCTATTTTA	900
	GCATCAACGA TAGGTTTTTT AGGTTCTTCA ATTTATTTAG TAGCGCACCG ACCTTTTAAG	960
15	TTTAAATGG TAAATAACAC TGCAAAGATC GTTTGGAAAC AGTTCGCACT TTCGGTTTTG	1020
	ATTTTCGCTA TCAGTCAATT AATCGTAATT TTATGGCAAG TGATTGATAG TGTTACTATT	1080
20	ATTAAGTCAC TTCAAGCGAT ACGCGTGCCA TTCGATGTTG CCATAACTGA AAAAGGAGTC	1140
	TATGACCGTG GTGCATCATT TATTCAGATG GGATTGATTG TAACTACAAC ATTTAGTTTT	1200
	GCGCTCATT CTCTGTTAAG TGACGCAATC AAAATGAATA ATCAGGTACT TATGAATCGT	1260
25	TATGCAAATG CGTCATTAAA GATTACGATT TTAATAAGTA CAGCAGCGGG AATAGGATTA	1320
	ATTAATTTAT TGCCTTTAAT GAACGGTGTG TTTTAAAGA CGAATGATTT AACCTTAACG	1380
	TTAAGTGTTT ATATGATTAC GGTCAATTGT GTATCGTTAA TTATGATGGA TATGGCATT	1440
30	TTACAAGCGC AACATGCTGT GAGACCTATT TTTGTGGTA TGACGGCAGG ATTGGTTATT	1500
	AAATTTATAC TTAATATCAT TTTGATTCGT TTAAGTGGCA TTATTGGTGC GAGCATTAGT	1560
	ACTGTTGTAT CATTAAATTAT ATTCGGTACG ATTATCCATA TTGCTGTCAC GAGAAAATAC	1620
35	CACCTATATG CGATGAGACG ATTTTTATC AATGTTGTTT TAGGTATGGT ATTTATGTCG	1680
	ATTGTTGTTT AATGCGTGT AAACATAGTG ACAACACACG GTAGAATCAC TGGACTCATT	1740
40	GAATTATTAT GTGCAGCAGT ATTAGGTATC ATTGCATTGT TTTCTATAT TTTTAGATTT	1800
	AATGTTTTGA CATATAAAGA GTTAACTTAT TTACCATTGT GTTCAAAGTT GTATCAAATT	1860
	AAGAAAGGAA GACGTTGATG GCACATACCA TTACGATTGT TGGCTTAGGA AACTATGGCA	1920
45	TTGATGATTT GCCGCTAGGG ATATATAAAT TTTTAAAGAC ACAAGATAAA GTTTATGCAA	1980
	GAACGTTAGA TCATCCAGTT ATAGAATCAT TGCAAGATGA ATTAACATTT CAGAGTTTTG	2040
	ACCATGTTTA TGAAGCACAT AACCAATTTG AAGATGTCTA TATTGATATT GTGGCGCAAT	2100
50	TGGTTGAAGC TGCTAATGAA AAAGATATTG TCTATGCGGT TCCGGGTCAT CCTAGAGTTG	2160
	CTGAGACAAC TACAGTGAAA TTAAGTGGCTT TAGCAAAGGA CAATACTGAT ATAGATGTGA	2220

	ATGATGGCTT CACACTGTTA GATGCGACAT CATTACAAGA AGTAACACTT AATGTTAGAA	2340
	CGCATACATT GATTACGCAA GTTTATAGTG CAATGGTTGC TGCTAATTTG AAAATCACTT	2400
5	TAATGGAACG ATATCCTGAT GATTACCCTG TTCAAATTGT CACTGGTGCA CGAAGCGATG	2460
	GTGCGGATAA CGTTGTGACA TGCCCATAT ATGAATTGGA TCATGATGAA AATGCATTCA	2520
	ATAATTTGAC GAGTGTATTC GTACCAAAAA TCATAACATC GACATATTTG TATCATGACT	2580
10	TTGATTTTGC AACGGAAGTG ATTGATACTT TAGTTGATGA AGATAAAGGT TGTCCATGGG	2640
	ATAAAGTGCA AACGCaTGmA AcgCTAAAGC GTTATTTACT TGAAGAAACA TTTGAATTGT	2700
	TCGAAGCTAT TGACAATGAA GATGATTGGC ATATGATTGA AGAACTAGGA GATATTTTAT	2760
15	TACAAGTGT ATTGCATACT AGTATTGGTA AAAAGAAGG GTATATCGAC ATTAAAGAAG	2820
	TGATTACAAG TCTTAATGCT AAAATGATTC GTAGACACCC ACACATATTT GGTGATGCCA	2880
20	ATGCTGAAAC TATCGATGAC TTAAGAAGAA TTTGGTCTAA GGCGAAGAT GCTGAAGGTA	2940
	AACAGCCAAG AGTTAAATTT GAAAAAGTAT TTGCAGAGCA TTTTTTAAAT TTATATGAGA	3000
	AGACGAAGGA TAAGTCATTT GATGAGGCCG CGTTAAAGCA GTGGCTAGAA AAAGGGGAGA	3060
25	GTAATACATG AGATTAGATA AATATTTAAA AGTATCACGG TTAATAAAGC GACGTACGCT	3120
	AGCAAAAGAA GTAAGTGATC AAGGTAGAAT TACAATAAAT GGTAATGTTG CTAAAGCTGG	3180
	ATCGGATGTT AAAGTTGAAG ATGTGCTGAC GATTGCTTTT GGTCAAAAAT TAGTAACAGT	3240
30	TAAAGTAACT GCATTAAATG AACATGCATC TAAAGATAAC GCGAAGGGTA TGTATGAAAT	3300
	CATTGAAGAG CGTCGACTTG AAGAAGCGTA AATTGGAGGT GACAAGCAAT GAAAAATAAA	3360
	GTAGAACATA TAGAAAATCA GTACACGTCG CAAGAGAACA AGAAAAACA ACGTCAAAAA	3420
35	ATGAAAATGC GTGTTGTTcG TAGGCGTATT ACAGTATTTG cGGGCGTATT aCTTGCGATA	3480
	ATTGTTGTTT TATCaATCTT GCTTGTTGTC CAAAAACATC GCAATGATAT TGATGCACAG	3540
	GAGCGAAAAG CGAAGAAGC ACAGTTTCAA AAGCAACAAA ATGAAGAAAT TCGGTTAAAA	3600
40	GAAAAGTTGA ATAATCTGAA TGACAAAGAT TACATTGAAA AAATTGCGCG TGATGATTAT	3660
	TACTTAAGCA ACAAAGGTGA AGTGATTTTT AGGTTGCCAG AAGACAAAGA TTCGTCTAGC	3720
45	TCAAAATCTT CGAAAAAATA AATCCAAATT GATTCAAAAT TATCCGAGTA TAGACATTGT	3780
	GAAAAAA	3787

(2) INFORMATION FOR SEQ ID NO: 376:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1644 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

5	TAAACCATTT CAAACTGAGG AACGCnAAGA CGGsACgTTT CCAGATTTAG AAGTATTTAA	60
	AAATGAATGT GATTTAAGCT ATGACATAAC GTCACTTTAT ACTTTTAAGC AACCTGTATC	120
	ACCAACACCTT GCATTTAAAA TGACAGATCA AATTTTCTA AATAAGCAGC GTGTATTAGA	180
10	TAAGGTAAAA GTTTTAGATA AGGAATTGA TTTTATCTTA ATTGAGGGTG CTGGGGGAAT	240
	TGCCGTACCA ATATATGAAG GTACAGATGA TTTCTACATG ACTAAAGATC TAATCAATGA	300
	TTGTGCAGAT TGTGTCATCA GTGTGTTGCC ATCAAAATTA GGTGCTATTA GCGATGCCAT	360
15	TGTTCAACCA GATTATGTTA ATCAGAATGT ATCGGCGAGT AATTTTTTAA TAATGAATCG	420
	CTATACAGAC AGCTATATTG AAAAAGACAA TCAATGACG ATTGGAAAAT TAACAAATAA	480
20	AACAGTCTAT ACATTTGAAG AACATGCCAC GTATGAAAAT TTCTCAGAAG CATTTTTAAA	540
	ACAATTAATA GGAGTTAAAA ATGAATTACA CACAACAACT TAAACAAAAA GACTCAGAAT	600
	ATGTTTGGCA TCCATTTACA CAAATGGGTG TATATAGCAA AGAAGAAGCA ATCATCATTG	660
25	AAAAAGGAAA GGGTAGTTAC CTTTACGATA CGAATGGCaA TAAATATTTA GATGGTTATG	720
	CATCGTTGTG GGTCAATGTG CATGGTCATA ATAACAAATA CTTnAATAAG GTAATTAAAA	780
	AGCAACTCAA TAAAATTGCC CATTCTACGC TGCTAGGATC ATCAAATATT CCGTCAATAG	840
30	AAC TTGCGGA AAAATTAATC GAAATCACGC CAAGTAATCT AAGAAAAGTA TTTTATTCTG	900
	ATACAGGCAG TGGCTCTGTT GAAATCGCAA TAAAGATGGC ATATCAGTAT TGGAAAAATA	960
	TTGATAGAGA AAAATATGCC AAGAAAAACA AGTTTATAAC GCTAAATCAC GGTTATCATG	1020
35	GGGATACGAT TGGTGC GGTA AGTGTGGTG GTATCAAGAC CTTTCATAAA ATATTTAAAG	1080
	ACTTAATATT TGAGAATATT CAAGTAGAAA GCCCATCTTT CTATCGCAGT AATTACGATA	1140
40	CTGAAAATGA AATGATGACA GCTATTTTAA CGAATATAGA GCAAATTCTA ATTGAAAGAA	1200
	ATGATGAAAT CGCAGGGTTT ATATTGGAAC CGTTGATTCA AGGTGCGACA GGCTTGTTTG	1260
	TTCATCCTAA AGGCTTTTTG AAAGAAGTCG AGAAATTGTG CAAAAAATAC GATGTCTTAT	1320
45	TAATTTGTGA TGAGGTAGCA GTTGGTTTTG GGAGAAGTGG AAAGATGTTT GCATGCAATC	1380
	ATGAAGATGT TCAACCGGAT ATTATGTGTT TAGGTAAGGC GATTACTGGT GGCTACTTAC	1440
	CACTTGACAGC TACATTGACA TCTAAAAAAA TATACAATGC ATTTTAAAGT GATTTCGCATG	1500
50	GTGTGAATAC CTTTTCCAT GGTCaTACAT ACaCCGGAAA TCAATcGTT TGTaCGGTTG	1560
	cATTaGaAAA TATaAGaCTT TATGaAAAAC GTaAGTtnAT TGTgCACATa TTGaAACGaC	1620

55

(2) INFORMATION FOR SEQ ID NO: 377:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 431 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

ATGATTTTtA aAAATCATTa AGTTAAGGTr GATACACATC TTGTCATATG ATCAAATGGT 60
 TTCGCCAAAA ATCAATAATC AGACAACAAA ATGTGCGAAC TCGATATTTT ACACGACTCT 120
 CTTTACCAAT TCTGCCCCGA ATTACACTTA AAACGACTCA ACAGCTTAAC GTTGGCTTGC 180
 CACGCMTTAC TTGACTGTAA AACTCTCACT CTTACCGAAC TTGGCCGTAA CCTGCCAACC 240
 AAAGCGAGAA CAAAACATAA CATCAAACGA ATCGACCGAT TGTTAGGTAA TCGTCACCTC 300
 CACAAAGAGC GACTCGCTGT ATACCGTTGG CATGCTAGCT TTATCTGTTC GGGCAATACG 360
 aTGCCCATTG TACTTGTTGA CTGGTCTGat ATCcGTGAGC AAAAACGGCT TATGGTnTTG 420
 CGAGCTTCAG T 431

(2) INFORMATION FOR SEQ ID NO: 378:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2006 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

TTTnTATAAC GTATTATAAA TCGTTAAAAA TTTTGTTGT GTTTGCGTCA CGTAGACAAC 60
 CTCCATAAAG TTACTTAATC ACTCTCATCA TACAATAATT TTTACTCAA TTGGAAnAAT 120
 TATAAAATT AAATATAGAT AGGCTTTGAA AATTAGTTTT ATACAAGGTT AGTAGCTGTA 180
 ACTGTAAAAAT GTTCTTAATA TTGTCAAAT GTAATGCTTG AAAGCGCTTT TAAAnAATAT 240
 TATTATATAC ATGGTTAGAC AAATAGACAA ATCACTATAC AAATATTGGG AGGAATATTT 300
 TATGAAATCA ACACCACACA TTAAACCAAT GAATGACGTC GAAATTGCAG AAACGGTTCT 360
 ATTGCCAGGA GATCCGTAA GAGCTAAGTT CATTGCAGAA ACTTATTTGG ATGATGTGGA 420
 ACAGTTCAAT ACAGTGCAGAA ACATGTTTGG TTTTACCGGA ACATATAAAG GTAAAAAAGT 480
 TTCTGTCAATG GGTTCCAGGTA TGkGTATGCC ATCTATTGGC ATTTACTCTT ATGAATTAAT 540

	CATTGATTTA TATGATGTGA TTaTTkCACA AGGTGCCTCT ACTGATTCAA ATTACGTTCA	660
	ACAATATCAA TTACCAGGTC ATTTTGGCGCC AATTGCTTCT TATCAATTAT TAGAAAAAGC	720
5	AGTTGAAACA GCACGTGACA AAGGTGTACG TCATCATGTA GGTAAATGTGT TATCAAGTGA	780
	TATTTTCTAT AACGCGGATA CAACAGCGAG TGAACGTTGG ATGCGTATGG GTATTTTAGG	840
	TGTAGAAATG GAATCaGCTG CaTTATACAT GaATGCaATT TACGCTGGTG TCGAAGCATT	900
10	AGGTGTGTTT ACAGTGAGCG ATCATTTAAT TCATGAAACG TCAACAACAC CTGAGGAAAG	960
	GGAACGTGCA TTTaCAGATA TGATTGAAAT TGCACGTGCA TTGGTGTAGA TGATTATGAA	1020
	TGTTGAATAT TCTAAAATAA AGAAAGCAGT ACCTATTTTA TTATTCTTAT TTGTATTCAG	1080
15	TTTGGTTATA GACAATCAT TTAAATTGAT TTCTGTAGCC ATTGCTGATG ACTTAAACAT	1140
	ATCTGTAACG ACAGTAAGTT GGCAAGCGAC ATTAGCCGGT TTAGTAATTG GTATTGGCGC	1200
	TGTAGTATAC GCTTCATTAT CTGATGCCAT TAGTATACGC AACTATTTTA TTTATGGCGT	1260
20	GATATTAATC ATTATCGGAT CAATTATTGG TTACATTTTC CAACATCAAT TCCCATTACT	1320
	TTTAGTTGGA CGTATTATTC AAAGTCCCGG TTTAGCTGCT GCAGAGACAT TATATGTGAT	1380
	ATATGTTGCA AAGTATCTTT CTAAAGAGGA CCAGAAGACT TACCTTGGCT TAAGTACGAG	1440
25	CAGTTATTCC TTGTCATTAG TTATCGGTAC ATTATCAGGT GGATTTATTT CTACGTATTT	1500
	AACTGGACA AATATGTTTT TAATTGCATT AATCGTAGTA TTTACGTTGC CATTCTTATT	1560
30	TAAATTATTa CCAAAAGAAA ATAATACGAA TAAAGCTCAT TTAGATTTTG TTGGCTTAAT	1620
	TCTAGTGGCA ACTATTGCTA CAACAGTCAT GCTGTTTATT ACGAACTTTA ATTGGTTATA	1680
	TATGATTGGT GCCTTAATTG CGATTATCGT TTTTGGCGTA TATATTAAAA ATGCGCAACG	1740
35	TCCATTAGTA AATAAATCAT TTTTCCAAAA TAAACGTTAT GCTTCATTTT TATTTATAGT	1800
	ATTTGTAATG TATGCTATCC AATTGGGTTA TATTTTACG TCCCATTCA TAATGGAGCA	1860
	AATTTATCAT CTGcAACTAG ACACAACATC ACTGTTATTA GTACCGGGgT TaTATAGTAG	1920
40	CAGTCATTGT TGgtGgCACT AAGTGGgTta AAATCGGgCG rAATATCTGG AATTCCAAAA	1980
	CCAAGCGGAT TATCACAGCC AATTAA	2006

(2) INFORMATION FOR SEQ ID NO: 379:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4858 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

	TGGGGAAAAA AAGACCCAGC AGTATTAGAA GAATCGTTAA ATATTTCTAT TGAAGAAATG	60
	AATCGTATCA TAAAATTAGT CGAAGAATTA CTTGAATTGA CTAAAGGAGA TGTAAATGAC	120
5	ATTTCTTCTG AAGCACAGAC CGTGCATATT AATGATGAAA TTCGCTCGCG AATACACTCA	180
	TTAAAACAAT TGCATCCTGA TTATCAATTT GATACGGATC TGACATCTAA AAATCTAGAA	240
10	ATTAAATGA AACCTCATCA ATTCGAACAA TTATTTTTTAA TCTTTATTGA TAATGCAATC	300
	AAATATGATG TGAAGAATAA GAAAATTAAA GTTAAGACAA GGTAAAAAA TAAGCAAAAA	360
	ATAATTGAAA TTACAGATCA TGAATTGGT ATTCCAGAGG AAGATCAAGA TTTCATTTTT	420
15	GATCGCTTTT ATCgAGTGGA TAAATCTCGT TCAAGAAGTC AAGGCGGTAA TGGACTCGGA	480
	TTATCTATTG CTCAAAAAAT CATTCAATTA AACGGAGGAT CGATTAAAT TAAAAGTGAA	540
	ATTAACAAAG GAACAACGTT TAAATCATA TTTTAATCAT GACTGAGACG TCAATCAAAG	600
20	TCATAGGATC AATTTTTTAA GTACACATTA GCTGTGACTA ATGTATAAGA ACAACTATAA	660
	AACAAATAAA CAGTGGTTCT TTATCATTTC TGTTGTACTC CCAAATTTA CAATAAAATA	720
	CATCTATAAA CCTAGAAGAA TCAACGCTTT TGTTGATTCT TCTTTTTAGC AGATAAATAG	780
25	GTAAATCTAC TTTAACAAAT AACTAAATAG TGATATTATT ACATTGTAAG CGTTTCAACA	840
	TTTTGTGGA GGGTGTA AAA TGACTAACGA AAGAAAAGAA GTTTCAGAGG CTCCTGTAAA	900
30	CTTCGGTGCG AATTTAGGTC TAATGTTAGA TCTATATGAT GACTTTTTTAC AAGATCCATC	960
	ATCTGTACCA GAAGATTAC AAGTCTTATT CAGCACAATT AAGAATGATG ACTCAATTGT	1020
	ACCAGCTTTA AAAAGTACAA GTAGTCAAAA TAGCGACGGC ACAATTAAGC GTGTCATGCG	1080
35	TTTAATTGAT AATATTCGCC AATACGGGCA TCTTAAAGCC GATATTTATC CTGTAAATCC	1140
	TCCAAAAAGG AAACATGTAC CTAAATTAGA GATTGAAGAC TTTGATTAG ATCAACAGAC	1200
	TTTGGAAGGT ATATCAGCAG GAATTGTTTC AGATCACTTT GCCGACATTT ATGATAATGC	1260
40	TTATGAAGCA ATTTAAGAA TGGAAAAACG TTACAAAGGA CCAATTGCAT TTGAGTATAC	1320
	ACATATTAAT AACAATACCG AACGTGGTTG GTTAAAAAGA AGAATTGAAA CGCCATATAA	1380
	AGTAACGTTA AATAATAACG AAAAAAGGC ACTATTCAA CAATTAGCGT ATGTTGAAGG	1440
45	GTTTGAAAAA TATCTTCATA AAACTTCGT TGGTGCAAAG CGTTTTTCAA TTGAAGGGGT	1500
	AGACGCACTT GTACCGATGT TACAACGTAC TATTACGATT GCTGCGAAAG AAGGTATTAA	1560
50	AAATATACAA ATAGGCATGG CTCACCGTGG ACGTTTAAAC GTTTTAAACG ATGTCTTAGA	1620
	AAAACCGTAC GAAATGATGA TTTCAGAATT TATGCATACA GATCCAATGA AATTCTTACC	1680
55	TGAAGATGGT AGCTTGCACT TAACTGCTGG ATGGACTGGT GATGTGAAAT ATCACCTTGG	1740

	AAGTCACTTG GAAATTGTTG CACCTGTTGT TGAGGGGCGT ACGAGAGCAG CACAAGATGA	1860
	TACACAACGA GCTGGGGCTC CGACGACTGA TCATCATAAA GCAATGCCAA TTATTATACA	1920
5	TGGCGATGCT GCTTATCCTG GTCAAGGAAT TAACTTCGAA ACAATGAACT TAGGAAACTT	1980
	GAAAGGCTAT TCTACGGGTG GTTCATTGCA TATTATTACT AACAATAGAA TTGGATTTAC	2040
	TACAGAACCA ATTGATGCAC GTTCAACAAC TTATTCTACA GATGTGGCCA AAGGTTATGA	2100
10	TGTGCCAATA TTCCATGTCA ATGCAGATGA CGTTGAAGCT ACTATTGAAG CAATTGATAT	2160
	TGCAATGGAA TTTAGAAAAG AGTTTCATAA AGACGTCGTT ATTGATTTAG TAGGTTATCG	2220
	TCGTTTCGGA CATAACGAAA TGGATGAACC ATCAATTACT AATCCaGTTT CTTATCAGAA	2280
15	TATTGCGAAA CATGACTCTG TTGAATATGT GTTTGGTAAA AAGCTTGTTA ATGAAGGTGT	2340
	CATTTTCAGAA GATGAAATGC ATTCATTTAT AGAACAAGTC CAAAAGGAAC TAAGACAAGC	2400
20	TCATGATAAA ATTAATAAAG CTGATAAAAT GGATAATCCA GATATGGAAA AGCCTGCAGA	2460
	TCTTGCAATTA CCGTTACAAG CAGACGAACA ATCATTTACT TTTGATCACT TGAAAGAAAT	2520
	AAATGATGCA TTGTTAACAT ATCCGGATGG CTTTAACATT TTGAAAAAGT TAAACAAAGT	2580
25	TCTTGAGAAG CGTCATGAGC CGTTTAATAA AGAAGATGGT TTAGTTGATT GGGCACAAGC	2640
	AGAACAACCTT GCATTTGCGA CAATTTTACA AGATGGTACA CCGATTCGCT TAACTGGTCA	2700
	AGATAGTGAA CGTGGTACAT TCAGTCATAG GCATGCCGTG TTACATGATG AGCAAACAGG	2760
30	TGAAACATAT ACACCTTTAC ATCATGTTCC TGATCAAAAA GCGACATTTG ATATACACAA	2820
	TTCTCCGCTT TCAGAAGCAG CAGTAGTTGG TTTTGAATAC GGCTATAATG TGGAAAACAA	2880
	AAAAAGCTTC AATATTTGGG AAGCACAATA TGGTGATTTT GCAAATATGT CACAAATGAT	2940
35	TTTTGACAAC TTCTTATTCA GTTCTCGCTC AAAATGGGGA GAACGTTTCAG GATTAACATT	3000
	ATTCTTACCT CATGCATATG AGGGTCAAGG GCCTGAACAT TCATCAGCAA GATTAGAGCG	3060
40	ATTTTTACAA TTAGCTGCTG AAAATAATTG CACAGTTGTC AACTTATCTA GTTCAAGTAA	3120
	TTATTTCAC TTATTGCGTG CACAAGCGGC TAGTTTATAGT TCTGAACAAA TGCGACCATT	3180
	GGTTGTTATG TCACCAAAAA GCTTACTGAG AAATAAAACA GTTGCAAAAC CAATTGATGA	3240
45	ATTTACTTCT GGTGGATTG AGCCAATTTT GACAGAATCA TATCAAGCGG ATAAGGTTAC	3300
	AAAAGTTATT TTGGCAACTG GTAAATGTT CATTGATTTA AAAGAAGCAT TAGCTAAAAA	3360
50	TCCAGACGAA TCAGTATTAC TCGTTGCGAT TGAAAGATTG TATCCATTCC CAGAGGAAGA	3420
	GATTGAAGCA TTAGTAGCAC AATTGCCAAA CCTTGAAGAA GTGTCATGGG TACAAGAAGA	3480
	ACCTAAAAAT CAAGGTGCAT GGTTATATGT CTATCCATAT GTTAAAGTGC TAGTTGCAGA	3540

55

AGAAATTCAT AAACCTTGTTT AAAATAAAAT TATAGAAAAT GCATTAAAAA ATAAC TAGGG 3660
 GGAAATAAGT CATGCCAGAG GTTAAAGTTC CAGAATTAGC AGAATCTATT ACAGAAGGTA 3720
 5 CCATTGCAGA ATGGTTGAAA AACGTAGGGG ATAGCGTAGA AAAAGGTGAA GCTATTCTTG 3780
 AATTAGAAAC TGATAAAGTT AATGTCGAAG TTGTATCTGA AGAAGCAGGT GTATTATCTG 3840
 AACAACTTGC AAGTGAAGGC GACACTGTAG AAGTTGGACA TCGCAATTGCT ATCATCGGCG 3900
 10 AAGGTAGTGG CAATGCTTCT AAAGAAAATA GTAACGACAA TACTCCACAA CAAATGAAG 3960
 AAACAAATAA TAAAAAAGAA GAAACAACAA ATAATTCGGT AGATAAAGCT GAAGTAAATC 4020
 AAGCAAATGA TGACAATCAG CAACGTATTA ATGCTACGCC TTCTGCGCGT CGATATGCTC 4080
 15 GTGAAAATGG TGTGAATCTT GCTGAAGTAA GTCCGAAAAC AAATGATGTG GTTCGTAAAG 4140
 AAGATATTGA TAAGAAACAA CAGGCACCGG CATCAACACA AACAAACAA CAAGCATCTG 4200
 20 CAAAAGAAGA GAAAAAATAC AATCAATATC CTACAAAACC AGTGATTCGT GAAAAATGT 4260
 CACGTAGAAA GAAACAGCT GCCAAAAAAT TATTAGAGGT ATCTAATAAT ACAGCTATGT 4320
 TAACAACATT TAACGAATG ACATGACAAA TGTTATGGAA TTGCGTAAAC GTAAGAAAGA 4380
 25 ACAATTTATG AAAGATCATG ATGGTACTAA ATTAGGATTT ATGTCATTCT TTAATAAAGC 4440
 TTCTGTAGCA GCTTTGAAA AGTATCCAGA AGTTAATGCA GAAATCGACG GCGACGACAT 4500
 GATTACGAAA CAATATTATG ATATTGGTGT AGCTGTTTCT ACAGATGATG GATTATTAGT 4560
 30 ACCATTTGTA AGAGATTGTG ATAAAAAGAA TTTTGCAGAA ATCGAAGCAG AAATTGCTAA 4620
 TTTAGCAGTT AAAGCAGCAG AGAAAAAACT TGGCTTAGAT GATATGGTTA ATGGTTCATT 4680
 TACGATTACA AATGGCGGTA TTTTGGATC AATGATGAGT ACGCCAATTA TCAATGGTAA 4740
 35 TCAAGCTGCA ATCTTAGGCA TGCATTCAAT TATTACAAGA CCAATTGCGA TTGATCAAGA 4800
 TACAATCGAA AATCGTCCAA TGATGTATAT TGCATTAAGC TATGATCATA GAATTATT 4858

40 (2) INFORMATION FOR SEQ ID NO: 380:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2222 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 380:

ATCAGTCACA CGGTAGGCAT ATAAATGAG TCGTTTCTAC AACATTTTAA AACAGTTCAT 60
 TCAATATTAT TTTTATCTAA TAATATATT GGGAGGATTA TACCTTTATA CACACCATGC 120

	ATGCTATTTA GCTAAAGCTA AAAGACCAGA CACTATGCAT ATTTCAACTG GAAATATGTG	240
	GCGATACTTA GTTGCAATTA TTGCCTGTAT GATTTGGTAC CTTAATAAAG CGCATGTAAG	300
5	TATCATCGGT ATAATTATTG GTTTAATGAT TTCATATGTT GTAGTTATCA TACGTCCTTT	360
	ACTAAAGGTG AGCAAATAAA TTAAGAAAGA GGTGAGATTA TGGATCACAA ATCCCCGCTC	420
	GTGAGTTGGA ATTTATTCGG TTTTGaTATC GTTTTCAATT TATCAAGTAT ATTGATGATA	480
10	CTTGTTACGG cGTTTCTTGT TTTTCTACTT GcTATCATTT GTACGCGTAA TTTGAAAAAA	540
	AGACCAACTG GCAAACAAAA TTTTCGTTGAA TGGATTTTTG ATTTCTGTGAG GGAATCATT	600
15	GAAGGTAACA TGGCTTGGA AAAAGGTGGT CAATTCCACT TCTTAGCAGT AACGCTGaTT	660
	CTGTACATTT TTATAGCTAA TATGTTAGGT CTTCCGTTTT CTATAGTAAC GAAAGATCAC	720
	ACATTGTGGT GGAAATCACC GACAGCnGAT GCAACAGTGA CTTTAACGTT GTCTACAACG	780
20	ATAATACTGT TAACTCACTT TTATGGAATT AAAATGCGTG GTACGAAACA ATATCTTAAA	840
	GGTTATGTAC AGCCGTTTTG GCCATTGGCA ATTATTAATG TTTTGAAGA GTTCACTTCA	900
	ACATTAACGC TTGGTCTGCG TTTGTACGGT AACATATTTG CAGGTGAGAT ACTATTAACA	960
25	TTACTTGCTG GCTTATTCTT TAACGAACCA GCATGGGGTT GGATTATTAG TATCCCAGGA	1020
	TTAATTGTTT GGCAAGCATT TTCAATATTT GTAGGAACAA TCCAAGCATA TATCTTTATT	1080
	ATGCTTTTGA TGGTTTATAT GTCACATAAA GTGGCAGATG AACACTAAAA ATTTCAATAA	1140
30	TTATATACAA TCACAGGAGG AAATTAAATT ATGAATTTAA TCGCAGCAGC AATCGCAATT	1200
	GGTTTATCAG CATTAGGAGC AGGTATCGGT AACGGTTTAA TCGTTTCAAG AACAGTTGAA	1260
35	GGTGTAGCAC GTCAACCAGA AGCACGTGGT CAATTAATGG GTATCATGTT CATTGGTGTA	1320
	GGTTTAGTTG AGGCATTACC TATCATCGGT GTAGTAATTG CATTATGAC ATTTGCTGGA	1380
	TAATTAACAG ATAAAAGAGG TCGGGACAAA GCGCATAGGA CATAATTCAT GATGCATATA	1440
40	TAGTAATATC TTTGAACTTT ATTAAATAGT TGAGATATGA ACGCACCATG CCTATCGCAT	1500
	AAATTCAGTA GGTCTTAACC TCGTCGTTTT TTTCTATATA AACTAGCGA TTATTTTAAT	1560
	GAAAGGAGTG TCATGAACCC GTGACTGAAA CAGCTAACTT ATTCGTTCTT GGTGCAGCTG	1620
45	GAGGCGTTGA GTGGGGTACT GTGATTGTAC AGGTCCTAAC TTTTCATCGTG TTAcTTGCGT	1680
	TACTTAAAAA GTTCGCATGG GGTCCATTGA AAGATGTAAT GGATAAACGT GAAAGAGATA	1740
	TTAACAGAGA TATCGATGAC GCAGAACAAG CTAAGTTAAA TGCACAGAAA CTTGAAGAAG	1800
50	AAAATAAACA AAAACTTAAA GAAACACAAG AAGAAGTTCA AAAGATTTTA GAAGATGCTA	1860
	AGGTTCAAGC ACGTCAACAG CAAGAACAAA TTATTCaTGA AGCAAACGTA CGTGCAAACG	1920

55

TTAATAATCa AGTATCTGAA CTATCAGTGT TAATTGCTTC TAAAGTTCTT AGAAAAGAAA 2040
 TTTCTGAACA AGACCAAAAA GCATTGGTTG ACAAGTATCT AAAAGAGGCA GGCGATAAAT 2100
 5 AATGGTAAAA GTAGCTAACA AGTATgCTAA AGCATTATTT GACGTGTCAT TAGATACAAA 2160
 TAATTTAGAG ACTATTAATG AAGAATTAAC AGTTATAAAT GAAGCAGTAA AAGATAAAAT 2220
 TG 2222

10

(2) INFORMATION FOR SEQ ID NO: 381:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1143 base pairs

(B) TYPE: nucleic acid

15

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

TAACAATTCC TaTATTCATG TTAAATACGA AACACTACAT TTACATTGTA ATTCACTATC 60
 TTTTGAAGTA ATAAAGTGAT TTGTTCAATC GATAGCTCAT TGCTTGTGTC GATTGTAACA 120
 25 ATTATATGCA AGTTTTTCAGG ATTAACACCT AATCTTTGAA TGATTTGTTT AATAGTATAA 180
 TAATCCATCC AATAAAAATT ACTTCCTTTA ATATAAATGT TTTTAGGTTG ATACATTTGA 240
 CTCCTTTTTG TAGGCTCAAA AGGTATATCA ATCTCGCGCA TACTTGAAGA ACTTTGATTA 300
 30 GTATCATCAA ATAATTCAAT TATATTTTTA TCAATTTCAA CTAATTGGGA ATGGTTAATT 360
 GAATGAGATG TTGGACTATA TCTTTTTCTA ATTAATTTAG GAGTGTTACC ATACGTTTCT 420
 TTAAAGAGGT GTATAAACG AGAATAATGA TTAAACCAT GACTACTTGC GATTCTTCA 480
 35 ATAGGCTTCT TAGrAGTTAA AATATCAATT AAGCAATGCT CCAGTCTAAT ATGATTTAAA 540
 TATTGAATAA AATTACTATA AGGTGTCGCT TTAAACATGT CACTTAAAGC TTTGTTTGTA 600
 40 ATACTAACTT GATTAATGAC ATCTTTCCTA TTTATCTTTT TATGGTGGTT GTTTGTTAAA 660
 TAATCGTGCA CTTCTTCGGC TACTAAATGA CGACTACCAT CGTATGTATT TAATGACATT 720
 AATTCAACAC ACATGTAATT AATAATCTTA TCATTAGCAT TATAAGACTG TTGTTTAATT 780
 45 TTGCTGTAGA TTAAATACTT AATCAAGATT CTAAAAGTGC TAGCAACCTC ATTTGTTAAT 840
 ATCTTGCCAC GCATCATATA GTTGCATTG ATAACTTCA AAAATGTCTT TGCACTTACT 900
 TCAACAACAC AACATACACT ATCATCATTA CCATCAATTT GATACAAATC ATTCAACATA 960
 50 ATAATGGTGA CATCATTTCT TTTTACATCA TATTGTTGTA GATTGATTGA AACTACACCT 1020
 GTACCTTGTA GCCAATATAT AATTTTAAAG TTCGCTTCCT TCACACTACT CATTTCATA 1080

55

ATT

1143

(2) INFORMATION FOR SEQ ID NO: 382:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 506 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 382:

15 GGACCACCTT TTTTACCTGT ACGACTAAAG CCAACTGTTT TAAAGGCACC AACAGCATGC 60
 CACACTTGAA TAATTTCTTG TGATGGTCTA AAGCGAACCG TATAAATTAA TGGATGGAAA 120
 TCATCAACAA AAATATAATC GGCCTTACCA AGTAAATATG GCAATCTAAA CTTGTCTCTC 180
 20 CATTTGCGTC TATCCGTAAT ATTCTCCTTA AATACCGTTT TAATATCATA ATCAAAATCT 240
 ACTTTTTGGC GTAGTAACTC ATCATATACA TACTTGAAAT TCCCTGATAA ATTCGGACGC 300
 GAATCTGATG TGAATAATAT TGTTTTGCTT CTTTTAATAT GTAGTAACTT TGTAATATTA 360
 25 AAAATAGCTT TAAATAAGAA ACTTCTACTT TCAAATGAAG CTTTATGACC TTGTTTatGA 420
 AGCCAGTGTG cACTTgTCGC AATGaCCCCCT GaTTTCyCyT GagGtAAGGk GaTTTCmATA 480
 TCAAATACAA ATTCGTTAAC GTCAC 506

30

(2) INFORMATION FOR SEQ ID NO: 383:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 421 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

40 AGCATCTGCA AtGAGTcTCT AATACAACGA TACGTTTTGC ATCTTTAGGT ACTTTTACTG 60
 TACCATTTTC ATCTTTTACC GAAATAGTAT CTTTAGTTGA TGATTCTTTT TTAATTGAAT 120
 45 TATCCGTATT ACCACAAGCT GCAACTAAAA GTAAGGCAAC TATTAATCCC AATATACTAA 180
 AAGTTTTTAG ACCTCTCATC GTTCCACTCC TTAATATGTA TAACTTCATT TATTATTTTA 240
 TTGATAACAA TTATCATTGT CAAGTAGCGT TCAATCTTTT TTATATTTCT AAAATGTATG 300
 50 ACTATATATT TCCTCTAATA ATTATGACTA CAATTAGCAC ATTTCTTAG ACAAATACT 360
 GATAATGTAT CATTGCTATA TCATCTTTGC ATTAATACAA TTGACACCAT TTAGCATGAC 420

55

(2) INFORMATION FOR SEQ ID NO: 384:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 862 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

TGCACTTGAT ACTTTAGCAC ATGAAATTGC TAATCGATTA GTTGCAAATG ACAAGAATGA 60
 AGCAACTTTG GAAATGACTA ATAAATGGC AACGATTCGT TTTACAGAAC CTACGCTGAT 120
 TGCATTAGCA GGGGGTAATG TCAAAGCTTA CACTGAGCAT ATGACTATAT CTCCATATAA 180
 ATTGTATTTG TTAGATAAAG GCGATGTTTT AAAGTTTAGA GAAACAAGTT ATACATCGCG 240
 AGTGTATTTA GCTGTgGGAG GCGGATTTGA ATTAGATGCA TGGTTAGGAT CTAACCTAAC 300
 CGACTTTAAT GTAAAAATTG GTGGTTTTAA AGGTAGAACA TTACAAGATG GCGATGAAAT 360
 AAAGCTTAAG AGAGATTATA CAGCTCGTCA TCATAAGTTA TTTGAAAACC TTGCTCACAC 420
 GAAACAAACA GATTGGGGTA TTGATGGATA CGCCTTGTC TTTAATTATA TGTCTGATGT 480
 ATTCATGTC GTTAAAAATA AAGGTACGGA AGATTTTAAA GAAGATGCCA TTCAAAGATT 540
 TGTGAAACAT GATTATAAAG TAACGAGCAA AGCAAATCGC ATGGGGATGA TGCTTGAAGG 600
 TGAAAAAATC AAAGCTTTTT ATGAAGATAT GCCACCGTAT CAGACTGTCA AAAAAGGAAC 660
 GATACAAATT AAGCGTGATG GCACACCTAT TATCCTATTA AATGATCATT ATACGCTAGG 720
 TAGCTACCCG CAAATCGGTA CAATCGCAAG TTATCATTTA ACGAAATTAG CACAAAAACC 780
 GCAAGGATCA CGTTTGAAAT TTCAATTTAT AGATATTTTA ACGGCTGAAA AGAACCTTGT 840
 TAAgTATAGT AACTGGTTAA AC 862

(2) INFORMATION FOR SEQ ID NO: 385:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1027 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:

AAATGAGTTG ATGATGGCTC CTATTCCATC TATCTCTAGG TATGACAGAA SATAAACGGT 60
 CTTCTACCTG TCTsACATTA TCTTTCCAAC GATTAATACC TAAGCGTTTA GAAACACGTT 120

CAGTTTTACG TCCTACACCT GCTAAACTTT CTAATTCCTT ATGTGTTTGT GGTATTTCTC 240
 CATTAAATTG ATCAATCAAA GATTGACAAA GTTCTTAAAT ATTCTTAGCT TTGTTACGAT 300
 5 ACAGACCGAT AGAACGAATA TCATTCTATA GTTCTTCATC ACTGACTGCC AAATAATCTT 360
 CAGGCGTTTT GTATTTTTTA AACAGCTCAG TTGTTACTCT ATTTACTAGA ACGTCTGTAC 420
 ATTGCGCTGA CAATAATACA GCAATAGTTA ATTGGAACGG ATTATCATGT TTTAATTCAC 480
 10 ATTCTGCATC CGGAAACATA TTTGCTATAA CATCAATCAT TTCTAATGCT TTTTCTTAC 540
 TTACCATCAA GGTTCCTCCC GTTTAACCAA TCAAATTTAG GTACCGTTTT AACTGTGTGC 600
 GTCATTTTCG GTTTATTGAA TTTTCTCTT ATTTTCTAG AATCGTCAAT TGTTTTGACA 660
 15 TTGTTTTTCT TCCAATTAAG TAAAATACGA TCTATATATT TAAAGCTAAG TTTATTCAAA 720
 CTATTCGCCT CGTCTAATGC CGCTTGATA ATTGCAGTAT CGTGTTTATC AACATCAATC 780
 20 CATTGATTTA ACGTTTCTAT TTCATATGGA GATAACGGCC TTGCAAATGT ATCCTCTAAA 840
 ACTCTAAATA ATTGTTTAAA TTTTCTTTA CTATTTAGCT CTTTCGTTTC CATACTTTGT 900
 TGCTTCAATA TATGACTTAA TTTTTCGAAA AAAGGATCTA GATTCATATA TTCGGKAAAT 960
 25 CTACCTTCyT CATCyTTTTG aACTkGtAAT tCTAGCAATT CACGTgTATC AAATTTTGGA 1020
 TACCATT 1027

(2) INFORMATION FOR SEQ ID NO: 386:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1006 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

40 AAGGnTTGGA GGAATTAAT nGATGGCAAA TCCCAGaTAA AGTACACaAC GCATGGAATG 60
 TGATGCATTT AATATTTCCA GTAGTTAGTA CGACGTTTGC AAGCTTTAAA TCTATGTATG 120
 GGGGCATACC AAAAGATTTT ATAGACTACT TATTTATTGA TGAAGCAGGA CAAGCAATAC 180
 45 CTCAAGCAGC TGTGGGAGCA TTATATCGTT CAAAAAAGT TGTAGCTGTA GGTGATCCGA 240
 TTCAAATAGA ACCGGTTGTG ACTTTAGAAA GTCATTTAAT TGATAACATT CGTAAAAATT 300
 ATCATGTTCC GGAATATCTA GTTCTAAAG AAGCTTCTGT GCAGTCTGTT GCAGACAACG 360
 50 CCAATCAATA TGGTTTTTGG AAATCTGATG CTA CTGATAG TAATCAAAAA ACCTGGATAG 420
 GCATACCTTT ATGGGTGCAC AGACGATGTT TAAAACCTAT GTTCACGATa GCTAaCCAAa 480

55

GGTATGACGT TAAAGGAAaC GCAGTTCAAA AACAAATTTGT GAAAGAGCAT GGTGAAAAaG 600
 TAGTGGGATT ATTAGCTGAT GATTGGATTG AAGCAATTAA GGAAGGTAAA AATGAACCGA 660
 5 GCTCATTGT AATATCGCCT TTTTCAGCAG TACAGCAACA GATTAAACGT ATGTTAAAGC 720
 AACAACTACC GACTAGAATT GATATTGAAC GTACAAAAAT TAATCAATGG GTCGATAAAT 780
 10 CCATTGGTAC TGTTCACTACT TTTCAAGGTA AAGAGGCTCA GAAGGTGTAT TTTGTAATAG 840
 GTACTGATAA TACCCAAGAT GGTGCTGTGA ACTGGTCATG CGAAAAACCA AACTTGTTAA 900
 ACGTTGCACT GACAAGAGCT AAGAAAGAGT TTA TGTAATT GCGGACATGC AAAGAATACA 960
 15 GATGAaACCA TTTTATGAGA CGATTTTTan AGnAAGAAAT GTAAAA 1006

(2) INFORMATION FOR SEQ ID NO: 387:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 662 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387:

CGTTTCATCA AGTnATCGAT CCTAATTTAC AAAATATTCC GGTTAGACTT GAAGAAGGGC 60
 GTAAAAATAAG AAAAGCCTTT AAACCAACTT CAAAAGATAG CGTTATATTA TCAGCAGATT 120
 30 ATTCTCAAAT TGAATTGCGT GTATTAGCAC ACATTACACA AGATGAGAGT ATGAAAGAAG 180
 CATTTATCAA CGGCGATGAT ATTCATACAG CAACTGCTAT GAAAGTATTT GGTGTAGAAG 240
 35 CTGATCAAGT CGATAGTTTA ATGCGTCGTC AAGCAAAAGC AGTTAACTTT GGAATTGTTT 300
 ATGGGATAAG TGATTATGGT TTAAGTCAAA GTTTAGGTAT TACTCGTAAA AAAGCAAAAG 360
 CATTCATTGA TGATTATTTA GCTAGTTTCC CAGGTGTAAA ACAATATATG TCTGATATTG 420
 40 TAAAAGATGC CAAAGCTTTA GGTTACGTGG AAACATTGCT ACATCGTCGA CGCTATATTC 480
 CTGATATTAC GAGTCGTAAC TTTAATTTAC GCGGCTTTGC TGAACGTAAT GCTATGAATA 540
 CGCCAATACA GGGCAGTGCT GCAGATATCA TTAACTGGC AATGGTTAAA TTTGCTCAAA 600
 45 AAATGAAGA GACAACATAT CAAGCTAAAC TATTATTACA AGTACACGAT GAATTAATTT 660
 TT 662

(2) INFORMATION FOR SEQ ID NO: 388:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 669 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

5 TCATCCAAAT TTTGGAAATT CCACATTTTA CATATCGTAA TTTTITAGGA AACTAGTGAA 60
 TATAACAAAT CCCTCCTCTC ATTTTAAAAA TAGATATATC ACTTCCCCAC TTTTACTTAA 120
 CTAAACTGCA ACGGTTCTTA ATACCAAAT CCTGCCCTCT ATTTTATCA ATTCAAGCAT 180
 10 ACTTATTGAA AAATGTTAAC GTTTTCTTGA TAATCATTGT AAGCGCATTI ATTTTATAAA 240
 CTAACGTTTG AAATATACTA CAGGAGTGAC ACGTAATGAC TCAAATTACT GAAAAAGAAT 300
 TAAAAAGAA GTATTTAGAT TTAATATCCC AAAATTTTGA TACTCCAGAA AAACCTTGCAA 360
 15 CTGAAATTAT CAATTTAGAA TCAATTTTAG AATTACCTAA AGGTACGGAA CATTCGTCA 420
 GCGATTTACA TGGTGAATAT GAAGCTTTCC AACACGTATT ACGCAACGGT TCTGGGAACG 480
 TCGGAGCGAA AATCAATGAT ATTTTCAAAG AGAGACTTTC AACTAAGGAG CTTAATGACT 540
 20 TAACTGCTCT TGTCTACTAT CCAGAAGrCm AATTtAAAAT TgATTAAAAG TGATTTCCAA 600
 AATTgCGGtC mActTAATGt CyGGtATATC ACaACmATCG aACATTTAAT TGAGTTAATT 660
 25 AAATATTGT 669

(2) INFORMATION FOR SEQ ID NO: 389:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 1249 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 389:

35 CACATGGCTG TTAGAGATAT GAATGGCCAT GCGTTACCTT TAACAAAAGA TGGCAATTTT 60
 TATCAAACGA ATGTAGATGC AAATGGTGTT AATCATGGTG GTAGTGAAAT GGTGCAAAAT 120
 40 AAAACAGGTC ATATGAGTCA ACAArGCCAT ATGAATCAGA ACACACATGA ACCAACAGCC 180
 ACACATGCAA CAAGGTCATA TGCAATCATC AAACCATCAA ATGATGAGTC CAAAAGCAAA 240
 45 TATGCATTCA TCAAATCATC AAATGAACCA AAGTAACAAA AAAGTTTTAC CAGCTGCTGG 300
 TGAAAGTATG ACATCAAGTA TTCTTACTGC AAGTATTGCC GCACTACTAT TAGTATCTGG 360
 GTTATTCTTA GCATTTAGAC GACGTTCAAC AAATAAATAA ACATAATACG ATTAATAATA 420
 50 GAAAAATCGT GTGATTATCT GaGGGAGCCT AGGACATAAA TCAATGTCCT AGGCtChCTA 480
 AtGTTATATT GGCAGTAGTT GACTGAATGA AATTGCGCTT GTAACAAGCT TTTCCATTTC 540

55

CGATTGTCTT ATACGTGTCA GTGTTAATTC AGATATTTCC TGTGGAATAT ACCACTTATT 660
 AATCATAATT GGATAAGGTG TTTGTGCGTA CAGTGTTCAT ATAATCAGCC AACAAATGTGT 720
 5 ATCACCATCA AACACGTGAC TATGATTTTt GAAGTGGGGC GCTTTGGTAA TAGACATTTT 780
 TAAATCTGAT TGATATGCAT TGCTATAAAT CGTTTGCTCA ACGAATGTCT TCATGTCGTC 840
 TTCGTTTTGT GTATTCACCTT TAAATGTGTC AATGACATTT AACGGTATAA AGGTAAAGCA 900
 10 AAATGCATCA GCTTGCTTAG AATGATTGTC CTTTTTTTGA TAATAGCGTT CCATTGCAAT 960
 GACGGCAGAA GGATGGTTTG CAAACAAATG ATTTGTATAT TCACTTTCTA AATCAACACG 1020
 ATAATTAATT GATGACATAG ATACGCGAGC TAGCAATATT TGATCAAGTG GATGCTTAAA 1080
 15 TTGATCCATA CTTGAAGCGT GTTGGGCATT TGTTTGTGGA ATAACAAAGT GTCCCTTCCC 1140
 TCTTGACTC TCTACGATGC CATCTCGGC TAACAATTT: ATAGCTTGGC GCAAAGTCAT 1200
 20 ACGACTGGAC ATCAAAGCGC GCACAAAGTT CCTTTTCAGT AGGTAATGC 1249

(2) INFORMATION FOR SEQ ID NO: 390:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1788 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

AATGATGATT CATTCAAAGA AATTAACGCT CGGTATATGC TTGGTATTAC TCATTATATT 60
 GATTGTAGGT TATGTCATTA TGACAAAAAC AAATGGTCGA AACGCCCCAA TTAAGACAC 120
 35 ATTTAATCAA ACATTAAAT TATATCCAAC CAAAAATCTC GATGATTTT ACGATAAAGA 180
 AGGCTTTTGA GATCAAGAAT TTA AAAAGGG TGATAAAGGT ACTTGGATAG TTAATTCTGA 240
 40 AATGGTAATC GAGCCAAAAG GTAAGGATAT GGAAACGAGA GGAATGGTGC TCTATATCAA 300
 TCGCAATACT AGAACCACAA AAGGGTATTA TTTTATAAGT GAAATGACAG ATGACAGTAA 360
 CGGCAGACCA AAGGATGATG AAAAAAGGTA TCCGGTAAAA ATGGAACATA ATAAAATCAT 420
 45 ACCAACGAAG CCACTACCGA ATGACAAGTT AAAAAAGAG ATTGAAAAC TTAAGTTCTT 480
 TGTACAATAT GGCAACTTTA AAGATATTAA TGATTATAAA GATGGTGATA TTTTATATAA 540
 TCCTAATGTA CCAAGTTATT CGGCAAAATA TCAATTGAAT AATGATGATT ATAATGTCCA 600
 50 ACAGTTaAGa AAAAGATATG ATATTCCAAC CAAACAAGCG CCGAAACTAT TATTGAAAGG 660
 CGATGGAGAT TTA AAAGGTT CATCCGTAGT TcTAGAAGTC TTGAATTTAC CTTTGTGCGAA 720

	AGGTATGAGT CAAACTGAAT ATCAAATAAA ACCTGGTAAT ATAACAAGTA ACTCTGAAGA	840
	AACAAGTTTCG ATATCTAAAG TGAGCTGTGA AATATAGGTA GCCATTTCAA AAAATTTAAA	900
5	GGTGAATTTG ATAATGTAGC TCAAGGAGAT TGGGTTAAAA AGGCGAAGAA TGAAGTGGAT	960
	GATATTAGTA AGAAATTAAA AAATATTCAA AGAACGGAAG TTTAATAGCT TATATGATTC	1020
	TTGGAGCTAA GACAGCATGC GTTCATTCAT GCCATTATTA ATATAAGCAC CGCAACAAAA	1080
10	AAGCTTCTAA TGTGATACAG GAACCTCATA TTCCGTATCA TGTTAGAAGC TTTTAATGTC	1140
	TAAAGAACAT CTACATTTTA TCATATTTTC TGACTTATTA AACTTTTATA TAATTAAATA	1200
	TTTCTTAATT TTCCAAAATA GTGATAAATT TGTGAAATAC ATCACAATC CCTTTATTTA	1260
15	TTTGGAAATT CATGTAATAT TAGACTTGTA AGAAGTTAAT AAATAGAGAG AGACGAGAGA	1320
	GTTTATATAA ATACTATATA AACATTGGAG TGATGATTAT GAGAAAAGAG ATTGAAGCGC	1380
20	TTATTTTCTC AGACGTAATA GCTATGATAT TTACGTGAAC ACTGGTGTA ATCAAGGATT	1440
	AATTGGTGAC ATCAAAGATG GTTACCTAAC TATTGATTCT ATGCCTTACA TTGATGCTGA	1500
	GCGTTGTAT CACTTTGCTA TGGAAAGTAA ATCGTTAGTC ACTAACTAGT TCTTATTGCC	1560
25	AATGATTACT ACCCCTAGTC GCGGCAATT GAAGTGTGAT TGATGTAAT TGCCCTCGTT	1620
	GGTGAGCAAT TGAGGCGAGA CCCCTTTAAT TAAGTAAACC CTAACCTCCC ACAAACTGG	1680
	AACGATACTA AAAGCCACGT CCTATATTGG ATGTGGCTTT AGTCAKACTT ATATTATTTT	1740
30	tAAAACGATT ACCTACAAGA TTTACATATA AAATTCTATC ATGnCTGC	1788

(2) INFORMATION FOR SEQ ID NO: 391:

(i) SEQUENCE CHARACTERISTICS:

- | | |
|----|-----------------------------|
| 35 | (A) LENGTH: 2407 base pairs |
| | (B) TYPE: nucleic acid |
| | (C) STRANDEDNESS: double |
| | (D) TOPOLOGY: linear |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

	GCAAGTTTAG TTAACAGAT ATTAAAACCA GATCAATATA CAGATGCAA TTCTCTATTA	60
45	GAGATTAGTT TGCAACAGG TATGTTTATT GCAGGTGGAT TATCAGGAAT ATTATATAAA	120
	ATAAATGGAT TCACTCTAAT TATAGCGATG ACTATAATGA TGTTTCTAAT CAGCATTTTT	180
	ATGTTATTTA GATTGCATGT AGATAAACCA ACTCATTCAG AGGAAGAATC AACAAATAGC	240
50	TTATTGCAAG AATATTTATT AGGATGGAAA TTTTAAAAG ACAACATGAT GATATTTATT	300
	TTTGGAGTTA TTTCAATCAT ACCAATGGTG TTTACAATGA TCTTTAACAT ATCATTACCA	360

55

	TATGGCATTG GAGGATTATG TGCAGGTTTA ATTTTCAGCTA TTCTTTTCGAA GAAAATCTCA	480
	ACTAAAGTAT TGATATTTTT GTTATATTTT ATATTAGTCA TAAATTCAGC ACTATTTATT	540
5	TGGATAAACT CAGCATTTTA CTTATTCATA GGATCATTTA TACTAGGCTA CTCAATTTCA	600
	TCAATCAGAA TTTATATGAA TACAGCTATA ATGAACACTG TTTTCAGATAA ATATGTCGGT	660
	CGCTCATTTA CGATATGGAC GTCAATTTCA TTGTTACTGC AAAGTTTAAT TGCTCCATnT	720
10	TTAGGAAGAT GGATTAATGA AATTAATGAT AAATTCGGTT TCTATATTAT ACTCATTTTA	780
	TCCTTACTCA TATTTGTaCA CTGCTGCTTG TTAACAAAAC AGACAAAATA AAATATGCAC	840
	ATAAGAAGA GTGACCGTCA CTCTTCTTTA ACAAGCGACC ATTTATCGaT GGGCTTAGTT	900
15	CTCTCTGCAC CCACACTTCA CTACTTCACT TTTTCAAATC ATTTTTTATG GTCTTAAATA	960
	AATCAGTGAG ATTTGTTGCT TCGGTAAAGT CTAGAATTAA TATCATTTCT TTAGAACCTG	1020
20	GATATGGCGA AACTAATGTA TTATCTTGCA ATTTCTGCTG GGCACTTTTA GTCGCCTTGA	1080
	CCAATAATCT ATTATCATAC AAACCACCTA TAACCACGCC ATCATAATAA ATAATATATT	1140
	CTCCCATCAT CTTTCTTGTC TTAACCGCGT TTGAATTCAC ATGATTTAAA AATAAATCAT	1200
25	GTACATCTTT CTTAGTCGCC ATTGTAATCG-CTCCTTCAGT TTTATGTTTA ATCACATTGG	1260
	TATTAATGAT TCATTTCTGT TTGCTCTTAA TTTTATCTAT AATTATATGC GTAGTTAAAA	1320
	TCAAACCTAT GGAAAAGAAA ATAATGATAA CAATGTTAAG AAATATAGTT ATAAAATTAT	1380
30	AGTTTGAAG GTATGCGAAT AGCAGAAAAA TAGGTATCGC AAAAAATAAA TCCCACCAAC	1440
	CTAAACTTTT TAAAGAATGC TTTAAACCTT CCATAATATC ACCTTTTATAA ATTTGTCTTT	1500
	GTTATAAGAT AACTAAAAA TCGCTTTACT GTAAAAGTAG CCAAAGAAAA TTCTGAATCA	1560
35	TATTCATAAG TAGTGTATCA TTAATAATGA ACAATTTAAT ACTATAATCC TTGATCTTTG	1620
	TATTGATCAA CTTACCACAA CATTATTTTT AGACTACTCT TAGACTTCCC TTTCAAATGG	1680
	TTGCATCTAT TGAAATTCCT TTTGTATAAG TTAGGCTTTT GTGGTAATAT CATCATGCAT	1740
40	AAAAAATCGA GATACTAATT ATAAAGAGGG TATAAATATA TTATGAAAGA AAATTTTTGG	1800
	AGTGAATTAC CACGTCCATT TTTTATTTTG GCGCCAATGG AAGACGTTAC AGATATCGTC	1860
45	TTTCGACACG TTGTAAGTGA AGCAGCTAGA CCGGATGTGT TTTTCACTGA ATTTACAAAT	1920
	ACTGAAAGCT TTTGCCACCC TGAAGGCATA CATAGTGTGC GCGGACGCTT AACTTTTACT	1980
	GAAGATGAAC AGCCGATGGT CGCTCATATA TGGGGAGATA AGCCAGAACA GTTCCGTGAA	2040
50	ACGAGTATTC AATTAGCTAA AATGGGCTTT AAAGGCATAG ACTTAAATAT GGGATGTCCT	2100
	GTAGCAAATG TTGCTAAAAA GGGTAAGGGT TCCGGCTTAA TCTTAAGACC TGACGTTGCT	2160

55

GGCTACTATG AAATCGATGA ATGGAAAGAT TGGTTGAAGC ACGTCTTCGA ACAAGACATT 2280
 GCCAATTTAT CTATTCATCT TCGTACACGT AAAGAAATGA GTAAAGTAGA TGCACATTGG 2340
 5 GAATTAATCG AAGCTATTAA AAATTTACGT GACGAAATTG CACCAAATAC ATTGTTAACA 2400
 ATTAACG 2407

(2) INFORMATION FOR SEQ ID NO: 392:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2424 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 392:

20 ATGATGnATG GcncCGCCAA GAAGTTGAAC CAGTCTATTG GCTAGAATTG CTCAAAAAGC 60
 GAGAcATGTG GTATTCaTAT GTTaGTAGCt ACGCAAAGAC CATCTGTCAA TGTAATTACA 120
 GGTtTAATTA AAGCCAACAT ACCAACAAGA ATTGCATTTA TGGTATCATC aAGTGTAGAT 180
 25 TCGAGAACGA TATTAGACAG TGGTGGAGCA GAACGCTTGT TAGGATATGG CGATATGTTA 240
 TATCTTGGTA GCGGTATGAA TAAACCGATT AGAGTTCAAG GTACATTTGT TTCTGATGAC 300
 GAAATTGATG ATGTTGTTGA TTTTATCAAA CAACAAAGAG AACCGGACTA TCTATTTGAA 360
 30 GAAAAAGAAT TGTGAAAAA AACACAAACA CAATCACAAG ATGaATTATT TGATGaTGT 420
 TGTGCATTTA TGGTTAATGA AGGACATATT TCAACATCAT TAATCCAAAG ACATTTCCAA 480
 ATTGGCTATA ATAGAGCAGC AAGAATTATC GATCAATTAG AGCAACTCGG TTATGTTTCG 540
 35 AGTGCTAATG GTTCAAAACC AAGGGATGTT TATGTTACGG AAGCAGATTT AAATAAAGAA 600
 TAATTATGAG TAAGGAGTTT TATATAATGA CACACTATCA TTTTGTCGGA ATTAAAGGTT 660
 CTGGCATGAG TTCATTAGCA CAAATCATGC ATGATTTAGG ACATGAAGTT CAAGGATCGG 720
 40 ATATTGAGAA CTACGTATTT ACAGAAGTTG CTCTTAGAAA TAAGGGGATA AAAATATTAC 780
 CATTTGATGC TAATAACATA AAAGAAGATA TGGTAGTtAT ACAAGGTAAT GCATTGCGGA 840
 45 GTAGCCaTGA AGAAAtAGTA CGTGCACATC AATTGaAATT AGATGTTGTA AGTTATAATG 900
 ATTTTTTAGG ACAGATTATT GATCAATATA CTTCAAGTAGC TGTAAGTGGT GCACATGGTA 960
 AAACTTCTAC AACAGGTTTA TTATCACATG TTATGAATGG TGATAAAAAG ACTTCATTTT 1020
 50 TAATTGGTGA TGGCACAGGT ATGGGATTGC CTGAAAGTGA TTATTTGCTT TTTGAGGCAT 1080
 GTGAATATAG ACGTCACTTT TTAAGTTATA AACCTGATTA CGCAATTATG ACAAATATTG 1140

55

TGGCACATAA TGTTAAAAA GGTATTATTG CTTGGGGTGA TGATGAACAT CTACGTAAAA 1260
 TTGAAGCAGA TGTTCCAATT TATTATTATG GATTTAAAGA TTCGGATGAC ATTTATGCTC 1320
 5 AAAATATTCA AATTACGGAT AAAGGTACTG CTTTGTATGT GTATGTGGAT GGTGAGTTTT 1380
 ATGATCACTT CCTGTCTCCA CAATATGGTG ACCATACAGT TTTAAATGCA TTAGCTGTAA 1440
 TTGCGATTAG TTATTTAGAG AAGCTAGATG TTACAAATAT TAAAGAAGCA TTAGAAACGT 1500
 10 TTGGTGGTGT TAAACGTCGT TTCAATGAAA CTACAATTGC AAATCAAGTT ATTGTAGATG 1560
 ATTATGCACA CCATCCAAGA GAAATTAGTG CTACAATTGA AACAGCACGA AAGAAATATC 1620
 CACATAAAGA AGTTGTTGCA GTATTTCAAC CACACACTTT CTCTAGAACA CAGGCATTTT 1680
 15 TAAATGAATT TGCAGAAAGT TTAAGTAAAG CAGATCGTGT ATTCTTATGT GAAATTTTTG 1740
 GATCAATTAG AGAAAATACT GCGCGATTAA CGATACAAGA TTTAATTGAT AAAATTGAAG 1800
 20 GTGCATCGTT AATTAATGAA GATTCTATTA ATGTATTAGA ACAATTTGAT AATGCTGTTA 1860
 TTTTATTTAT GGGTGCAGGT GATATTCAAA AATTACAAA TGCATATTTA GATAAATTAG 1920
 GCATGAAAAA TCGTTTTTAA TATGTTTATA ATAGAGTAGT ATGGGTATTT ATTATTAATG 1980
 25 ACATTATTAC ATGTTAATTA GGAGGCGTTT TTAATGGATT GGATTTTACC AATTGCTGGA 2040
 ATTATCGCTG CGATTGCATT CTTAATTTTA TGTATCGGTA TCGTAGCTGT ATTAAATTCT 2100
 GTTAAGAAAA ACTTAGATTA TGTTGCAAAA ACACTTGACG GTGTAGAAGG TCAAGTTCAA 2160
 30 GGTATTACTC GTGAAACAAC AGATTTACTT CATAAAGTAA ACCGTTTAAC TGAGGATATC 2220
 CAAGGTAAAG TAGATCGTTT AAACCTCAGT GTAGATGCTG TTAAAGGTAT CGGTGACTCA 2280
 GTACAAACGT TAAACAGCTC TGTAGATCGT GTAACAAATT CAATTACACA TAATATTTCT 2340
 35 CAAAATGAAG ATAAATCTC ACAAGTTGTT CAATGGTCAA ATGTTGCAAT GGAAATTGCA 2400
 GACAAATGGC AAAATAGACA CTAC 2424

(2) INFORMATION FOR SEQ ID NO: 393:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 738 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 393:

50 CnATAATATT CnTCAAnCCT TTAAATAGAG GAATAGACTG CCGACAGAGT CCGAGACTTT 60
 GTGGGTAGTT TTTTtagTTT TGATAACGGA AGTTAGAGGC TCTCTGTCAA ATTGGGCAGA 120

AGTGTAAGT TTTCAACATA ATACTATTAG TTCGGTCATG TATCGGACTG ATGGAAAAGC 240
 GTTTCACCTT TAATGACTCA TTAAGAACGG CCTGAAAATG TTTGGCGTAT TAAGTGCAAT 300
 5 GATAGTTTGT ACATTAGTGT TCTAATTGGT CATTACTGCC GAGCAAATCT AGTAGAGTAA 360
 TCATGTAAAT CTTTAATGTG CCATTGATT CACTAGCGGT GTTAATAACT ACGGAAATTG 420
 CATTTCGGAC TGAAATTTTT GAAAAATATC AACGTACGCT ACAAATAAAA TTTTAACTG 480
 10 TTATAAATGT GTCTCAATTT CATATGTTCA TCGACGATAT GAAGCGTATT ATGGTAAAAT 540
 GAAGAAATAA TAACTTGTT AATAAATAAA ACATCAGCAT TTGACTAAAG CACTTTATTA 600
 15 TTGTGTAGAT AATAGTTTTT TAACGAAATA AAAATGGCGA CTGGTTTTAA TAAATCAGCT 660
 AATGAATCAC TACACCTATA AGTATGAATA TAGTGATTAG AATGCTTTGT ATAGTTGGAT 720
 TTTGCAAAAT TGATGTTA 738

(2) INFORMATION FOR SEQ ID NO: 394:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1270 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 394:

30 AAAAGTTGTA ATTAAAAGTG GGATTTTACT TAAGnnAGAA GGAAACTATT TATATGACTA 60
 ATAAAAGAGn AGATGTCCGC AATATAGCAA TTATTGCTCA CGTTGACCAT GGTAAAACAA 120
 CTTTAGTAGA TGAGTTGTTA AAACAATCTG GTATATTCAG AGAAAATGAA CATGTCGATG 180
 35 AACGTGCAAT GGACTCTAAC GATATCGAAA GAGAGCGTGG AATTACGATT CTAGCCAAAA 240
 ATACGGCTGT TGATTATAAA GGTACACGTA TTAATATTTT GGATACACCA GGACATGCAG 300
 40 ACTTTGGTGG AGAAGTAGAA CGTATTATGA AAATGGTTGA TGGGGTTGTC TTAGTAGTAG 360
 ATGCGTATGA AGGTACAATG CCTCAAACAC GTTTTGTACT TAAAAAGCG CTAGAACAAA 420
 ACCTGAAACC TGTGTTGTT GTTAATAAAA TTGATAAACC ATCAGCACGT CCAGAGGGTG 480
 45 TTGTAGATGA AGTTTtagAT TTATTTATTG AATTAGAAGC AAACGnTGAA CAATTAGAAT 540
 TCCCTGTTGT TTATGCTTCA GCAGTAAATG GTACAGCTAG CTTAGATCCT GAAAAGCAAG 600
 ATGATAATTT ACAATCATTa TATGAAACAA TTATTGATTc ATGTACCAGC TCCAATTGAT 660
 50 AACAGTGATG AGCCCATTAAC AATTTCCAAG TAGCATTGTT GGAATACAAT GATTATGTTG 720
 GACGTATTGG TATTGGTCGT GTATTCAGAG GTAAAATGCG TGTCGGAGAT AATGTATCAC 780

55

5 GATTA⁵AAACG TTTAGAAATT GAAGAAGCAC AAGCTGGAGA TTTAATTGCT GTTTCAGGTA 900
 TGGAAGACAT TAATGTTGGT GAA¹⁰actGTAA CACCACATGA CCATCAAGAA GCATTGCCAG 960
 TTCTACGTAT TGATGAGCCT ACTCTTGAAA TGACATTTAA AGTTAACAAT TCTCCATTG 1020
 CTGGCCGTGA AGGTGACTTT GTAACAGCAC GTCAAATTCA AGAACGTTTA AATCAACAAT 1080
 TAGAAACAGA TGTATCTTTG AAAGTTTCTA ACACAGATTG TCCAGATACA TGGGTAGTTG 1140
 10 CTGGTCGCGG TGAATGTCAT TTATCAATCC TTATTGAAAA TATGCGTCGT GAAGGTTATG 1200
 AATTACAAGT TTCAAACCA CAAGTAATTA TTAAAGAAAT AGATGGTGTA ATGTGTGAAC 1260
 CATTGAACG 1270

(2) INFORMATION FOR SEQ ID NO: 395:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1365 base pairs
 20 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 395:

25 AAGACCAGGA GAACAmGTAA AACAAATATAC AGTTGaAATC GCTCGTAAAT TAATGGAATT 60
 TGATATAAAA TGCTCGTGAT TGCTTGTA³⁰AA TACGCAACTG CTGTnGCTTT AGAATATTTA 120
 CAAAAGACCT TATCAATCCC CAGTGATTGG CGTAATTGAc CAGGTGCTAG AaCAGCAATA 180
 ATGACTACTA GAAATCAAAA TGTATTAGTA CTAGGAACGG AAGGCACAAT TAAATCTGAA 240
 GCATATCGTA CGCATATTAA ACGTATCAAT CCACATGTAG AGGTACATGg CGTTGCCTGT 300
 35 CCAGGTTTTG TGCCACTTGT AGAACAAATG AGATATAGTG ATCCAACAAT TACAAGCATT 360
 GTCATTCATC AAACACTGAA ACGTTGGCGT AATAGTGAGT CTGATACTGT CATTTT⁴⁰TAGGA 420
 TGTACCCACT ATCCATTGCT CTATAAACCT ATCTATGATT ATTTTGGTGG TAAAAAGACA 480
 GTGATTTCTG CTGGATTAGA AACGGCTCGT GAAGTTAGTG CATTGCTAAC ATTTAGTAAT 540
 GAACATGCAA GTTATACTGA ACATCCAGAT CATCGATTTT TTGCAACAGG TGATCCTACT 600
 CACATTACTA ACATTATCAA AGAGTGGT⁴⁵TA AATTTATCTG TCAATGTGGA ACGTATATCA 660
 GTGAATGACT AGGAGGATTT TTAATGAAAG AGATTGTTAT TGCATCGAAT AATCAAGGGA 720
 AAATAAATGA CTTTAAAGTA ATATTTCCAG ATTACCACGT AATAGGTATT TCAGAACTAA 780
 50 TACCAGATTT TGATGTGGAA GAAACAGGAT CAACATTTGA AGAAAATGCT ATATTAA⁵⁵AAT 840
 CAGAAGCTGC TGCAAAAGCA TTGAATAAAA CGGTCATAGC TGATGACAGT GGA⁶⁰CTAGAAG 900

GCGATGAAGC AAATATTGAA AAATTATTAA ATAAGCTTGG TAATACAACT GATCGTCGTG 1020
 CGCAATTtGT TTGTGTCATA AGTATGAGTG GCCCTGATAT GGAAACAAAA GTATTTAAAG 1080
 5 GTACTGTTTC AGGTGAAATT GCAGATGGAA AATATGGCGA AAATGGTTTC GGATATGATC 1140
 CGATATTTTA TGTACCGAAA TTAGATAAAA CCATGGCTCA ACTTTCAAAA GAACAAAAAG 1200
 GGCAATTAG TCATAGACGA AATGCGATTA ATTTACTTCA AGCTTTTCTT GAAGGTGATA 1260
 10 AAAATGTCTA AATGGATTAT TGTGAGTGAT AACCATACTG AATCAGGCGT TTTATATCAA 1320
 ATTTATGAAA TGCACCCAGA TGCAGATGTA TATTTACATT TAGGA 1365

(2) INFORMATION FOR SEQ ID NO: 396:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1383 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 396:

25 AATTCCTGGT GCAATAATAA ATAGGATGAA AAAGATTCGG AAAATATGAT AACTCGTAAT 60
 CATAGCAACA TCGGCACCAG TAGCTAATGC AACTAAAACAT ATCTGATTAA CCCCTCCTGG 120
 TGCTGCACCA AGaACAATT CATTAAATAGG ATTATTATCA AAGAAATGTA TGATATAAAC 180
 30 CATGATTAGC GCACCAATTA TCAACATAAT ATTTTGAATT GTAATTGCGA TTGCTAGTCT 240
 ACCTTTTAAA TCTGACAATA AATGCGCAAT TTGAACTCCA ATTCTAATCA TATATATTAG 300
 TTGTGCCATG TTCAACAACC AATGATCTAG TGTAATGTTT AAACCTGTAG AAAAATTCCA 360
 35 AACAATTAAT ACAATGAGTG GTGCTAATAA TTGAAATGTT GGAAACTTTA TTTTAGACAT 420
 AATTAGATAA ACTATAAAGA TAGCTATCGC TAAAATAACT ATTTGCCCTA TGTTTAATAC 480
 TTGTGATAAA GGCAAGACTT TTGTTAACTT TCCATTGCGA TGCATGTTAC CATCATGAAA 540
 40 AAAATATGAA ATGAACGGTA CTAAAACAAC AACAAATATA ATTCGTGATG TTTGCGTTAA 600
 GCTAACAACT AACAAATTAG CACGTTTGTC TTGTTAGGCC ATGACCAGCA TTTGTGTTAG 660
 TGCTCCTGGT ATAACACTTA AAATAGCTGT TTCTGTATTA ATACGTGCAA TTTTTTTAAA 720
 45 AACAAATGCC ATTACTATTG CAATTAATAA TATCGAAATA GATACAACAA TAATCGAAAG 780
 CCAATTGTTT TTAATATCCA TAACGACATT TTTGTAAC GTTGATCCGA TTTGCACACC 840
 50 TAATAGTACA ATACCTAATT CACTAAGTAA GAATGGCCAT TTAATATCAA GTTTGAAAAC 900
 TTTTACACAA ATGATTGATG CGATAATAGG ACCAAACATA AATGGAAGTA ATACGTGCGA 960

TATCATTGCC ATGTTTTCCA CTTCCTTCAA TAAAAAATAA AATGACTAAA TTGCTGCTTG 1080
 AGCTTCACGT TTGTTAAGAT AACAAATATCC GCTAGCAGTT tTGACTACAA AGCATATATG 1140
 5 GaCTTTCACT ATCAAGTCGC CGCCCATGCC TTATATACAT TTAAAAAGAG CCTGAACAAA 1200
 GTTCAGGCTC TCAATTTGTC CGTATATTTA TTTTACAATA CGACTTAAAG CCGTATCAAA 1260
 TGCTTGAATC GTTTTTCAAT ATCTTCTTTC GTGTGTGCCG TAGATAAGAA TGTACCTTCA 1320
 10 AATTGAGATG GnGGnAAAAA CACACCTCTT TnCATTTCTC GGTACATTTC TGCAATAATT 1380
 TCC 1383

(2) INFORMATION FOR SEQ ID NO: 397:

15

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 415 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 397:

25

TCCACTAANA TGATTACAAT TGCATTAGTT TGGTGGAGTG CATTTACAAT CTTAACGGGT 60
 ATGATTAAGA ACCACGGTTT AATTaTTTAG TGAGATTCTT ATTTGGTGTA GGTGAGGCGC 120
 CAATGTACCC TTCTAATGCT GTGTTTAATT CATTTTGGTT CTCTAAAAAT GAAAAAGGTA 180
 30 GAGCATCAAG TGCATTATTA GCAGGATCAT ATTTCCGGACC TGTATTAGCA CCAATAGTTA 240
 CAATTGCTAT TGTTAACGCA TTTAACTGGC AAGCAGTATT TTACATTTT GGTGCACTAG 300
 GTATTTTAAT GGCKGTATtA TGGGCGATTA TTGCCAAAGA CTTACCTGaG CrACATAGwa 360
 35 TGGTTAATGA AGCGGAGAAA CGTTTCATTA TGGAAATCG TGATATCGTA GCTAC 415

(2) INFORMATION FOR SEQ ID NO: 398:

40

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1141 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 398:

50

TTTTAGaTaA aTyCAATTnT CyATaCTaAA TgATTnTCTT ATTaCGTCAA TTcGCCTTTT 60
 aTTTTATCGT AATCTTTCCa CTGCAAAGCT AAAGCTTCTC CTATTCTAAG ACCAGAATAA 120
 AATAACAGTC TAGTTAGCTG ACGAGAAGTA TCATTtGTGA TTTGTTCTAC TTTTTCATCA 180

55

AATGTGGGGT CGTATAAGAG CTTGTAATGC TTTTGGCGT AATTGATAAC TGCTTTAAAA 300
 CCTGCCCCACA CAGATCGTGC ATAGTCAACA GAAAGACCTG CATCGTTTAA CAAATAATTC 360
 5 CTGAAAGCAG TACATTGCGT AGTAGTGATT TTGCCAATAG GGATATTTCC GAACCTTTCT 420
 TTTATGTGAG TATTATATTC TGTAGTTCGC TTTTCTATTG AGCGTGCAGA AAGATTTTCA 480
 TTTTTTAAAC GATCAAAAAA TATATATTCA AAGGGTTGAT TGTCCGAGTA TCCATATTTA 540
 10 ACATTTTGTG TAAATTCGCT TTCAGCTAGT TTGGCATCTT TCTTACGTTT AAACCCACGC 600
 TTCATTTTTT GTTTGTATT ACCGTATACA TCTTTATATC TAATGGAAAA ATACCATTTA 660
 CCTGTATTAT CATCCTTATA TACTGGCATT TTGCTTCTCC CTCCTCAAAA TTGGCAAAAA 720
 15 ATAATAAGGG TAGGCGGGCT ACCCGAAATT TAGTACTAGG TACTAAATGT GATATAATAA 780
 AATAAAAAGT AGGTGATGTT ATGACATTTA AAAACAATCA TAATTTCAAT GAATTAGTTT 840
 20 TAACGAATGA AGACATTAGA ATTTTAAAAA ATGTCTTAGA AGATGCAGTC AGTGTTTATG 900
 ATGAATATTC GGTATGTAAT GAAGAATCCG ATTTTGCTTA CTGTTTATTA AGAGACTTAT 960
 ATACATTAGA CAGCTTAGCT ATTTCTGCAA ATAATGTTTG AATTATCGAA TTGTACTCTT 1020
 25 CGATTTTAAT ACCATGCATA ATAGAGTTTC TGTGTTCAAT AGCAGCTTTG ACTGAATGTC 1080
 TTAAATGTTT TTCTATTAAA TCGTTGTTTt CCAtTTCGtK TaaAAATGtT CytATATTCC 1140
 T 1141

30 (2) INFORMATION FOR SEQ ID NO: 399:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 706 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 399:

40 ACTGTGGTAG GTTTTTTATT TTGAAGTATT AATCATAACA GACTAATAAT CATGAGGTAA 60
 CTAATAACAC ATATTTAACT TGTATTCTTA AACTGGTATA ATAAATTTAT GTTGAAATGA 120
 45 ATATTGTATG ACAGGGTATT CACTTTTATT AAAAGGTAAA ATTAAATAAA GGTTTTATAG 180
 AACGTATTTA AATATATGAG GAGTAAACAA ATGGCTGATA GAACGAATAA AGAAATTAAA 240
 ACAGGACGCT TTATTGCAAC TGCATCAATC GTATTCTCAA TATTATTGAT TATTCATTAC 300
 50 TTTGTTTCGT TGGATAATGC GACTGCCAAA GCATTACTTA ATTTAACGAA TCAAAACACT 360
 TCAGATAAAG CGATTGATTA CATTTTAAAC AGCTTTAGAT TCACTGGTAT TATGTATATT 420

55

ATGTTTGCAG TTTATGTATC AAATAGTTTG TTTACGTTGA TTAATTTATC AATCACAATT 540
 CAAGCAATAA AAGCTGCACA CGGTGCGTAC TTAACATTGC CAATTTTAAT TGTTATTATA 600
 5 GGTTCGGTTG CATTAGCGAT TTATATGCTT GTTGTTCCTA TCAAACGTAA AAGTACATTT 660
 AATCGCTAGA AAATTGATTT TAACAATAAA AATATGAAAA AAAAnn 706

(2) INFORMATION FOR SEQ ID NO: 400:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1187 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 400:

20 ACACAACTCG AAGATTCACG TTGTGGTGCT GGACATGATC GAAAAATTAG AGCTGAACAA 60
 ATGAAAGAAA TCAGTGATTT TGTTAAAAAG AAAAATATCC CTAAAGATGA AACGGTATAT 120
 ATAGGTGGCG ACCTTAATGT CAATAAAGGC ACTCCAGAGT TCAAAGATAT GCTTAAAC 180
 25 TTGAATGTAA ATGATGTTCT ATATGCAGGT CATAATAGCA CATGGGACCC TCAATCAAAT 240
 TCAATTGCGA AATATAATTA CCCTAATGGT AAACCAGAAC ATTTAGACTA TATATTTACA 300
 GATAAGATC ATAAACAACC AAAACAATTA GTCAATGAAG TTGTGACTGA AAAACCTAAG 360
 30 CCATGGGATG TATATGCGTT CCCATATTay aCGTTTACAA TGATTTTTCa GATCATTACC 420
 CAATCAAAGC CTATAGTAAA TAGTGCTCAA CTAATAATA ACTTGCTtCG TTCTAAAAGG 480
 ACGAAGCGAG TTATATTGTT AAAATTGAA TTGACTTACA TTTTAATAAA ATCATCTTAA 540
 35 CAACTTTAAT TTTTCaTTAA TACAaGTCTT TACTCTACAC TCAAACnAGA TTCATACACT 600
 GCACGTCATA ATAAATCTAT CTATTCAAAT ATAAATAAAA GTTACCTACT ACATTCTATG 660
 40 TAGCAGGCAA CTTTTATTAC TTATTTCTTT TCATTATCAT TAAGTACTTT TACAAACTTC 720
 ACATTATGTG TCTTCCAATC AACTTCATAT AATGCTGATA ATTTTCTTC TTTTATCT 780
 ACATGGTTTT CACCAGACCA ATAGCCCCAG AAACCATGGC GATTCCAATC TATTTTAAAC 840
 45 TCATCCATTG ATCTTTTATA ATGAACAACA AATTGTGATT TACCTTTGTC TTTTATCA 900
 TGTGACATAA CAGCTAAAAA TTCTGGATTA AACCCTTCAG ACACAGTTAC AGGCATTTTG 960
 TCTTTAGGTG TGAAATTATC TTTCGCCCCAT AAATTTCCAT TTCGTGTTAA AGAAAAAATT 1020
 50 TCACTTTtag TTCTATTATC ACTATCATTa GTTAATTGTC TCGTATGGTC ATGTCCTATA 1080
 TTATTTATCA AATGTGCTTC TACTTTCCAA CCTACACCTT TATGTGACGT AGATTGATCA 1140

(2) INFORMATION FOR SEQ ID NO: 401:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 847 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 401:

CCAGAATTAT TTTTCAAAA AGGACAATTT AACAAATGTCG ATAACGTTAT CATAAGCAAA 60
 CCGATGAAAG GGACAATGCC TAGAGGTAAA ACGGAAGCTG AAGATCAACA GTATTATAAA 120
 ACATTGCAAA CTTCTTCGAA AGATCGTGCA GAAAATGTCA TGATTGTTGA TTTACTAAGA 180
 AACGATATAG GGAGAATATC ACAGAGTGGC TCAATTAAGG TGTATAAACT ATTTTTTATT 240
 GAGGCATATA AACTGTATT TCAAATGACT TCGATGGTAA GTGGAACTTT AAAAAATAAT 300
 ACAGACTTAA CTCAAATTTT AACATCGTTA TTTCTTGTG GTTCGATTAC AGGTGCACCG 360
 AAACTGAATA CAATGAAATA TATTAAACAA TTAGAAAGTT CACCTCGTGG TATATACTGC 420
 GGACAATTGG ACTATTACTT CCAACTGAAG ATGATAAAAT GATTTTTAAT ATTCCGATTC 480
 GCACTATTGA GTATAAATAT GGACAAGCGA TTTATGGAGT CGGAGCAGGT ATTACAATTG 540
 ATTCTAAGCC AAAAGATGAA GTGAATGAAT TTTACGCAAA AACCAAGATT TTGGAGATGT 600
 TATAATGCAA TTATTTGAAA CAATGAAAAT TGATAATGGA CATATCCCTA GACTTACTTA 660
 TCATACTAAT CGCATAAAAT GTTCTTCTGa GCGATTAAAC TTTAAATTTG ATGAACATGC 720
 ATGGCGAAAT GAATTAAACG ATGTAACAAC AAAGTATCAC AGTGGTCAAT ATAGACTTAA 780
 AATCGTATTA AATGCTGAAA GCAAATTTGA AACGATAGTG TCACCTTTAC CTGAGAAAAG 840
 TAGTTTT 847

(2) INFORMATION FOR SEQ ID NO: 402:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 740 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 402:

TGAAGATGAA GCAGAAGCAG AAGACAATCT GCTACGAGTA CAATCGAAGA AAGAAGAAGT 60
 GTATCGTCGA TTAATTGCTT CTAACCTAAC AAGCGTTCCT GAAAGTTTAT CATCATGAAA 120

GTTAAACAGT TAAAAGATAA AGTGTCTAAA ATTGTGATTC AAATGAATAC ATTTGAAGAT 240
 GAAGCAAATG ATGTTCTTGT TAATGCTGTT TATGCAGAGA AATTAATTCA ATATGGAAAT 300
 5 AGATATCGTA AGGACTATAG CAATGTTGAT AAGAGCTTAA ATGAAGCTGA ACGATTATTT 360
 AAAAATAATC GCTATAAGCg TGCGATTGAA ATTGCAGAGC AAGCTCTTGA AAGTGGTTGAG 420
 CCAGGTGTTA CTAAACATAT TGAAGAAGAA GTTATTAAGC AATAGAAACT AGTATGTAGT 480
 10 TATACTTAAA TAATATGAGC ACTCTGTCAA ATTGGACTGA TGAGTTTAAT AATTGAAGTT 540
 AGCCAACGAT ACGTTGTCTA GCTTCTTTTT TATATGGATA AATGaAAGGG ACAATAAATA 600
 TAAATAGCAA TTGTTTAAAG ATAAACGTAA TCAAATGTGT TGTTTTAATT AATATAAGTA 660
 15 GTGAAAAAAG CATAATCACA CAGCTGTTTA AATAGAGTGA AATAGTCTAA TTCTTATTTA 720
 ATAAGTAGAA ATAAGATTAT 740

20 (2) INFORMATION FOR SEQ ID NO: 403:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 630 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 403:

30 ATGCCAATAA TTAAACCATG TAATAAATAT ACGTATAGCG TACGACTACC AATATAAGTA 60
 TATAATTTTT TCTTTGTTGA CATTAAATTT AGAAACGCAG TCATTGCGAT TAATATAATT 120
 35 CCATATAATA TAAGTCGTTT AAAAGGACTG AATATACTCT GTCCTTCATT TTCAAGTGAA 180
 GTATATGGTG AACTTCCCAA TAACCAATCT GCATTGATAG GATGAATCAC GTAAACGATA 240
 AAAAACAAAA TAAAGGTAAT GATAGATACT GGTATTAGTT TTTTATTTTT AAAAATAGCC 300
 40 GTATGTTTTT TGGTGAAAAT GTAACCTAGA TAAAATATTG GGAAAAATAC GATTGTCCTT 360
 GAAATGCTTA AGTAGCTATC GATGTTATCT GAAAACCTG CTCCAATAGA TATAATAATT 420
 GAACTGATA GCACTTTATA TGGATTAAAT CTTCTAACTA TTACTAAAAT GACATGAAAG 480
 45 AAAAATAGCG TGATCAAAAA CCATAACGCA AATACTGGGT TAAAAGGATC AAGTTGTAAT 540
 TCGTCACTTT TACCTGTTAA GAAATAATAA ATTGAAAAGA ATGCAAAAAA TATCATATAA 600
 GGTACTATCA AACGTTTTGA AATTTTTTCT 630

50 (2) INFORMATION FOR SEQ ID NO: 404:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6254 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 404:

	ATTTAAACGT TCAGTTTCTA AAAGTGAAC AATCCAAGAA GTACGTAAAC GTGAATTTTA	60
10	CGAAAAACCA AGCGTAAAC GTAAAAAGAA ATCAGAAGCT GCACGTAAAC GTAAATTCAA	120
	ATAATTAATA CCTCTGTTGA CTCCCTCAAC ACGAATATTA ATTATATAAA ACAAACATCA	180
	CAAGTTAGTG TCTGACACTA ATATGTGATG TTTTTTTGTT GTCAATTTTT AATTAAAAAA	240
15	AGTTATATAG TTTATAAATA ATCAAATTGA TATTCTATAG GTTCTTATAA CTATAAAGTA	300
	TATTCAATTT CATGTATAAT TAATGTGAGG GCGAGGTGAA ATTGTGAGTT ATAATAATTT	360
	TTTACAAATG ACAACTATCT TGGAATCAAC GGCTGGAGAT ACTTGGGTTG AACAAGTTAG	420
20	CAATATAATT GTTCAACCTA TTTTACGTT AATATTAACC TGTTTGACAT TCTTAGGATT	480
	TGTATATCAA CTTTACTCTA AAAAAATCAA TGCAGCTGGT ATTATCGCCA CATTATCATT	540
25	ACTTATTTTA TTTTGGGAT TTCTAATCCA AGGAAATGTC AATATGCATT CTATCTTAAT	600
	ATTCTCAATT GGC GTTATAT TAGTTGTAAT TGAATTATTT GTAGTTGGTG CAGTAATTGG	660
	TATTATTGGC ATGATACTGA TAACTATAAG CATAACAACG CTCGGTGATA ATTTGCTATT	720
30	TATGCTTGCG AATGTTATCG TTGCCTTGAT TTTAACGATT GTAGAATGGG TGATATTAGT	780
	GAAGATTTTC AACAGAAAGA TTCCGTTTTT GGATAAAGTT ATCTTAAAAG ATTCAACTAA	840
	TTCTGAGTCA GGTTACAATT CTCATGATAA CCGCTCGCAC CTCGTAGGAA AGACTGCTCA	900
35	AACAGTTACA GATCTTCGAC CTGCAGGGAT TATTTTTTGT GAAAATGAAC GTATTGATGC	960
	TGTTTCAGAT GGCAACTTTA TTTTGCGCAA TAAACGGTA AAAATCCTTG AAGTTGAAGG	1020
	AACAAGAGTA GTTGTGAGGG AAGTAGATTA ATTAAGGA GCGATACCAT GTTTAGTTTA	1080
40	AGTTTTATCG TAATAGCAGT TATTATAGTA GTTGCATTAC TTATTTTATT CTCATTTGTA	1140
	CCCATTGGTT TATGGATTTC AGCGTTAGcA GCTGGCGTTC ATGTTGGTAT AGGTACATTG	1200
45	GTTGGTATGC GTTTACGTCG TGTATCTCCA AGAAAAGTTA TAGCGCCATT AATTAAAGCG	1260
	CACAAAGCAG GACTAGCATT AACAACAAAC CAATTAGAAT CGCATTATCT AGCAGGAGGA	1320
	AATGTTGACA GAGTTGTTGA CGCTAATATT GCTGCACAAC GTgcTGACAT TGATCTTCCT	1380
50	TTCGAACGTG CTGCTGCAAT TGaCCTTGCA GGACGTGACG TATTAGAAGC GGTTCAAATG	1440
	TCTGTTAATC CTAAAGTCAT TGAAACACCA TTTATCGCAG GTGTAGCAAT GaACGGTATT	1500
	GAAGTGAAAG CCAAGCTCG TATCACAGTT AGAGCTAATA TTGCTCGACT TGTTGGTGGT	1560

55

	AGTAAGCATC ATACAGAAGT ACTTGAAAAC CCAGATAATA TTTCTAAAAC AGTTTTTAAGC	1680
	AAAGGTTTTAG ATTCAGGTAC TGCATTTGAA ATTTTATCAA TTGATATTGC TGACGTTGAT	1740
5	ATTAGTAAAA ATATTGGTGC AGACTTACaA ACTGAACAAG CATTAGCAGA CAAAAATATT	1800
	GCACAAGCAA AAGCTGAAGA ACGTAGAGCT ATGGCTGTAG CAACTGAGCA AGAAATGAAA	1860
10	GCGCGTGTAC AAGAAATGCA TGCTAAAGTA GTTGAAGCCG AATCTGAAGT ACCATTAGCT	1920
	ATGGCTGAAG CATTACGTTT AGGTAATATC AGTGTAAAG ATTATTATAA TTTGAAAAAT	1980
	ATCGAAGCTG ATACAGGCAT GAGAAATGCA ATTAATAAAC GAACTGATCA AAGTGATGAT	2040
15	GAGTCACCTG AACATTAAGT CGAGAGGTGA TTAAATGAGT GTCGGTATTC TAATTTTTGT	2100
	CATATCAGTG ATCATTTCTA TCATTACTAC TATGCGCGAA AATAGTCATA AAGATAGACA	2160
	AAATCAAAAG CCACCTCAA AAACATCTAC CGATAATGAA CCAAAAAAAG GTGGCTTTTT	2220
20	TGAAGAAATT GAGCGAACGT TTAAAGAAAT AAGTGAAGAA TTAAATGAAG AAGAAAAGAA	2280
	ATCATCGAAA CGAAAATATG ATGATACGTT ACCACCTTTA TTCGATGAAC TTCCAAAGGA	2340
	AGAGCCTAAA TCGAAACCTG TTGTAGAACC TATGGCACCT AAAAAACAAC AAGAAACAAA	2400
25	ACCGATGACA GAGAAACCAA TCACAGTGCC TAAAGCAGAA CCGGTGGAGC AGAAACATAG	2460
	ACCTTCTAGA CAAGATAATT CTGACGAAAT TAGACGTCAA TTAGAAAAAT CACTTAGAGA	2520
30	TGATATTAAA ACGATTCGTA CTGACATTGA TAGAGAAAAA GAAAAGCAAA TTGCTAAAAAT	2580
	GGAAAAACGT GCTAGAGATA TTATTGAGGA TAAATACTTA TCTGAACGTA CAAAACGTTT	2640
	GAAATTAAAG CAGCTGCTTA ATTCTCAAAA TGTCGAAAAA GATTTGACTA AATCAGCGTT	2700
35	CCAATTTGAT AAAGATGAAG TAATCAATGG TATGATATGG TCAGAAATTT TAGCTAAACC	2760
	AAAACAATTA TAAAATTTTT TGAAAACAAG CACTATCGTA ATGGTAGTTG CTTGTTTTTT	2820
	TACGTTAAGG AAAATTAAAA AACAAAGAGA ATTTTTCGAG AAATATTAGT TATTTAAATT	2880
40	ACAGCAAAAA ATTGATTAGT CTAAAATTGA ATCTGCTTTT ATGACAAGGT GAAAAGTATA	2940
	AATGATTATT TTAAATTAAA GAAAAATGAG TAAGTCAATG CAAAGATGTT TAAATCAATC	3000
	AATTGCATGA TATAATTAAAG TAGATATTAA AGCATCATAG AATGAATATA AATGATATAT	3060
45	GAAAAGGAGC GCGTGTATGC CTGGAATTAT ACAAATAGAC GATATGAACC AATCTCAAGC	3120
	TTTAATTGGA AATAATGATG AACATTTAAA AGCAATTGAA GAGAGTTTCG ATGTTGTCAT	3180
50	CCATGCAAGA GGACAAGaAG TTGCCGTTAA AGGTACAAAA ATAGAAAACG TAGAAAAAGC	3240
	GGAATCAGTA TTAATCAATT TGCTGAAGGT TATTGATTTA GGTAAATAATA TTACAATTAA	3300
	AGATGTTGAA GCAGCTATTA AAATGGCGCA TAATAACACA ATTCAACATC TGTTAGATTT	3360

55

	GCAACGTATA TATGTTAATG CCATGAAAAA TAATGATTTA GTATTTGGTA TAGGTCCTGC	3480
	TGGTACAGGT AAGACATTCT TAGCTGTAGT TTATGCAGCA AAGCAACTCC GTAAAGGTGC	3540
5	TGTTAAACGT ATTGATTATA CAAGACCTGC TGTTGAAGCA GGAGAGTCAC TTGGATTTTT	3600
	ACCAGGAGAT TTGAAAGAAA AGGTAGATCC ATATTTAAGA CCTTTATATG ATGGTCTATA	3660
10	TACTGTTCTT GGGCGTGAAC AAACAGAGCG ATTTATTGAA AGAGGCATTA TCGAAATAGC	3720
	GCCACTTGCA TATATGCGCG GACGAACATT AGAAGATGCA TTTGTAATTC TTGATGAGGC	3780
	GCAGAATACG ACACATGCGC AAATGAAAAT GTTTTTAACA AGACTAGGTT TTGGCTCAAA	3840
15	AATGGTAGTT ACTGGTGACC AAATCAAAT CGATTTACCT AAAGGTGTTA AAAGTGACT	3900
	TAAGGAAGCG GTCAGTAGGT TACACAACGT TAAAGGTATA AGTATATTGA AATTAGATCA	3960
	GAGCGATGTA GTCAGACATC CATTGGTAAG TAAGATCATT GAACATTATG AAGGAGAGAA	4020
20	TTAAATGTTT ACGATAGATT TTAGCGATCA CACAGGCTTA GTTAAAGATG CTTGGTATAA	4080
	ACAAATTGAA GATTTATTAG AATTTGCTAA AAAAGAAGAG CATATAGAAG ACGATGCTGA	4140
	GCTTCTGTT ACATTTGTAG ATAAACAAGA AATACAAGAA ATTAATCGAA CATATAGAGA	4200
25	TAAGwTAAR GTTmCaGATG tAaTCyCaTT tGCTTTAGrA GrAGATGAGC CmGaGATkGA	4260
	TtTAGTGGT CTTGATATAC CACGTGTTTT AGGGGATATA ATTATCTGtA CGGATGTAGC	4320
	GCAAGrACAA GCAAACAATT ACGGACATTC TTTTGAACGA GAATTAGGAT TTTTAGCATT	4380
30	ACATGGATTT TTGCATCTAT TAGGTTATGA TCATATGACT GAAGCGGATG AAAAGGAAAT	4440
	GTITGGTCTGA CAAGATACAA TATTAAACGC ATATGGATTA ACACGAGACT AATTATGAAA	4500
35	AGGTTTAAAT ATGCACTTGA TGGGCTGAAA ATCTTAATTC AAAAAGACTA TAAATTTCTT	4560
	TTACATGTGT TTGCAATGAT TGTTGCTATT GTCTTTGGTC TCGTACTAAA TATTAATCGG	4620
	ATTGAGTGGG TATTTATACT CATTGCTATT GCATTAGTTC TCACTGTTGA AGCTTTAAAC	4680
40	ACTGCTATTG AATATGTTGT CGATTTAGTG ACCGTTGAAT ATCATGATTT AGCTAAATAC	4740
	GCTAAAGATA TTGCGGCTTT TAGTGTACTT ATAGTTTCAA TATTAGCATT TATTATAGGT	4800
	TTAATAGTAT TTTTACCACA TTTTATAGCG TTATTTTAGG GAGGCATATA TGAGTTATCA	4860
45	ACCTCATTAT TTTCAAGAAG TTAGAAAAGC ACAACAAGAA TCATATTCGC CATAAGTCA	4920
	ATTTAAAGTA GGGGCTTATT TAAAAmCGAA AGACgGTAGA ACTTTTTATG GTACCAATGT	4980
	AGAAAATGCT TCTTATCCAT TATCGATATG TGCTGAACGA GCTAGTTTGG TATCGGCAAT	5040
50	TTCTCAAGGA TACAGACCAG GTGATTTTGA ATCAAtAACT GTAACCGTAG ATGCAGATAA	5100
	ACCGTCATCA CCTTGTGGTG CATGTCGTCA AGTTTTGAAG GAATTATGTG ATGATGATAT	5160

55

ACCATTGGA TTTTCAGGAA AGGATTTAGA ATAAATGACA GAACATAAAT CAGGATTTGT 5280
 TTCAATTATA GGTAGACCAA ATGTAGGAAA GTCAACATTT GTTAATAGAG TGATCGGCCA 5340
 5 TAAAATAGCA ATCATGTCCG ATAAAGCTCA AACAACTAGA AATAAAATTC AAGGTGTTAT 5400
 GACAAGAGAT GACGCGCAAA TTATATTCAT TGATACGCCA GGTATTCATA AACCTAAACA 5460
 CAAATTAGGT GACTATATGA TGAAAGTCGC TAAAAATACA TTATCTGAGA TAGATGCAAT 5520
 10 CATGTTTATG GTTAATGCCA ATGAGGAmAT TGGACGAGGC GATGAATATA TTATAGAAAT 5580
 GTTGAAAAAT GTTAAGACAC CAGTATTTTT AGTATTAAAT AAAATAGATT TAGTGCAATCC 5640
 AGATGAATTA ATGCCAAAGA TTGAAGAATA TCAAAGTTAT ATGGACTTTA CAGAGATTGT 5700
 15 ACCTATTTCA GCATTAGAAG GGCTAAATGT CGATCATTTT ATTGATGTTT TAAAGACGTA 5760
 TTTACCCGAA GACCTAAAT ATTATCCAGA TGATCAAATT TCAGACCATC CTGAACAATT 5820
 20 TGTAGTGGGT GAAATCATTC GTGAAAAAT CCTTCATCTT ACAAGTGAAG AAATCCCTCA 5880
 TGCGATTGGT GTTAATGTGG ACCGTATGGT TAAAGAAAGC GAAGATCGTG TTCATATCGA 5940
 AGCAACTATA TATGTTGAAA GASGTTGCA AAAAGGAATT GTCATTGGAA AAGGCGGTAA 6000
 25 AAAGTTAAAA GAAGTAGGaa AACGTGCGAG ACGTGAAtATA GaaATGCTtTC TAGGCTCTAA 6060
 AGTTTACTTA GAATTATGGG TCAAAGTTCA AAGAGACTGG CGAAACAAAG TTAACTTTAT 6120
 TCGCCAAATT GGTATGTTG AAGACCAAGA TTAATCTTAA AAGTGGTGAA GATAATTGTT 6180
 30 AATGCGCCAA AAAGGATTA TCATCAAAGC AGTTGATTAT GGTGAATCTG ATAAAATTAT 6240
 CACGATTTTA AATG 6254

(2) INFORMATION FOR SEQ ID NO: 405:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3710 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 405:

45 GTTGTTCTAA ATGTTTCTTG nATGAAACGA GTCAATGTAA ACTGATATTG CTGTATTTGT 60
 TGCAGCAATT CATATTGGTC TGGTGTGCA ATAACAGCAG CTTGAGTTGG AGTCGCAGCT 120
 CTGATGTCTG CAGCAAAATC ACTTAATGTA AAGTCTGTTT CATGACCAAC TGCTGATATA 180
 50 ATCGGTGTCT TACAATTATA TATTGCACGG ACGACAGCTT CTTGTTGAA ATTCCATAAA 240
 TCTTCTATGG ATCCACCGCC TCGACCTACA ATAATGGTAT CTACACCTAA ACTATCTGCA 300

	TGTATTTGTT	CAGCTAATGG	AAAACGACTA	TTTATCGTTG	AATGGATATC	TCGAATTGCG	420
	GCACCTGTAC	TCGCTGTTAA	AACTGCAATT	TTTTTAGGAA	ACTTAGGTAT	TGATTTCTTA	480
5	TTGCTTTTAT	CAAAACAACC	TTCTTCAGTT	AATTTTTTCT	TTAATGCTTC	TAATTTTTGA	540
	TATAAGTTCC	CTATACCATC	TAATTGCATT	TTATTTACAT	AAATTTGATA	GTTTCCACGA	600
10	CGTTCAAAAA	CAGAAACACG	TGCTTCTAAT	AAGACTTCAT	CTCCTTCTTT	AGGTTGGAAG	660
	TTTAATTTAG	AAGCACTACC	TTTGAACATC	ATGGCACTTA	TAACGCTTTC	TTTATCTTTC	720
	ACATTAAAGT	ATAAATGACC	ACTTGaATGc	TTTTTGAAAT	TTGAAAGCTC	ACCTTTAATC	780
15	AATACAGATT	GGAGATGTGG	ATCTTGATCA	AATTTATATT	TAATATATTT	CGTTAAAGCT	840
	GAAACACTTA	AATAATCTGA	CATATAACAT	CACTCAATTT	TATTTTTTTA	TATTACTCAA	900
	TACACCATT	ATAAATTTAT	AATGATCATC	ATCACTGAAT	TGTTTTGTTA	ATTCAACTGC	960
20	TTCATTCATT	ACGACTTTAG	CAGGTGTATC	ACTGTGTAAT	ATTCATATG	TTGCCATTCT	1020
	TAAAATAATA	CGATCCGTTT	TTAATAACG	TGCAATAGTC	CAATCTTTTA	AATAAGGACT	1080
	AATTGTCTCG	TCTAATACAG	GTTGCTGATC	TTTAACGCCA	GAACTAGCC	AATGAATAAA	1140
25	TTCGAAGTCT	AAATCTGGAT	TATCGTCTTT	AATAAAGCTT	ATCGCTTCAT	TTATCGTTAA	1200
	ATCACTGTCC	TTCATTTCTA	ATTGAAATAA	AGTTTGAAAA	GCTTGCACTC	GGGATTCTTT	1260
30	ACGACTCATT	TTTAACTCCT	TCAAACGTTT	GTATTTTTCT	TTATTTAATT	ACTGAATTAG	1320
	GTATGACATT	ACTTTTCAAT	AACGATTTGT	GTAATGTGAA	TATTAATTG	CTTAGGTTCT	1380
	ATCGCTGTCA	TATTAGAAAT	TGAATTAAAA	ATTGACGTTT	GAATTTTGTT	TGCAGTTTTT	1440
35	GAAATATTAA	CACCATGTTT	TAATGCACAA	TATACATCTA	TATATATGCC	ATCTTCTTTA	1500
	CTCTCGATTT	TTAAATCAG	GCTTAAATTT	TTACGACTAA	CTTTTTCTAA	ATTTGTTTCT	1560
	TTTAATTCAG	CAAATGGCC	AGTGATGCCT	TCGACTTCCG	AAGTAGCTAT	ACTTGCAATA	1620
40	ACAGATAGCA	CTTCTGGCGC	TATTTCTACT	TTACCTAATT	TTGAATTGTA	ATAATCAGTT	1680
	ACTTTGACCA	TGGATTGACC	TCCTATTAAC	CTTCATCATT	CATAATGCTA	TTTTGCTCTA	1740
45	AAAAGTTTGT	ATTAAATTTA	CCGCTTCTAA	ATATATCGTT	ATTCAATAAT	TTAATATGGA	1800
	ATGGAATAGT	TGTATCAATA	CCAAGAACCA	CAAATTCATC	TAGTGACGTA	ATGCCAGCCA	1860
	TAATCGCTTC	ATCTCGTGTC	GGTTCATGTA	TGATTAATTT	CGCTACCATC	GAATCATAAT	1920
50	ATGGCGGTAT	CGTATAATTA	GTATAACATG	CTGACTCTAT	TCGAACACCA	TATCCACCTG	1980
	GTGCAAGATA	TTGCTCmATT	TTACCTGGTG	ATGGCATAAA	GTTCTTGTA	GGATTTTCAG	2040
55	CATTAATTCT	AAATTCAATT	GCGTGTCCTG	TTAATTTAAT	ATCTTCTTGT	TTATACGGTA	2100

CAGTTACAGG ATGTTCTACT TGAATACGTG TATTCATTTT CATAAAATAA AATTTATTAT 2220
 CATTTAAATC ATATATAAAC TCAATTGTTC CCGCATTTTC ATAATTTACA GCTTTGCTG 2280
 5 CACGAAGTGC GGCATTTCCC ATTTACACGAC GTGTTTCATC ATCTAAAATT GGGGAAGGTG 2340
 CTTCTTCCAC TAATTTCTGC ATACGTCTTT GAATTGTACA ATCACGTTCT CCTAAATGAA 2400
 TTACATTACC ATAGCTGTCC CCAACAATTT GGATTTCAAT ATGGCGGAAG TTTTCGATGA 2460
 10 ATTTCTCCAT ATAAAGTCCA CCATTACCAA ATGCAGTTTG AGCTTCTTGT TCTGTCATTC 2520
 GGAAGCCAGT TTCAAGTTCT TTTTCATCAC GAGCAACACG GATACCTTTT CCGCCACCGC 2580
 15 CAGCAGTAGC TTTAATGATG ACCGGATAGC CAATTTTTTTT GGCGATTTTC TTAGCTTCTG 2640
 AGACGTCTTT CATTAACCG TCACTACCAG GAACAACCTGG AACATTGGCT TTGATCATTT 2700
 CTGCCTTAGC AACATCTTTG ATACCCATTT TTTGGATAGA TTGATAACTT GGTCCAATGA 2760
 20 ACTTCAATTG GcATgctTCG CATAATTCTG CAAAATCAGC ATTTTCAGCT AAAAAGCCAT 2820
 AACCCGGATG AACGCCATCA CAACCTGTAG AAGTTGCAAT AGATAAGATG TTCGGAATAT 2880
 TTAAATATGA ATCTTTAGAC AAAGTGGGAC CTACGCAATA TGCTTCATCA GCAATTTGAG 2940
 25 TATGTAGCGC ATCTTTATCC CCTTCAGAAT AGATTGCAAC AGTTTGGATG CCTAAATCAC 3000
 GACAAGCGCG AATAATCCTA ACTGCGATTT CACCGCGGTT TGCAATTAA ACCTTTTTTCA 3060
 TTATTTTACC TTAAATAACG GTTGGCCATA CTCTACCATT TGTCCGTCTT CTTACTAAGAT 3120
 30 TTCAACAATT TCACCTGAAA TTTCTGCTTG AATTTCAATTA AATAGTTTCA TTGCCTCTAA 3180
 AATACACACT GTTGTTCAT TTGAAACAGT GTCCCAACT TGCACATATG CTTCTTCGTC 3240
 35 TGGAGATGGC GATTGTGAAA ATGTACCTAC CATAGGTGCA TTAATTGTTT TGTGATTATC 3300
 TGAAGTTGGC TTTGGAGCTT CAGTTTATT GCTATCAGTT GATTGTGCTT GAGGCATAGG 3360
 CATTGCCGCA GCTTCAACTG GCAATTTGTGA GATTTGTGGC GTGATAATCT CAGTTTCTTT 3420
 40 TTCTTTCTTA AGCGTCACTT TGCCTTAGT ATCTTCAATA TTGATTTCCG TTAAAGTTGA 3480
 TTTATCCAGA ATTTCAATTA ATTCTTGAT TTCTTTAAAG TTCATTATTA CTGACTCCTT 3540
 CAGTTTGTTC TCATCTACCC GTCTATTTTA CTTGAGACAA CTCTTCAATT CAAGCATGTT 3600
 45 CATATTGCTG GCGACATTAT AAGTCTATCC CAAAGTTATA ATAAAACCAC ATTTTAAATT 3660
 AAAAACACTT GTGTATTTAT TACTTAACAT TGAATCATCT TAACTCTTGA 3710

(2) INFORMATION FOR SEQ ID NO: 406:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1705 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 406:

5	GCTGATG TTC GTTGCCCTTT TCCACCAGAC AATTCAGAGG GATATTTATC ACTAATATCC	60
	AATATATTTA ATGCTTCTGC TACTTTTTCA TAACGATTTA ACATATGTTT TTTATCTAAC	120
	TTCTGTACCG TTAGTGGTAA CATTATGTTT TCTTTAACAG TCAATGTATG CAGTAAATTA	180
10	TACTCTTGAA AAATAAAACC AATATCATGC TTGCGTATAT CAGATWATTC CTTGTTTGAA	240
	AGCTTTTCTA ATTTTTTTCC TTTAATGTA ATAGAACCTT GTGAAATATA ATCAATTGAA	300
	CTTAAACAT TTAATAATGT CGTTTTCCCA GATCCAGAGG GACCCATAAT AGCAATAAAC	360
15	TGCGCTTCTT CAATAGACAT ATTGATATCT CGCAACACTT CTTGTGCCAT TTTTTAGTT	420
	CCATATATTT TTGTTAATTG TTTACTTCT AAAATTGCCA CTTTAACACT CCTATAATTT	480
	ATCTTAACTT CATTTCTTTT AGGCTTTGGC ACTTGATCTT TCAATTTAAC ATATGACTAA	540
20	CATCTATCTT ATTATAACGT TGAAGCTGCA TTGATGTATC AATTCTAAGT AACAAAACGC	600
	ATGTTTAAAA TGACAAATTT GTCACCTCCG ACATGCGTTC AACAAATTCA TTTTGTAATG	660
25	GGAAATCAA TCTGACAGTT GTCCCCCTAC CAACAGTCGA CGTGACTGTC AGGTGAATAC	720
	CTAATTGATC CTTTACACTA TTTACTAAAT ATAGACCCAT ACCTGAAGAC GTCGTTTCAT	780
	TTCTGTTAGC CGTTGACGTA AATCCTCGTT CAAATATTCG CGGCATATCT TTTTTACTAA	840
30	TACCTCTGCC ATAGTCTTTA ATATATAACG AAACATGTTG ATCATTTAAT TCTGTCCCAA	900
	TTTCAATATT AAAATTCTCA CTATATTTCA ATGCGTTTGA CAAAATTGT CTAATAATCA	960
	TACGACACCA TTTTATATCT GTATAAACAT AATCATCCAC TTTAAAGTCA ACATCAAAAC	1020
35	CAATACCTTT AACCTGACTA ATATGTCTTG TTAATTGTAT TTCATCAATG ACCATGCGTT	1080
	TAAGTGACAC GTAATCAAAA TACATATCTT TACGTTGAGA TTCTAATCTA GTAATATACA	1140
40	GCTGTGTATC TAGCATCGAG TTTATACGAG ACCATTCTA TAGTAATGCT TGTWTTCTTT	1200
	CTTGATTTTT TTCTTGATCA ATTAATAATT TCATAGCTGT CACAGGtGTT TTTATGTCGT	1260
	GCACAAATTC TGTAATGGTT TGTTCATGCA TGTTCAATTG CAACTGTTGC TCAACAACCT	1320
45	TTTCTTTGTG CGCTGAGATT TGACGATATA AATAATCAAC TGTATGACGT TGAAATGGCG	1380
	TTCCGCTAA ATCTTTATGT TTAATTTCTT CTATTTCTTT ATCTTTGTCA AAATGCTTAT	1440
	ATAATTTTAC TTCTTTAAAA TATGTCAATA AAAGAAAAAT CATTGTTAAA CTTAAATTCA	1500
50	AAGAAACAAT ATAAAATAAA CTGTCTATTG GAAAATCATA ATCGATTAGA CTATGTCCTA	1560
	ACATAAGGAA GTTTAAAAAC AATATCCAAA AtATCCAGTT CATGCGAGAT TTCAAAAAAT	1620

55

GCACTATCCA TACTAATTTT AGATA

1705

(2) INFORMATION FOR SEQ ID NO: 407:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1722 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 407:

15	TCAATATATC TTAAATTC A TGATTAAAT CnATTATCAC TAGACATnAA ATACATAAAT	60
	CCTATTCCCC ATTTTCATTT nTTAATTCAT AAATGAATCA ATAACCACCT AATAACAAAT	120
	CATATTATAC ACCTTTGTTC TCTATTTTTT TAAGGTTTAA AAAATATTTT TAGGTAAACC	180
20	TAAAAATAGA TGTAATAAAA ACGCCTCCTC AGATATTTAT ATATCTATGA AGACGTTTAA	240
	ATACATTATA GATGGTCTGG TTCTGGGTGA ACGTATACTG AGGAAALACC TTTTTGTGC	300
	AAATGATGTT CGACATTGTC ACAAATTTGA TCGCTTCTA CTAAGGAAAG GTTAGCATCT	360
25	ACAACAATTG TGACATCAAT AAACACACTA CTTCCATGGT AACGCCCTTT AATACTTTTA	420
	ACTTCTTGTA CTTTCATCAAC TTCTAAAATA TCATTGCGAT ACGCTTCTAA TTCAGTTTCA	480
30	TTGAAACCAT CACTCAACAT AAAAATTGCT TCTTTAAAAA TACCAAAACC AGTATAAACG	540
	ATTAGTAAGC CTAGTAATGT TGCTAAAATA ATATCGACAA TTGGGAAACC GATTTCGTA	600
	AAAATTAATC CTATCGCTGT TCCAATGCTG ACTAACTAT CCGATAAATT ATCTTTGGCA	660
35	GCCGAATTTA AAGAACTACT TTTGTTCTT TTCGCTAGTC TTTGATTGAC TGCAAATACA	720
	ATCAACATTA CAAGACCACT GATTAAGCTG ACGATAATTG TTATTGCGTT AGGTACAACG	780
	TCATCTTCTT TGAACAAACG AGGTGCATTT TGAATAACTA CTTGGATACC TACAAACATA	840
40	ATGACAAATG ACACCAATAA TGAAGAAATA TTTTCAGACT TCAAATGGCC ATAAGGATGA	900
	TTTCGATCGG CAGGTTTAAT TGAAATTTTC AATCCAATAA TAACAGCTAA AGAAACGATA	960
	ATATCTGTCA TATTGTTTAA TGCATCGGCT CTTACAGCTG CAGAGTTAAA GACAAAACCC	1020
45	GTGACATACT TAACAATAGA TAAGATTATA TATACAATTA AACTCAAATA AGCACCGCGT	1080
	TGCGCCAATT TAAGATTTTC ATTATGAGAC ATGCGTTGAA CCACCTTGAA TTAGTATAGT	1140
50	AACAATATTA TGAATGATTC ATTTTAATTT TACAACGTTT TTAATTTTTT TAAATTTTTT	1200
	TAAAATTAAA CTAATTTATT CATTGCAAC CCCTAAAAAT AATTTTTAGC CTTTCTGCGA	1260
55	ATTTTATGAG CTAGAAAGGC GCCCAACTCT CCCTGTTTGT TAACTTTTCG CTCGAAAGTT	1320

CGAAwTTAT GAGCTAGAAA GGCTTATGCA GTTGACGTTT TACGTCCAAC TGC GTTCCTC 1440
 CGTCTTCTTC AAATTTATTT GTnAGAAAGG CACCCAAC TC CCTGTTTG TTAAC TTTTCG 1500
 5 CCTCGAAAGT TTCTATGTTA GAACCTATG CATGAGTTGC GAANTATCTA ATGTCGTGAA 1560
 CTAATTATAT AGAAGAAAAA GTGCATCAAT GACAAATTAA ATGAGATTTC TACTCTACCA 1620
 AACTCTCTTC GAAAGACAAT TTTCTCCTCT ATTTATTAGC AACTATTGCA TTTCTCCATA 1680
 10 TAGTACTTCC TTA CT TAAAA TACGCTGAAT GTCTGAATTA AA 1722

(2) INFORMATION FOR SEQ ID NO: 408:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5521 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 408:

GCGAGACCCC CTGAGGGAGC AGTGCCAGT CGAAGACCCG AGGCTGAGAC GGCACCCTAG 60
 25 GAAAGCGAcC ATTyCAATAC GAaTTgTGat AAATAGAGAA CAGCAGTAAG ATATTTTCTA 120
 ATTGAAAATT ATCTTACTGC TGT TTTTTTA GGGATTATG TCCCAGCCTG TTTTTGTGA 180
 TTTTAAATAA TTTGAATATG GAAAATGTAT TATTCTCTCA TTTGTATAGA TTGTATTTAA 240
 30 TAAGTTAATG TAATCCTTGA GCTCAGGATT AATAAAATTC TATAACCTTA ATTATTTTCT 300
 CGATACAAAG GGT TATTAAAC TTTAATATAA GTATAATGAT GTGCCTCATC TTCAAGACGC 360
 35 ATTGTTGTAA CACACTTATA ATCTATAAAT GCGCGAACA TGGTATCTTT AATTTCaTTT 420
 ATGCGATTCT CATTTACTTG ATTAGATTG TGTGTCGAAA GTACAAGTTG ATCaAAAATG 480
 TTATCTAGTA CATCACGAAC GATATACCAC ATATGTCTTT CTAAGTTTGA ATCATTTGAT 540
 40 GCTTTAGAGA TTGTAAGAAT TAATTCGCCT AAATGGTTTT GAACGGTAGA ATAAAAGGCT 600
 TTGTTAAACA CAGACGTTTT TGAATCAGTA AGAATTCTTG ATTTTTCATG GAAATGAGAT 660
 GTACTGTATA CCATTTCAAT TAGTTGTGCT TTATCAATTC TTAAACCTTC AAAGTCTCTA 720
 45 ATATACATCG TGTCCAATAG ACCATCTTTC CGAAATGTTG CAATAGCATT TTGCAAGTGA 780
 GCCTCTAATG CAATGCCATA TTTAGTAACT AGTGGGATTA CGAGACCAAG CAATGCTTTA 840
 CTATAAGTTT CAACCCACGA TTTGCGCGAT GATTCAAAAT CAGACAATGA TGCAGCTGAT 900
 50 TGATAACGTT TAATCAATGT CACGATAGGT GATTCATTGT TAAATGGGTA GGTGCAACT 960
 AAGCTTGAAG GAATCAATGG TGTGACTTCT TGTGGAATCA TTTGGTATAT ATTTTTTCTA 1020

	AAAAATGAAT ACCAGCAACT TCATCAATAA TTGTTGATGC ATAGGACTTA AATATGACAT	1140
	CTTTCTCCAA AATATCATT AAAATACGTG TCATTAGTGG ACCATTGTGC GTCGTTTGT	1200
5	CTGATAATGT ACGAATCTCA CCTGTAATAT GAACGTTTGT CGACAATTG ATGTGTGGCG	1260
	ACATAGCTGG GTATTTAGGA ACTAATGTTT TGAAAGATAA ACCAGCATAA TAATCCaACG	1320
	TATGTTTTGC TTCAATGATT AATTCTTTAT CTACTTCTGC TTGATAATCA GAATGTAATA	1380
10	CGTCATCTAA TTGCCATGGA TGAACAATCA TAATGTGATA ATCATTAAAG TTAAACTTTG	1440
	GCGTAAATTC ATTTTCTAAT TGTTTAATTA AGTCCGGAAA TAGTTGATGA ACAGTTGTAT	1500
	CATAATCTTT AGACAGTGAC ATAGTACGGC TTAATTTACT GTGAATCAGT ACTATTTTCA	1560
15	ACTTAATAGG TTGATTGAAT TCTGAAGAAT ATAGGAATGT TTGTAATGCA TTTAAACCTT	1620
	TACGTAATTT AGCCCCAGGA TGTAGCGGAT GACCTTCAAT AACGGCTTGC TCTGAACGCA	1680
20	AGTAACTATC TTCGCTATTT TCGATAATAT TAAATAAAGG TGCAGAATCA TGTTGCATTG	1740
	ACAGTGCTTG ATAGCTAATT GCAAATGTCA TATtAGTtGC ACTGTTTATT AAATCTTGCT	1800
	GAAATTGATC ACTAGCAGCA TTTTTTAAAT CTGGTGCTTC AATTAAATA CACTCAAGAA	1860
25	TTTCATTGG ATGGTGTACT CGTGTAATCG TATTTGTAAT GTCATCTTTA ATGTAGAAAG	1920
	GGCCTTCAAC ATCAATTGCA TCAAAGCGT GTTCTCCAGT GATAGGAGCA TATAATGTTT	1980
	GCTTAGCTTG TGGGAAGTGG ATTTCTAGTA TATGAGTCGT TGAGATATCT AACATAATCA	2040
30	AATCAGGACT CAATATTTTC TTAATTTGAG TCGCGGCTTT AACTAAGTTT TCGCGATGCA	2100
	TTGATGTGAC CAATCTCTGA GTGACTTTAT CTCTTCCTTG TAAAATCATC TCTTTAAAAA	2160
35	TATTAGCCCA ATCGCTATTA TGTTGTTGTA AAAATAAATA TGTTTCTTGT TCTTCTTTGT	2220
	TAAATTTTAA TGTCTGTTCT TTAAAAATTA AGTTCAAGTT CATAATTCAC CTCTATGAAA	2280
	TATTTTACAA AAGCAAGATA GATTTGTATA ATCCATATTA ATGATAATGA YTCTTATTAT	2340
40	CAACAGAATG CGGGTGTAG TTTTATGACA AAATATTTT TTAGCAGTTC TTTTCTACTA	2400
	TTTCTAGGTA ATTGGATTGG ACAAAATAGGG CTAAATTGGT TTGTACTTAC CACTTATCAT	2460
	AACGCAGTTT ATCTGGGGAT TGTCAATTTT TGCAGACTTG TACCAATATT ATTACTAAGT	2520
45	GTGTGGGCAG GGGCAATTGC CGATAAATAT GATAAAGGGC GATTGCTGAG AATTACAATT	2580
	TCATCATCAT TTTTAGTAAC TGCAATTTTA TGTGTGCTCA CGTATAGTTc ACTGCAATTC	2640
	CAATTAGCGT CATTATTATA TATGCGACAT TAAGAGGGAT TTTAAGTGCG GTTGAAACAC	2700
50	CTTTAAGACA AGCAATCTTA CCAGATTTAT CAGATAAAAT ATCTACTACA CAAGCTGTmw	2760
	CATTTCAATC ATTCAATCATT AATATTTGTC GTTCAATAGG GCCTGCCATT GCTGGTGTCA	2820

55

	CAGTTTTATT ATGCTTACCA TTACATTTTA AAGTAACTAA AATACCTGAA GaTGCATCAA	2940
	GaTACATGCC GTTAAAAGTT ATTATAGATT ACTTCAAATT ACATATGGAA GGTTCGACAAA	3000
5	TATTTATAAC ATCATTATTG ATTATGGCGA CAGGTTTTTC ATATACGACA CTTTTACCAG	3060
	TTTTGACAAA CAAAGTATTT CCGGGGAAAT CTGAAATATT TGGTATCGCT ATGACGATGT	3120
	GTGCCATTGG TGGTATTATT GCAACGCTAG TTTTACCTAA AGTACTTAAA TATATTGGTA	3180
10	TGGTAAATAT GTATTATTTA AGTTCATTTT TATTTGGCAT TGCTTTGTTA GGTGTGGTAT	3240
	TTCACAATAT TGTCAATGTT TTTTATTTGTA TTACATTGAT TGGGTTATTT AGTCAATGGG	3300
15	CACGTACGAC AAATCGCGTT TATTTTCAAA ATAATGTTAA AGATTATGAA CGTGGTAAAG	3360
	TACTGAGTAT TaTTATGATG GgATAGAGGT ATGaTTCCAT kGGGAAGTCn ATTAATGAGT	3420
	ATATGTGCAG ATGTGTTTGG CATTGTTAGA ACTTTTTCAA TAATGGGAAT AAGTACTATA	3480
20	TGCATTACAA TGGTATTCTA TTTTATAAAT AGAAAGTTGA AGTTAAAGTT GGAGGAAAGT	3540
	AATCATGGTA TATCTTGAAT GGGCAAAGGC AGATAGAAAT ATTCAATATC GTGTAATTAA	3600
	CGCCATTATT AAAGAACGTA TTTACCCCGA GCAAACATTT ATTTGCGAAA AAGGATCTTT	3660
25	AATTGAAATT CAGTATCATA TGCATGTGTT GACTATTGAA GTTGTTAGAA AAAGTGCATT	3720
	AGAACGCTAT GAGTTTACAG GTGATATTAC TTATTTAAAT AAAGGTGAAA CGTCATTAAT	3780
	TATAACTTTA GAAGGTTTAT TAGATGTGTT GAATCATGAC TTTGATATCC CTATTTTACA	3840
30	GCGACTACGC GAAGAGTTAA TACACAGTCG AGATAGTTTA GTTGAAACAT ATAAGCAAAT	3900
	GTCTCACAGA CAAACGTTAA TAAGTCmAAG TTTTAAATTT TCAAGGTTAC CACAAGATAT	3960
35	TAACTTTTTT TCakGGTtAC AACATGTAAA AGATAGTGAT AAGACAGATG ATTTAACTTA	4020
	TTCTGAGAGT TTGGTACCAG AGGGGCATCC AACACACCCT TTAACCAGAA CGAAATTGCC	4080
	CTTAACATATG GAAGAAGTAC GAGCATATGC ACCTGAGTTT GAAAAAGAAA TCCCTTTGCA	4140
40	AATTATGATG ATTGAAAAAG ACCATGTTGT GTGCACAGCT ATGGATGGTA ATGATCAATT	4200
	TATTATTGAT GAAATAATTC CCGAATACTA CAATCAGATT CGTGTGTTTT TAAAGAGTTT	4260
	AGGTTTGAAA AGTGAAGACT ATAGAGCGAT TTTAGTACAT CCTTGGCAAT ATGATCATA	4320
45	GATAGGGAAA TATTTTGAAG CATGGnTTGC TAAAAAATA TTAATTCCAA CGCCGTTTAC	4380
	AATACTTcCA AAAGCaACTT aTCatTTaGG ACGATGTCTT TAATTGATAA ACCATACCAT	4440
50	GTTAAGTTGC CCgTCGATGC aCAAGCAACA AGTGCCGTTA GAACAGTCTC AACTGTGACT	4500
	ACTGTAGATG GACCAAAGTT AAGTTATGCT TTACAAAACA TGTTGAATCa ATATCCaGGA	4560
55	TTTAAAGTTG CTATGGAACC GTTCGGTGAA TATGCAAATG TTGATAAAGA TAGGGCACGT	4620

AGTGCAAGTC TAGTTAATAA AAATCCAATA GATCAAAAAG TTATCGTGGA TAGTTACTTA 4740
 GAGTGGTTAA ATCAAGGAAT TACTAAAGAA AGTATTACGA CATTATTGA ACGATACGCT 4800
 5 CAAGCATTAA TCCCGCCTTT AATTGCTTTT ATTCAAAATT ATGGAATTGC TTTAGAAGCA 4860
 CACATGCAAA ATACAGTAGT GAACTTGGGG CCACATTTTG ACATTCAATT TTTAGTGAGA 4920
 GATTTAGGTG GTTCTAGAAT TGATTTAGAA ACATTACAAC ATCGTGTATC AGATATTAAA 4980
 10 ATTACAAATG ATAGTTTAAT AGCTGATTCT ATAGATGCAG TGATTGCAAA ATTCCAACAT 5040
 GCTGTTATTC AAAATCAAAT GGCAGAATTA ATCCATCATT TTAATCAGTA TGATTGTGTT 5100
 GAAGAAACCG AATTATTTAA CATAGTACAG CAAGTAGTAG CGCATGCCAT TAACCCAACA 5160
 15 CTACCACATG CAAATGAGTT AAAAGATATT TTGTTTGGAC CAACAATTAC TGTCAAAGCG 5220
 TTGTTAAATA TGAGAATGGA AAATAAAGTA AAGCAATATT TAAATATTGA GTTAGATAAT 5280
 20 CCGATAAAAA AAGAGGTGTA GTACTACATG GCACACGTTA ACATAAATAT ATCGAAGATT 5340
 AAATATAACG CCAAAGTACT TCAAACAGTT TTCAAAGTA AAAATATGCA ATTCACACCA 5400
 GTAATTAAGT GCATAGCTGG TGACCGTACA ATTGTAGAAA GCTTAAAAGC GTTAGGTATC 5460
 25 AATCATGTTG CAGAATCCAG ATTGGATAAC ATAATTAGTA TTGCAGATAC AGGATTTAAC 5520
 A 5521

(2) INFORMATION FOR SEQ ID NO: 409:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1261 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 409:

40 AGGTCGTCTr GCaGmAGmGm TATTTGACGA TCTTGCTTTC CCAAAACACG ATGATGATTT 60
 TAACATACTG TCTGATTATA TTGAGACACA TGGTGATTTT aCATTGCCAA TGTCTGTATT 120
 TGATGATTTA TATGAAGAAT ATACGGAATG GCTAAAATTT TAATATAATT TTTAATAATA 180
 45 ATAGTTAGAA CCAGGGTGAT GCAATTCGTT ATCCTGGTTT TAATTTAAAA TAACTAAGT 240
 TGTGACTAAA AATTAATCAA TTATAGTGAA ATATGGTGCG CTATCTTGCA TAAATTGATA 300
 TGATTAAC TA CACAGAATTT AAAAGTACAT AATACATAAT AAGGAAGTGA TACAATGGAT 360
 50 GATAAGCAAC ACACATCTTC ATCCGATGAT GAACGCGCTG AAATTGCAAC AAGCAATCAA 420
 GACCAAGAAA CTAATTCATC GAAACGCGTT CACTTAAAAC GTTGGCAATT CATATCAATA 480

CAAAAAATAA GTGGTTTAAA CAAAAGTAT CAAGCAAAGT TAAATAAAAT TGAAATGTG 600
 TATAAATCT TAAATAGTGA TTATTACAAA AAACAGGACT CTGACAAGTT AAGTAAAGCT 660
 5 GCAATTGATG GCATGGTCAA AGAATTAAAA GATCCTTATT CTGAATATTT AACAAAAGAA 720
 CAAACGAAAT CCTTTAATGA AGGTGTTTCA GGTGATTTTG TAGGTATTGG TGCAGAAATG 780
 CAAAAGAAAA ATGATCAAAT TATGGTTACT AGTCCTATGA AGGGATCTCC AGCAGAACGT 840
 10 GCTGGCATTG GTCCTAAAGA TGTCATTACT AAAGTAAATG GAAAATCAAT TAAAGGTAAA 900
 GCATTAGATG AAGTTGTCAA AGATGTTTCG GGTAAAGAAA AACTGAAGT CACTTTAATC 960
 GTTCAACGAG GTAGTGAAGA AAAAGACGTT AAGATTAAAC GTGAAAAAAT TCATGTTAAA 1020
 15 AGTGTGAGT ATAAGAAAAA AGGTAAAGTT GGAGTTATTA CTATTAATAA ATTCCaGAAT 1080
 GATACATCAG GTGAATTGAA AGATGCAGTT CTAAGAGCTC ACAAAGATGG TTTGAAAAAG 1140
 20 ATTGTTTTAG ATTTAAGAAA TAATCCAGGT GGAAGCTAG ATGAAGCTGT TAAATGGCA 1200
 AATATTTTTTA TCGATAAAGG AAAAAGTGT GTTAAAGTAG AAAAAAGTAA AGATACTGAA 1260
 G 1261

25 (2) INFORMATION FOR SEQ ID NO: 410:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2488 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 410:

35 AAATATATTG AAnAGAnAAT TACTAAGATT AAATCnTCTT AAAATATCCC TGAAATAACG 60
 TCCTAAAGAT TAAAGGAAAG AGGTTATAAG TTATGCCAAA ATTAATTTTA TGTCGTCATG 120
 40 GACAAAGCGA GTGGAATGCT AAAAAGTTAT TTACTGGATG GGAAGATGTT AATTTATCTG 180
 AACAAAGGTAT TAATGAAGCG ACTAGAGCAG GTGAAAAAGT AAGAGAAAAT AACATTGCCA 240
 TCGATGTAGC TTTTACATCG TTATTAACAC GTGCTTTAGA TACAACGCAT TATATTTTAA 300
 45 CTGAATCTAA ACAACAATGG ATTCCTGTAT ATAAAAGCTG GCGTTTAAAT GAACGCCACT 360
 ATGTTGGATT GCAAGGCTTA AATAAGATG ATGCTAGAAA AGAATTTGGA GAAGAACAAG 420
 TACATATTTG GCGTCGTTCT TATGATGTGA AACCACCTGC TGAAACCGAA GAACAACGTG 480
 50 AAGCTTACTT AGCTGATCGT CGATATAATC ATTTAGATAA ACGTATGATG CCTTATTCTG 540
 AAAGTCTGAA AGATACTTTA GTTCGAGTGA TACCATTTTG GACAGATCAT ATTTACAAT 600

55

	TTAAATATCT TGAAGATGTG TCAGATGAAG ATATCATTAA TTATGAAATT AAAACAGGTG	720
	CACCGCTTGT TTATGAATTA ACGGATGATT TAGAAGTTAT AGATAAATAC TACTTATAAA	780
5	AaAAGAGCTG CATGTACACA AGCAGTGAGT GTATATGcAG CTCTTAAAtA TGTGAAGTAA	840
	TGTAAGGAAA TAGTTAAGTA TAGAGTTTAT ATTAACGAGC TAGGGATACT CGAAAATATA	900
	GTTAGACATA CAATATAGTC AAATTAAAAC AATTATTTTCG CTCTTTTATG TTGCTTAATA	960
10	ATCTTTAAAG CACGCTTTCT TGTTTTAAATG TTAGGGCTAT TTAAATTACG ACCAGCAGTC	1020
	TGTAAATCTA ATTTTCATCTC TATCCCTCCT TGTAATATA TTATGACCGA TAACTACTCA	1080
	TATGTAAATA GTAATGATTA CGTTTTAAAG AAATTGTAAT AAAGTCGTGC TAATTTTTTG	1140
15	GAAAATGGGT ATAATTACCG GATATCTAAA AATGTGTGTC GTTTTTTAGA TGGTGAGGGG	1200
	GAAGCTTTAA ATGTCGAAGA AACAAAAATT AACGATGATT ATTACTATGC TGATGGGTGG	1260
20	ATTTTTTGGA TTATTAAATG AAACACTATT AGTGACGGCT TTACCAAGTA TTATGAAAGA	1320
	TTTTGAAATT TCATATACAC AAGTTCAATG GCTGACAACA GCTTTTTTAT TGACTIONG	1380
	GATTGTTATT CCTTTGTCCG CGCTTGTTAT ACAACGTTAT ACAACAAGAC AAGTGTTTTT	1440
25	AGTGGGTATT TCTATCTTTT TCTTAGGTAC ATTACTCGGC GGCTTGAGTC CGCACTTTGC	1500
	AACATTATTA GTTGCTAGAA TTATTCAGGC GTTAGGCGCA GGTATTATGA TGCCATTGAT	1560
	GATGACAACG ATTTTGATG TTTTCCAACC ACATGAACGC GGTAAATATA TGGGGATATT	1620
30	TGGTTTGGA ATTGGTTTAG CACCAGCTAT TGGACCTACT CTTTCAGGTT ACCTTGTTGA	1680
	ATATTTTAAC TGGAGATCGC TTTTCCATGT TGTGCTCCA ATTGCAGCTG TGACATTTTT	1740
	AATTGGAaTTT AAAAcGATAA AAAATGTTGG AACTACAATT AAAGTACCTA TTGATTTTTAT	1800
35	TTCTGTCATT TTTTCTGTAC TAGGTTTCGG cGGGTATTG tATGGAACGA GTTCaATTc	1860
	AGAAAAAGGT TTTGATAATC CTAcGtATTA GSTATCTATGA TTGGAGGCGT TGTTTTAGTC	1920
40	GCATTATTTG TAwTACGTCA ATATCGGCTA TCAACACCAT TATTaAATTT TGCTGTATTT	1980
	AAAAATAAAC AATTTACAGT TGGTATCATT ATTATGGGTG TCaCAATGGT ATCGATGATT	2040
	GGTTCGGAAA CGATTTTACC TATCTTTGTG CAAAATTTAT TGCATCGTTC AGCTTTAGAT	2100
45	TCTGGATTAA CTTTATTACC AGGAGCAATT GTTATGGCAT TTATGTCGAT GACTTCGGGT	2160
	GCTTTATATG AAAAGTTTGG TCCTAGAAAT CTTGCTTTAG TAGGTATGGC GATTGTTGTT	2220
	ATTACTACGG CTTATTTTGT TGTAATGGAT GAACAAACAT CAACAATCAT GTTGGCAACA	2280
50	GTTTATGCGA TTCGAATGGT TGGTATCGCG TTAGGATTAA TTCCAGTAAT GACCCATACG	2340
	ATGAATCAAT TAAAGCCAGA AATGAATGCA CATGGTTCAT CTATGACAAA CACAGTACAA	2400

55

AACTTTTCTC CAACTATGTC AGACTATA

2488

(2) INFORMATION FOR SEQ ID NO: 411:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1105 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 411:

15 TTTTACAATT TCAGATATCT CTAAACACA TACCGATCCA ATGGCCAAAA TTATTCGTCA 60
 AAAATTGAAA AACTAGGAA TTCGTAAAGG GATTCCAGTT GTATTTTCAG ATGAAAGTCC 120
 AATTGTCATA AGAGAAGATG TAAAGATAT AGTTGGAGAT AAAAATGCTA TCAATCGAAA 180
 20 AGGACAAATG CCACCTTCTT CAAATGCCTT TGTGCCAAGT GTTGTGGAT TAATTTGTGC 240
 AAGTTATGTG GTGAATGACG TATTAAAAGA TATTCCAGTT CGTCGCATTA AAGACAAAGG 300
 GCAATAATTC ATTTTGAAAG GGATAATTTT CAACGTAAGG CAAGTGTAAC CACACCATAA 360
 25 AACTAAAAC TACTAGTTCTG CAAGCATAGT AGAATATGCC TCGTGACTA GTCAGTTTTG 420
 ATTTGATTAT AGCTAAAATA TCAAAGGTAA TCGATTTAAT CTTATTTTTC CGGCCCTTGT 480
 TTAAGTAAGT TGTTATATAT TTCTTTGAGT TGTGTTCAC TTTTAGATGT CGTTTTTGGT 540
 30 TCGTAATAAA TTTTGTTTTT TAGTTTATCA GGTAAATATT GCTGTGAAAC ATAGCCATTA 600
 ACATATTGGT GTGGATATTT GTAACCAATA GATCGGCCTA GATCTTTAGC ACCTTGATAA 660
 TGTCCATCTT TTAAATGGTT TGGAAATTTG CCCACATGAC CGTTTCTAAT ATCGGACAAT 720
 35 GCACTATCAA TGGCACTCAT TGCTGAATTT GACTTAGGTG ATAAGCATAG TTCGATTACT 780
 GCTTGGCTTA GTGGAATTCT AGCTTCTGGT AAACCTAGAC GTTCTGCTGA TTCAATAGCA 840
 GCAAGTGTTT TCTGACCAGC ATTAGGTGAG GCTAAGCCTA TATCCTCATA GCTAATTACA 900
 40 AGTAATCGTC GAACTATTGT AGGTAAATCT CCAGCTTCAA TTAATCGTGC TAAATAATGT 960
 AAAGCGGCAT TGACGTGCT ACCACGGATA GATTTTGGGA AAGCGCTCAT AACATCGTAA 1020
 45 TGCATGTCAC CATCCTTGTC ACTTACAAAT GCACCTTTTT GTAAACAGTC TTTAGCATCT 1080
 TGCAATGTAA CATGTCGATA ACCGT 1105

(2) INFORMATION FOR SEQ ID NO: 412:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 579 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 412:

5 TAACGTCGAA CTTGAGCTGT TACGTTATGA CTCATAATTA TTTTAGCATA GTCATTTAAA 60
TAAACTTCTG TTCTGTCTGT TGGATAAGCA AATTCAAGCA ATTGACTGTA ACTATCATT 120
ATGACTTCTT GATCAACATG ACTATCAAAA TATACAGCAT AATAATAAGT ACCATCAACC 180
10 ATATATAACA AATCTTCAAA CTCTGTAGTT ATTGGATTGC TATGATATGC ATAATTAATA 240
ACATCTTCTA AATCGTTAAA TTTCACAATG ATTGTTCTTG TATTTTACG TGCTGAAGAC 300
TTTTGACGTT TAGAACCTTG AGCTTCTTTT TCTTTTGTTT GTTGCTCGAA TAATTCTTCT 360
15 AATTGATCTT CACCTTCTAA TGTGTTGAGCT AACAATTCTT GAACTTGTTT ATCAAATKGA 420
TCAGTTGCAT CATCATCAGA CATATTCATC ATATCTTCAT TTTTAGATTT AGAAATTGTG 480
ACTTCGACAC CTTTTTCAAA GGCATGTACT TGAATCCATA ATGGACCTnC TACAACAAAA 540
20 TCTTCTACTT CGTTAATTTC ATCCATCATT GAnCAAAAG 579

(2) INFORMATION FOR SEQ ID NO: 413:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 1342 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 413:

GTGGTGAAAC TTCTTGCTTT GTAATTTTAT AAAGTGGATC AATATCACTC TGACTTACAT 60
35 CAGGTTGAAC TGTATCTTT TTAGTTACTT TGTGTTCAAG CGTAATATTT GCTAATTCTA 120
GGTTTACAG AATAATGAAT GTCATCAGCT AmCCCTTTmm CtTGATACTy CTCTGATAT 180
TTCCAGCTT TAGCATTGA TAAATCAATC ACTACTCTTA AATCTTCTGG ATTTTCAATT 240
40 TTTATTATCT TTGATTGTGG TCCTGAAATA GTCACATTAA CTGTTTCAGG CGCTTTTGTT 300
AAATGCAAAT CTTTAGTGTT ATAAAGAATT TCAACGGGTA CATCTTGAAT CGTTTACTA 360
GACTTTTGAC CAAGATTACC AGTGTTAAAG ATATTTCCAA AAACATTGTT AACAGATAAA 420
45 AAGAAAAACA ATGCCAAAAG AAAGGCAATA AATCTCAAGC CCCATTTACT TTCTAGCATA 480
TTATTTCACT CCTTCTTTT GAAAGCGTGT GCCAAACCAA TGTTGAGCAA GCAACTCTTC 540
AAAAATTTCG TTTGAAATGT CTCGTCGTAA TTTTCCATCA AATGTTACCG AAATATCACC 600
50 AGTTTCTTCA GATACAATAA CGGTAAATGC ATCAGATACT TCTGAAATAC CAACCGCAGC 660

5 TGCTGCTGCT GCAATCTTCG TGCCTTGAAT AATCATTGCA CCATCATGTA AAGGTGTGTT 780
 AGGTATAAAG ACATTAAATTA AAAGTTCTTG CGAAATATTT GAATCCATTG CAATACCTGT 840
 10 TTCAATATAA TCTTGAAGAC CTGTTTCTTT TTCAAAGACA ATTAATGCAC CTATACGTCT 900
 TTTAGCCATA TATTGCACAG CCTTTGAAAC CGATTGAATC AATTTCTCTT CATCTTTACT 960
 ATACGTATTA GAAGTATAGC GTTTTAAAAA GCTACCTCTA CCAAGTTGTT CTAACGCACG 1020
 15 TCTAATTTCT GGTGGAATA TTACTATTAA AGCTAATACC CCCCATTGAA TAACGATATC 1080
 GAATAATTTA GATGTGCGAG TCAAGTTCAA TATCATACTT ATCTGCTGAC CAATAACAAT 1140
 TACTAATATC CCTTTAAGTA ATTGTATCGC TTTAGTTCCC TTAAAGACCG TGATGAGAAG 1200
 ATAAAGTACA TACCAACTA TCAGTAAATC AAGGATACTC GTTACAATTT TTAACGTACT 1260
 GAGGTTTTGA AAAAAGTTGG AAAATCCAT AACATCTCCT CCGGTAATA TTTTCCATA 1320
 20 ATACCCATTA TACCAATCAT TT 1342

(2) INFORMATION FOR SEQ ID NO: 414:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1073 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 414:

CTGTATAAAG ATGGAGGTGT TTTGCATGGT AAAACGTACT TATCAACCAA ATAAACGTAA 60
 ACATAGTAAA GTTCATGGTT TCAGAAAACG CATGAGCACA AAAAATGGCC GTAAAGTTTT 120
 35 AGCGCGCCGT CGTCGTAAAG GCCGTAAAGT TTTATCTGCA TAAGATCACT GACTCATCAG 180
 TGAtCTkTTT TTTCGTTTAA ATTAAGAATA AATAGAAATT TATGTTATAA GCTCAATAGA 240
 AGTTTAAATA TAGCTTCAnA TAAAAACGAT AnATAAGCGA GTGATGTTAT TGGAAAAAGC 300
 40 TTACCGAATT AAAAAGAATG CAGATTTTCA GAGAATATAT AAAAAAGGTC ATTCTGTAGC 360
 CAACAGACAA TTTGTTGTAT AACTTTGTAA TAATAAGAA ATAGACCATT TTCGCTTAGG 420
 TATTAGTGTT TCTAAAAAAC TAGGTAATGC AGTGTTAAGA AACAAGATTA AnAGAGCAAT 480
 ACGTGAAAAT TTCAAAGTAC ATAAGTCGCA TATATTGGCC AAAGATATTA TTGTAATAGC 540
 AAGACAGCCA GCTAAAGATA TGACGACTTT ACAAATACAG AATAGTCTTG AGCACGTACT 600
 50 TAAAATTGCC AAAGTTTTTA ATAAAnAGAT TAAGTAAGGn TAGGGTAGGG GAAGGAAAC 660
 ATTAACCACT CAACACATCC CGAAGTCTTA CCTCAGACAA ACGTAAGACT GACCTTAGGG 720

55

TTTAGATACA ATTACGAGTA TTTCAACACC AATGGGTGAA GGGGCAATTG GAATTGTTTCG 840
 ATTGTCTGGA CCGCAAGCCG TTGAAATTGC TGACAAATTA TATAAAGGAA AACATCTTTT 900
 5 AAATGATGTT CCATCACATA CGATTAAATTA CGGTCATATT ATTGATCCAG tCTAAAGAAG 960
 TGGTTGAAGA AGTTATGGTG TCTGTGTtAA GAGCGCCAAt AACATTtACA CGCGArGATA 1020
 10 TTATAGaGAT TAATTGTCAT GGTGGtATTk TAACGATTAA TAGAGTGCTG GGA 1073

(2) INFORMATION FOR SEQ ID NO: 415:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3176 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 415:

CTTTACCAAT GCCAAATCCG AAGTAAAGTA TAGCAATAAA GATTACTAAT ACAATTCTGT 60
 AAATGGCAAA TGGAATTAGT TTGATTTTGT TAATTAGATG CAAGAATGTT TTGATTGCAA 120
 25 TTAGTCCAAC AGTAAATGCA GCTAAAAAGC CTAAATATA AAAAGGTATA TCAGCAATCT 180
 GAATATCTTG ATAATGTTTT AATAAAGATA AACCCTAGC TGCTAACATA ATTGGAACAG 240
 CCATAATAAA TGTAAGTCC GATGCTGCTT TATGATTAA TTTCAATTAAT ACCCCAGTTG 300
 30 AAATTGTTGA GCCTGAACGG CTGAAACCAG GCCACATAGC TACTGCTTGA gAAATACCAA 360
 TTACAAATGC TTGGAAATAA CTGATTTGAT CTACTGTTTG TGGGTTTTTA ACTTTAGCTG 420
 AGTATTTATC AGCAATAATC ATATAGATAG CACCTACGAA TAAGCCAATC ATAACAGTTG 480
 35 GCACACTAAA TAAATGTTCT TCGATGAAAT CATCAAATAG TAAGCCTAAA ATACCTGCTG 540
 GCACCATACC CACTAATACA TGTAAATAAT TTAAACGTCT TGGCTTTGAA CGTCTTTGTT 600
 GATCGTTATC TCCTTCAACA TGTTTGTGTT TACCAATATG TAAATCTCT AAGAAGCGTT 660
 40 CGCGGAACAC CCATGCTGCT GCAAAGACGG ATCCTAATTG GATGACGATT TTAAATGTAA 720
 ATGCTGACTG AGAACCTAAA AATTCAGATG ATTTTAACCA CATATCATCA ACTAGGATCA 780
 45 TATGTCCAGT AGAGGAAACA GGTGCAAATT CTGTTAATCC TTCGACGACC CCTAAGATAA 840
 TACCTTTTaT TAATTCAATG ATAAACATAA TGTACCCACT TTCATTACTC AATTTAATTT 900
 ATTTAAATAT CAAAATTACC ATATCATGAT AGCATATTCA TTTAAAGACA TGCTAGTTAT 960
 50 AGTTATAATA CTAGACTAAA GATGTATATA TTCATTTTCT TTTACATGTA AACTACAAT 1020
 ATTTTATTGA GCTATTTAAT TTGATTTTAA GGAAAACCTT TTATAATAGG TTTAGGTGAT 1080

	TTCTTGGTCA GTACTGGTCT CGGCATAATC GTTATAACGC AAAATATTTT AATAGCAGAT	1200
	TTTTTAGCTA AAATTATAAG ACATCAATTT CAAGGTTTAT GGATTGTATT ATTTATTTTA	1260
5	TTAGGTGTTT TACTTTTAAAG AGCAACTGTG CAATTCTAA ATCAATGGTT AGGTGATACA	1320
	TTAGCATTTA AAGTTAAGCA TATGCTTAGA CAGCGGGTTA TTTATAAAAA TAATGGTCAT	1380
10	CCAATCGGTG AACAAATGAC TATACTCACA GAAAACATTG ATGGTCTAGC ACCTTTTTAT	1440
	AAGAGTTATT TGCCTCAAGT GTTCAAATCA ATGATGGTTC CGCTCATCAT AATCATTGCA	1500
	ATGTTTTTCA TCCATTTCAA TACCGCATTa ATTATGTTAA TAACTGCACC ATTTATTCTT	1560
15	TTGTTTTATA TTATTTTCGG TTTGAAAACG CGAGATGAGT CAAAAGATCA AATGACTTAT	1620
	TTGAATCAAT TTAGTCAACG GTTTTTAAAT ATTGCTAAAG GTTTAGTGAC GTTAAAGCTA	1680
	TTTAATCGTA CAGAGCAAAC AGAGAAGCaT ATTTaCGACG ATAGTACTCa GTTTAGAACT	1740
20	TTAACAATGC GCATTTTaCG CAGTGCTTTT TTATCGGGAT TAATGCTCGA ATTTATAAGT	1800
	ATGTTAGGTA TTGGATTGGT TGCATTGGAA GCAACGCTAA GCTTAGTAGT ATTTCATAAT	1860
	ATTGATTTTA AAAGTGGCGC AATTGCGATT ATTTTAGCGC CTGAATTTTa TAATGCAATT	1920
25	AAGGACTTAG GGCAAGCGTT CCATACTGGA AAACAAAGTG AAGGTGCCAG TGACGTTGTG	1980
	TTTGAGTTTT TAGAACAACC GAACTATAAT AATGAATTTT TATTAAAGTA TGAGGAAAAC	2040
	CAAAAGCCAT TTATTCAGTT AACAGACATA TCATTTTCGAT ATGATGATTC TGATAGATTG	2100
30	GTATTAAATG ATTTAAATTT GGAAATATTT AAAGGTGATC AAATTGCACT TGTAGGTCCA	2160
	AGCGGGGCAG GTAAATCCAC TTTGACACAT CTTATTGCAG GTGTTTATCA GCCAACAATA	2220
	GGTACTATAA GTACAAACCA GCGTGATTTA AATATAGGAA TACTTAGTCA ACAGCCATAT	2280
35	ATTTTCAGTG CTTCTATAAA AGAGAATATT ACGATGTTTA AAGATATAGA AAATAATACT	2340
	ATTGAAGAAG TGCTAGACGA AGTAGGTTTA TTAGACAAAG TGCAATCTTT CACAAAAGGC	2400
40	ATTAACACAA TAATAGGTGA AGGAGGCGAA ATGTTATCTG GTGGACAGAT GAGACGCATA	2460
	GAACTTTGCC GTCTTTTAGT TATGAAGCCA GATCTCGTTA TATTTGATGA GCCTGCAACT	2520
	GGTTTAGATA TTCAAACAGA ACACATGATT CAGAACGTTT TGTTCACAA TTTTAAAGAT	2580
45	ACAACGATGA TTGTCATTGC ACATAGAGAT AATACAATTC GCCATTTACA ACGACGCTTG	2640
	TATATAGAAA ATGGAAGACT GATTGCTGAT GATCGCAATA TTTCAGTAAA TATAACAGAA	2700
	AATGGTGATG ACTTATGAAA ACACGACTAA AATTTCAGT AGATAAGGAT TTATTGTTAG	2760
50	CTATAGTTGT TGGTGTTTGT GGAAGTTTAG TTGCGCTCGC CATGTTTTTC TTAAGTGTTT	2820
	ATATGGTGAC ACAAAGTGCA CTTGGTGCGC CACTATACGC TCTGATGATT TTAGTCGTTA	2880

55

ATAAAGCTAC ATTTACAATG CTACGTGATA TTCGGGTACA GTTTTTCGGT AAATTAGTAA 3000
 ATGTCATTCC TAATGTTTAC CGTAAACTGA GTTCTAGTGA TTTAATTTCA CGTATGATTA 3060
 5 GTCGTGTTGA GGCATTACAA AATATAKATT TACGTGTTTA TTATCCACCA GTCGTCATCG 3120
 GTTTGACAGC GCTAGTTACA GTCATAGTTT TGGCGTTCAT TTCAATCGGC CATGCG 3176

(2) INFORMATION FOR SEQ ID NO: 416:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2109 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 416:

20 TCTTTATTTT AATTTCCAAT TGAATTTTTT TATTATTTAC GCATAGCTCT TAAAATTAAC 60
 GTTACGATTG CAATTAAGAT AATTGAACCA ATTAATGCTG GCAAGATGTA AATACTTCCT 120
 AATTCAGGAC CCCATTGTCC GAATAGTTTG CCACCTACCC ATGATCCAAT AATACCTGCG 180
 25 ATAATATTGC CTAAATACC ACCTGGGATA TCTTTACCCA TAATAGCACC AGCAGCCCAT 240
 CCAATTAAGC CACCGACAAT TAACATTCCA ATAAATCCAA ACATAATTTT CAGTCTCCTT 300
 TTTCTATTTA TTTTGCCTTA TTCTAAGTAG TACCCCTTAT TTACAATTCT AAAACAAATT 360
 30 CAAATTATTT TTATCCAAAT ATTTTAAAA GTAGTAATTG AATATCAATT TTATTCAATG 420
 TAGCTATCGT TATTTAAAGT CTCTGTACCG ATAATATCAT ATACATTTAC ATTATTTTTT 480
 CTGCCGAATT CATAGCTTGA TTATTTTATG TTATAGGACT AGAATATACA CATATTATTA 540
 35 GAGCATCTTT GAATTTTAAA TCAAGAAGCG AGGTAAATGA ACAATGAATA TGCATATTTT 600
 ATATAACTTA CGAACTAAAC ATAATTTAGA AATTGACGAA TTAGCACAGC AATTAAATGA 660
 GAAATATGGT ACTAAATATG AAGCACATCA AATTTGGGAA TGGGAGAATC ATCACCATGA 720
 40 ACCTAAATTT AAAGATGCCA TGCATTTAGC TGACTCTTTT GATGCACCAT ATGAAATGTT 780
 TTTAGAAAGT AAGGTAAAG AATATCAGAA ACATTTAGAA GAAGTCGATA TTCGCATGGA 840
 TAAATAGATG CAAATAAACC CTCACAACAC GTTTGGCATA TATCCTTTCA AATCTATACT 900
 GGATATATTA CATTACGTTG TGAGGGTATT TTATTAATTA ATATGAATTA AGACATTTTA 960
 CAAGCGTTAA TGCAACGAAT CTTTTTAGTG ATCTTGCTCA CTCTTTAATA CTTTACCGTT 1020
 50 CTTAGCATCA ACAGTAACTT CTTGTTTTTT ATTACCTTTT TTCAAATCGA TATTGTAAAC 1080
 AAGTTTGCCA TCATCTTTTT CAAGTGACCA TTCTTTAATA TCACCATCAA ATTCTTTTTG 1140

5
 10
 15
 20
 25
 30
 35

ATTCAATTGTA TCTTCTTTTT CAGTCTTTTT GTTAATCACT TTTTATTTT TATCAGCAAC	1260
AAGTACTTCT GACTCTTCAC CAGATTTTIG TTGCGTCACT TTATAAGCCC ATTCACCATT	1320
AGAATTTTCA AATGAAATTC CTTTCAACTT TTGGCCTTTG TAAGTTTCTT CAGCTTTTTT	1380
CACAGCATCT TCTGGGCTTG TTTTAACATC TTTTAAAGCA ATAACATCTT TTGTTGTATT	1440
AGTGTCTTGA TTAGTATTG ACTCTGTTGA TTTTGTTTCA TCTTTTGGAG TATCATTGCC	1500
ACATGCAGTA AGCACCACCG CTGACATTGA TAACACTGCT AATGATTTTA ATTTCATAAT	1560
ATCACTCTCT CTTCTATTT TTGAACTCA TAACAAAAGC TTATATGCTA TATAGATTGT	1620
ATTACCCCTT GTTTTTAATT TTATTCATAA TTATTACAA TATTTTAA TTAATCGTCA	1680
TGCGTTACTT TCGTTCGTAC TCCTTTTATA AATGAACCAT GTAATATAAG CATGCTATTA	1740
TCGACTCAGT TTGTCTAATG CTTTCTTTGG TACTTCTTCC TTTTCAACTT CTTCAAAAGT	1800
TTCTACATGA TGACCTTTAT GTGTAATTTT CAAGTATCTA TGCGGTTTAA CATCAAATGT	1860
AGCAGTATAC ATTAATTGCG TCTCTTTCCC TTCTTTATTA AAAACACTTT GACTATAACT	1920
TCGGAGTTGA TCATCCATTC TAGTAGACAC CTCTGTCGTT TTAACATAAG AATCATCCTT	1980
CTTTACTAAT GGGTTAAATT GATCTGTTAT ACCATGGGAG TCTATTGTTT TTAATAATGAA	2040
TAAAGCAGCA TAAGCGCCAA TGATAGTAAG TACAAGATA GTKATTGTTT TTAATAATCGT	2100
TTCAAAAGA	2109

(2) INFORMATION FOR SEQ ID NO: 417:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 813 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 417:

40
 45
 50
 55

GTTACATAAA TGAAACAAA AAAGATAATT TTAGTGCTTA TGCTACACCA GAACATAATT	60
ATCAATTTGG TGGTGCTATG ATAGAAAGTG AAAAATTAAG CGAGTTACTA AAGCCAGCCA	120
ATCAGTTAAA ATCACCAGAT GATATAAAAA AAGAACTAAA TAAAAAGAAG AGTCACTAAA	180
GTTAGGAGTT ACTTTAATGT CCAAAAAACA TGTTTTTATA ATTATTGGTG TCATATTGTG	240
TATATGTACA GTTCTACGG TCATGCATTT TAAAATGAAA TATGATGAAA AAGAAAAACA	300
AAAAGCGATT TACTACAAAG AACACAAGA ACGTATTACA CTCTATCTTA AGCATAATAC	360
TAAAGAAACG AACACGATTA AATCTGTACA TTTCAAAAC TTGGAACAA GTCCTATGGG	420

ATCGCCTGAA CATAATTATC AATTGGTGG CGCTATGATA AAAAGTGAAG GAGTAGATAA 540
 ATTATTAAAA CCAGCACATG AAAGAAAATC ACCAGAAAAA ATCAAAGAAG AATTAGATAA 600
 5 AAAAGAAGGC CACTAGGGTC TTCTTTATTT TTGATTTAAT CTTCCAATAA TCTATGTCAT 660
 TGCTATCGAA GGTGTTTCGC AATTAATATA AATCACTTCA TCATCACCAA TACTTCCCCA 720
 10 GTTTTGTACA GTACATTAAC ACAACGAAC CACGTTAATT TAAATGGAWT AaTAGTTTGG 780
 CCATTATAAG AACAATATAT ATCGANTAAC AAT 813

(2) INFORMATION FOR SEQ ID NO: 418:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 640 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 418:

AAAAGCAATC GTTGGTGGTG CTAAATTCAT CGGCAACTCA TATGTAAAAG CTGGTCAAAA 60
 25 TACACTTTAC AAAATGAGAT GGAATCCTGC ACATCCAGGA ACACACCAAT ATGCTACAGA 120
 TGTAGATTGG GCTAACATCA ATGCTAAAAT CATCAAAGGC TACTATGATA AAATTGGCGA 180
 AGTCGGCAAA TACTTCGACA TCCCACAATA TAAATAAGCA ACATGAACAT AGGATCAAAA 240
 30 GTCATCCCCC ACTATCAATC ATGGGGGATG ACCTTTGATC CCTTTTTTAT ACATACACAA 300
 GCAAAAATAG CGGTGATTGT TTACCATCAA TTTTAACAAT CACCGCTACT TTTGCTTGTA 360
 ATTCATGATT CAATTTTTGT TGTGTGCACA ACGACACTAA ATTATGTGTT TGCTATTGTC 420
 35 GTGTTACAAC GATATGCGTC GTTGATTAA CTTATCAAGT AATTGATTAA AATTGTCTAA 480
 TTCGACTTCC GATAACATT GACATCTTGC TTCAATCAAT TCGCAACGTG CATTATTTAT 540
 40 TTGTGAAATT AATGTACGTG CTTGATCAGT CAAAATTAAT TCTTTACATC TTAAATCTTC 600
 GCTAGATTGT TGAATTGTGA TGTACCCnnt CAAAACATAAT 640

(2) INFORMATION FOR SEQ ID NO: 419:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1159 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 419:

55

TTCTTGGATA CGTAATAAAT TACCTTTTGT CATAGGATTA AACATAGATC CAGGTAAAAC 120
 ATAAACATTT CTATTTTGCT CTAATGCAAA ATCTATCGTG ATGTGACTGC CACTTTGTTC 180
 5 CTTAGCCTCA GTAATTAAAA CCCCTTTTGA CAAACCGCTG ATAATTCTAT TGGCTCAGG 240
 AAATCTATAT TTAGCGATTG GTGTATGTGG TGGATATTCA GATATAACTA AACCTTTTTTC 300
 TTCTATTTTA TTTCTTAATG CTAATGTACT TTTGGGATAA TGTGTTTGAT GGCCAAAGGC 360
 10 TAAAACTGCA ATTGTAGGGA GATTGTATTT TAAAGCTATT TGATGTGCCA TTGCATCAGC 420
 TCCTTGAGCA AGGCCGGAAG CAATTGTTAA ATATTTGCTT TTATCATTTG ATAATAAAAA 480
 TTCTAAAGAC TGTGCGGTAT AACTTGTTAGA ATCTCTTGCA CCTACTACTG CCAAATGATG 540
 15 CATATTATTT ATTAATTTGA TGTTCCCTTT ATAGAAAAGA AGTAATGGAA ATTGATATAT 600
 TTCTTTTAAAT AGCACTGGGT ATTCAGAATC CATATATGTA ATGTAACCTA CTTTTAATTT 660
 20 CTGCAGTTCT TTAATAATAT CGTTATGATC AAGTTTTATA AAAGCATAGT ATTTACGTAA 720
 TAAATGAACA TTTTCTTCCC TATTCACCCA TTCACATAAA TAACTATCTT TTTTCTTCC 780
 CTCCTCTTTA ATTACATTAG GATATGCCAT TAAAAATTGA TGAATTGTT TAGTCGAAAA 840
 25 GTGTGCCCAG TATAACTTAA GCAAAAATAG TTTAATCAAT AAATCAACTC CTTTTGTAA 900
 TCATACAAAA TCATATTCTA TTTTGTGTTT ACATTTCTAA TACAAAAACA TTGTCGATGT 960
 AATGTTATTT TAAGGAGTAA AAATACTGAC TTAAAAAGTG AAAAGTATGT TGGAAAGAAT 1020
 30 TTAAAATTTT AATTTCCAAC ATACTTTATA ATTAAACCTT ATAAATAAGT TTTGCAATT 1080
 TATATAGAAT TGGTCTTACT GGTGTGATGA AATCACCAT TAATCTTCA ACATGTGCAT 1140
 TAAAACCCTT TTAAATTG 1159

(2) INFORMATION FOR SEQ ID NO: 420:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1879 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 420:

TCAATCAGCG ACTACAAGAA GTGCGGGTCT TCAACAATT GATGTGACAA CACTAAGTGA 60
 CCCCACTAAT ATTATCATGG GTATTTTAAAT GTTTATAGGA TCTTCGCCAA GTTCGGTTGG 120
 50 TGGCGGTATT CGTACAACAA CTTTCGCTAT TTTAATTTG TTTTAAATTA ACTTTAGTAA 180
 TAATGCCGAT AAAACATCCA TTAAAGTTTA CAATAGAGAA GTACACATTA TGGATATTCA 240

ATCAGCTACT GAAAATGGTA AGCTTACATT TTTACAAGTA TTTTTTGAAG TCATGTCTGC 360
 ATTTGGAAGT TGTGGACTAT CGCTTGGTGT CACAAGTGAT ATTAGTGATA TTTCTAAGGT 420
 5 CGTACTAATG ATACTCATGT TTATAGGACG TGTGGCTTA ATATCATTTA TCATTATGAT 480
 AGCAGGACGT CGAGAACCAG ATAAATTCCA TTATCCAAAA GAACGTATTC AAATAGGATA 540
 ATATAATAGC AATCTAAGTT TAGTTAATGT AGATTTTAAAC TGGAACTTAG ATTGCTTTTT 600
 10 TAGTTTGAT TTTTAACTTA TTTTATAAGA CGATTGGTTT CGAAAATGGT AAAC TAGTAA 660
 CAATGAGAGG TGTAACATGA TGGAAAAAAA TGAAACATT AATGTAGAGA TTTTAACTAC 720
 GTCAGATATG CATAGTCATT TCTTAAATGG TGATTATGGT TCAAATATTT ATAGAGCTGG 780
 15 TACTTATGTT AACCAAGTAA GAGCACAAAA TCATCGCGTC ATTTTATTAG ATAGTGGCGG 840
 AAGTTTAGCT GGCTCGTTAG CGGCCTATTA TTATGCTATT GTTGACCTT ATAAACGACA 900
 20 TCCAATGATA AAGTTAATGA ACAGAATGCA TTATGATGCT AGCGGTGTGA GTCCAAGTGA 960
 AITCAAGTTT GGTTTATCAT TTTTAACTCG TTCAATTGCT TTGGCACGTT TTCCATGGTT 1020
 ATCAGCAAAT ATTGAATACA ATGTTACTAA GGAGCCTTAT TTTTCAACTC CATATTGTAT 1080
 25 TAAACATTTT GGTGACTTAA AAATTGCTAT CGTAGGCGTC ACAGCAGATG GTTTAATGGA 1140
 AAATGAGTAT TCTGAAATGG AGCAAGATGT ATCTATTGAA AAGACATTAG TGGCATCAAA 1200
 ACGTTGGATT AGATATATCC ATGAAGTTGA AGAGCCAGAT TTTTGTATTG TAATTATCA 1260
 30 TGGTGGATTG AATAAAATTA GTAATAGTAC GAAAAATAAA AAGGCAAGTT CGAATGAAGC 1320
 TGAAAAATTA ATGGAAGAAC TCGGTGTTAT AGATTTAATG ATTACAGCTC ATCAGCATCA 1380
 AACAAATAGTA GGTCAAGATC ATGAAACGTA TTATGTTTCA GCTGGTCAAG ATGCCAAAGA 1440
 35 GCTTGATCAT CTTTCGATTA ATTTTAAAAA GAGAACAACA ACTTATGATG TTGAAAGCAT 1500
 TGATTCTAAA GTGATTGACT TAAATGAGTA TGAAGAGGAT CAAGAATTAT TAGATTTAAC 1560
 ATTCTATGAT AGAAAAGCAG TGGCTTATTG GTCACAGGAA ATCATAAGTG ATAAAGGTTT 1620
 40 GATGTTATCA GTAAATGGGT TACAAGATTT AGTCTGTCAA ACACATCCAT TTTGCAATT 1680
 ATTACATGAT GCAATTCACC TTGCATTTGA TAATGATATA ACATGTGTCC ACGTGCCTAT 1740
 GaACGGAGAG AAGGGGTTGA GTGGACAGAT TCGAAATGAA GrTTTGTATC aTGCATACCC 1800
 45 ATaTCCAGAT aAGCCATGGG tATGACAATT aGTGGTCAAA ATaTCAAGGT ATnTTGGGGT 1860
 ATAGTTATTC ACCATTAGG 1879

50 (2) INFORMATION FOR SEQ ID NO: 421:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2710 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 421:

	AATCCCCCT TTTCCCCAT CCATTTTnCC ATCCACTGGT CCTTACGGGA CCATATTATT	60
10	TnAAATTGGA nAAAATTATt TTAAAGaTTA TwACTACTCT TtAATCATTt TrGTGAATTA	120
	AAAAAAGTAG TGCAAAAAGC AAAATATACT TTATACACTA CAAATCATTt ATTTATAATA	180
	AAGTTTCACC AAAAAATGTT CCAACTAATG AAACCGCTTG TTCAGCAGTA TGATTATTAC	240
15	TGTCAATCAA TGGATTTACT TCAACTAAAT CCATTGAGGA AATTAAATGT GATTGATGCA	300
	GTAATTCCAA TGCAAAATGG CTTTCTCTAT AACTAAGACC ACCCAAAACT CTAGTACCAG	360
	TGCCTGGCGT TTCAAGCGGA TCTAAAGCAT CAACATCTAA AGATAAATGA ACGCCATCAA	420
20	CATTGCGTGA CTTCAAATAT TCTATTGTAT TTTCAATTAC TTCCTTTATC CCCAATTTAT	480
	CAATATCTGA CATAGTAAAT GTTTTAATAT TATGATCTTT GATAAATTGT CTTTCACCTT	540
	TATCTAAATC TCTCATACCA ATTAGTACGA TGTnTTTCTGG CTTGATTACA TTACnATTTA	600
25	ATTCTAAAAG TTCTTTGGGA CCTTCGCCTG TCAAAATCCT TAGAGGCATA CCATGAATAT	660
	TTCCACTTGG TGA CTCTTCA GGTATATTTA AATCACCATG TGCATCATAC CAAATAACAC	720
	CTAAATTATT ATAATGTTTA CTTATTGCTG ATACTGAACC TACCGCAATA GAATGATCTC	780
30	CACCAAGAAC TAGAGGAAAT CTGTTATTTT CAATTGATGC TGAAACCTCT TTATTTAATT	840
	TTTGATTAAC ATCTATAATT TCATCATAAT TTCTTAATCC TTTTGTTCa CTATGAAATT	900
	TTTCAATGTT CACAGCAGGt ACCTTAATAT CCCCCTTGTC ATATACATCA AGGTCTAATT	960
35	GCTTTAATCT TGAAATTAAT CCAGCATATC TAATTGCTGT TGGTCCTAAA TCAACACCTA	1020
	ATTTTCTTTG TCCAAATGTT GATGGTGCAC CTATAATATC AATTGCTTTT GTCTTTGTCA	1080
40	TAAGCGTCCC CCTTGCTCT ATGTAATTAA AGAATAATGT ATGCGCTTAC CATTATCAAG	1140
	CAATAGCTAC ACATATAATC TGTTTATCTT ATTACTTCAT AAAAAAGGT TCTTCATCTT	1200
	TTATGGTGGG AAGGTAAAC TTCTGCTTT TTTTAATACA CAAAAAGCGC AATTGCCTCT	1260
45	ATAATTTAAA GTGACCAAAC CCAAACTAAA GGAGACAAGT GCGCCTATGT GTAATGATAC	1320
	CTTAGAATTA CTAAGAATAA AAGATGAAAA TATAAAATAT ATAAACCAAG AAATTGACGT	1380
	CATTATCAAA GGAAAAAAG CAACAGTGGT TAATGCTGTA CTAACGTATA AGCCTTCGGC	1440
50	CTGTTATTGT TGTGGAGTTA AAAATGAAGG ACAAATTCAT AAACATGGTA AGCGTGTTC	1500
	TCGTATTACT TTA CTTAAAA CTCAAGGGTA TAACACATAC CTCAACTTAG CTAAACAACG	1560

55

EP 0 786 519 A2

	GTGTTTTATC TCAAGATGTG TTA CTCAAAA AGTTATAGAA GAAGCTACTA AAGTTAAAAAC	1680
	AGAGATTGAT ACTGCAGAAG ATA ACTGTAT CTCTCCATCT ACTGTAAGTC GTATTAGAAC	1740
5	TAAAGCGGCT AATTCATTAC GAATTAAACC CTTTAATTGT TTGCCAGAAC ACATCGCTAT	1800
	GGATGAATTT AAAAGCGTTA AAAATGTAAC TGGATCAATG AGTTTCATTT TTATAGATAA	1860
	TGATACTCAT GATGTTATAG ATATTTT TAGA AAATAGAACT ACAAGATTCT TGC GTGCCTA	1920
10	TTTCGAGCGA TTCGATTTAA AAAATCGACA ACAAGTTAAG ACGGTTACTA TTGACATGTA	1980
	TGAACCCTAT GTCCGATTAT TTCGCGACCT ATTTCTTAAT GCAGCTATTA TTTT TGACAG	2040
	ATTCCATATC GTTCAACATT TAAATAGAGA ACTTAATAAG TATCGTGAC AAGTTATGAA	2100
15	TGAATACCGT AATAAAAAAG GACCTGATTA TACAATTTT AAGAATAACT GGAAAGTCCT	2160
	ATTGATGGAT ACTAGTAAAA CCATATTTAG TAAATACAGA TGGAATAAAT CTTTTAAGGC	2220
	TTATAAACGC TCATCTGACA TTGTAGAATT CATGCTTTCA AAAGACGATA TACTACGACA	2280
20	CTCCTACGAA CTGTGCCAAG GATTACGAAA AGACCTAAGG TTATGTAATT GGCCTAAATT	2340
	TATTAATCGT TTAAATTCAG TTAGTAAAAA GTCTGTGAGT AAGGGTGAT GGAAAGTGGT	2400
	TAAATATTAT AGAAAACATC AAAGGATGTT AAGAAATACA ATTTATTACC CAGCATTTAA	2460
25	TAATGGTGCT ATAGAAGGAA TTAATAATAA GATAAAATTA ATCAAGTGAA TTTCTTTTGG	2520
	TTACAGAAAT TTCAACAACT TTAAAGCAG TATAATGATG aTTTTCAGCT TGTACAAAGG	2580
30	AGaAAAAAAG AAGACAACCA AGCCCAATAA TGGACTGGCC GCCTAATAWA nGGGskCTAA	2640
	AAGTTgTATT TTTAAAAATA GTTCCTTTAA ATTATATACC ChCCACATTT GGGGGAGGAC	2700
	CTAAAAAAGC	2710

35 (2) INFORMATION FOR SEQ ID NO: 422:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1027 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 422:

45	CATTTTAATT GtTaaaATTC CAAAAAAtGT ArGTGGATTw AAAGrAAACC CtGTGTTTTT	60
	AAAAGGtACC ATTaAAATag TTCCGATTGT ACCATCCCAC GTGAAAtTTT TAgTAtTGCC	120
	GGTGaGAGAA AATGCCAATG CAATCATCGC AGTTCATAAT CATCCATCCG GTGATGTAAC	180
50	GGCCTCACAA GAAGATATCA TAACAACAAT GAGGTGAAG GAGTGTGGTT TGATTTTAGG	240

55

	GGGTTACTTT GATGAAAATG ATTGAAGTTG ATAGAATTAA TGACGTATCT TGTGTATAAT	360
	ACCTACGAAG TACTTTTCATT GGAGGAAAAA TAGTGACTCT ATTTATTATT ATCGGGGTTC	420
5	TCGTGCCAAT GGTTTATACC ATGCAGTTAA ATATTAAAAA TGAACCTGTa ACAAAGCGCA	480
	ATCTTTTAAT AACATTAGCT TTATCTACGT TAGGTATTTT AGTAACCGCG TTAGCAGGTG	540
	TAATCGTTAC GAAACAAGCT TTTCCTTTAT TAAGTGTAGC AATTGGCTCA ATTTTTACTG	600
10	GAATCGTTTG GGGCCTTTTA CTAAGTGGTA GcTACCGCT GATACGATTT TTATCTAACG	660
	CATTTGGGCG TAAGTAATAG TAATCTGTTC ATCAAGTAGT ATCCGTGCTT GAAAACAAAC	720
	TAAACTCCT AATGTGGAAC TAGTTAAAAA ATTCTAAACT AGTACATTAG GAGTTTTGTT	780
15	ATGCAGAATA AATTTAATTG TTAAATTGAA AGTGCGGTAA AAATCCACTA TTTATTTGAA	840
	AAAAATCGAG CAACCAAATT AAACCTTGAT ATCCTAAGTA AATACATAAT AAAACAAGTC	900
	CGACATAAAT TAAAAATCGC AAAATAGAAA GTCCAACCTCT AAAAAGGAGG ATGACTAATA	960
20	GTGCTATTAA AATAATTGt AGTATACTCA ATGCTACAAA CCTCCTArTA CGCTTTTAAA	1020
	TCCATAA	1027

(2) INFORMATION FOR SEQ ID NO: 423:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 569 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 423:

35	CGGCACATTT TTAATTTATA CATATTTTAA AACTAAGTAA CAGTTTGAAG AAATCGTAGT	60
	TCAATAATGT TAATTGTGAA AATGTATATA AACATAAAAA AATCATGTAT AATATATGTT	120
	GTTAATTAAA CAGTTTCGAAA GCGAGATGAC ATTATGGGAC GTAAATGGAA TAACATTAAA	180
40	GAAAAAAGG CCCAAAAGA TAAAAACACA AGTAGAATAT ATGCGAAATT TGGTAAGGAG	240
	ATTTATGTTG CAGCAAATC TGGTGAACCC AATCCAGAAT CTAACCAAGC TTTAAGGTTG	300
	GTGCTTGAAC GCGCTAAGAC ATATTCAGTG CCGAATCATA TTATTGAAAA AGCAATAGAT	360
45	AAAGCTAAGG GTGCTGGAGA CGAAAACTTT GATCACCTAA GATATGAAGG ATTTGGCCCA	420
	AGCGGATCAA TGCTAATTGT TGATGCGTTA ACAAATAATG TAAATCGTAC TGCCTCTGAT	480
	GTGCGAGCTG CTTTTGGTAA AAACGGCGGT AATATGGGTG TATCTGGATC AGTTGCTTAT	540
50	ATGTTGTCAT GTGGCCACAT TTGGTATTG	569

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3264 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 424:

10 TGGACCTATA AAATGGATGC ATCCAAATAT GGATGnGGGG TanGGCGGGa AAATAGGaGG 60
 TTATATAGTA AGTCATTTAA GTCGAATAGC CATCCTTTTT AAAATGAAAA AAATAGAAAG 120
 CTCAATAGTT TGTTAAAGCC TTAAATAGC GTCGTTACGC AATTTTAGAA TGCTAAAAAT 180
 15 TGTCACAAAT TTCAAATTAC GTGCTAAAGC TTGAGATATC AATATTTATT GCGGATAGAG 240
 TGTAATTTGA CTCGCTTAC ATCTCAGTTT TATGTTTGTA ATTTGGTAGC ATAATATTAT 300
 AATAAAATAA AATTGTTAAT CTTTAATTTT AGTATAGATA TTTTACGTG TAGTCACGTG 360
 20 TAAATAAAT TCAATTAGGT TAGGAGACAT AATTATGAAA ACATTTGGTA AAAAGGTTGT 420
 ATTAATCGGA GATGGATCTG TAGGATCAAG CTATGCCTTT GCAATGGTTA CGCAAGGTGT 480
 TGCTGATGAA TTTGTAATTA TTGACATTGC AAAAGACAAA GTAAAAGCAG ATGTTCAAGA 540
 25 TTTAAACCAT GGTACAGTCC ACAGTCCTTC ACCAGTTGAT GTGAAAGCAG GTGAATACGA 600
 aGaCTGTAAA GATGCAGATT TAGTTGTTAT TACAGCTGGT GCACCTCAA AGCCAGGTGA 660
 AACACGTTTA CAATTAGTTG AAAAAATAC TAAGATTATG AAGAGCATCG TTAAGAGTGT 720
 30 TATGGATAGT GGCTTTGATG GATATTTCTT AATCGCGGCA AACCTGTAG ACATTTTAAC 780
 AAGATTTGTA AAAGAATATA CTGGATTACC AGCAGAGCGT GTTATCGGTT CAGGTACTGT 840
 ATTGGACAGT GCACGTTTAC AATATTTAAT TAGCCAAGAA CTTGGTGTTG CACCTTCAAG 900
 35 TGTGACGCT AGTATTATTG GCGAGCATGG TGATACTGAA CTGTCAGTTT GGTCAACAAGC 960
 AAATGTAGCA GGTATTTTAC TATATGACAC ATTAAAAGAA CAACTGGTA GCGAAGcTAA 1020
 40 AGCGGAAGAA ATTTATGTGA ATACACGTGA CGCTGCTTAT GAAATTATCC AAGCTAAAGG 1080
 GTCAACATAC TATGGTATTG CATTAGCATT GATGCGCATT TCAAAAGCCA TTTTAAATAA 1140
 TGAAAATAAT GTCTTAAATG TTTCTATACA ATTAGATGGT CAATATGGTG GTCACAAAGG 1200
 45 CGTTTACCTA GGTGTACCAA CATTAGTTAA CCAACATGGC GCAGTTAAAA TTTATGAAAT 1260
 GCCATTAAGT GCCGAAGAAC AAGCGTTGTT CGATAAATCT GTTAAAACAT TAGAAGATAC 1320
 ATTTGATTCA ATTAAATATT TATTAGAAGA CTAAGCCTAT TTTAAGTATT AATTAGAAAT 1380
 50 ATATTAATGG TAAGAGGATC TATTAGTGTT GCAGTAACAC GTGGCACTGA TAGGTCCATT 1440

55

	TATGACTCGT GAAAATTTAG TGACATGAAT AAAAATGTTG AGGCGTCATT GAGTATAAAG	1560
	GAAAGTAGTT CTGCATTAAT CACGAaGTAG AGCATGACma CGAGGAATAA CTATAGGGaG	1620
5	ATGGTTTTGG AATGACGAtG TCTTGATCA ACATGGTACA TTAGGTACGT TAATGGCTGG	1680
	CTTACTAGAA GGCACAGCTA CAATTAATGA ATTGTTAGAA CATGGGAATT TAGGGATTGC	1740
	AACGTTAACA GGGTCTGATG GCGAAaAATA TTTTtagACG GAAAGGCATA TCATGCTAAC	1800
10	GAGCATAAAG AATTTATAGA ATTAAAGGC GATGAGAAAG TACCGTATGC ATCGATTACT	1860
	AATTTTAAAG CGAGTAAGAC ATTTCCATTG CAACAATTAT CACAAGATGA TGTATTGCA	1920
	CAAATTAAAA ATGAAATGTT AAGTGAGAAT TTATTTTCGG CTGTTAAaAT TTATGGCACA	1980
15	TTTAAACATA TGCATGTACG AATGATGCCT GCTCAGCAAC CGCCATATAC ACGTTTGATT	2040
	GATTCAGCAC GCAGACAACC TGAGGAAAAA AGACAAGATA TTCGTGGTGC CATTGTTGGA	2100
20	TTTTTTACAC CAGAATTATT TCATGGCGTA GGGTCTGCTG GTTTTCATAT ACATTTGCG	2160
	GATGATGAAA GAGCTTATGG TGGACATGTT CTGACTTTG AAGTGGATGA CGTTGTCGTT	2220
	GAGATACAAA ACTTTGAAAC ATTCCAACAA CATTTCCCGG TAAATAACGA GACGTTTGTT	2280
25	AAAGCGAAAA TAGACTATAA AGATGTGGCA GAAGAAATTA GAGAAGCTGA ATAGTCTAAT	2340
	ATAATTAAAA GACCTTAGCG ATATTAGGAA CAGGTGGTTC TAAATGCATC GTTAAGGTCT	2400
	TTTTATATTA TATATGTGCT TACATATTTT GTTGATACGC GCGTAAAAAG CTATAAATAG	2460
30	TATCAAGATA TGGGGTGTCa AGAATGTGTT CACGTGCACG TCGATAAATA AAGCCTTGAA	2520
	TCGCTTCGAC TTCAAGAGGT TGCTGATGCG CAATGTCGTA ATACATGCTC GTTCCCATT	2580
	CGTCGGGATA ACCTTGATAT ATAGTCATAA TAGTATCGAC AGTTTGTCTT GAAAAGTTTA	2640
35	ACCCCTCTGC CTGTGCAACA CGACAACCAT CTAGTAATAG CTGTCTACAC AGGTGACGTA	2700
	TTTCAGGATT ATGCATGATT GCAACGGTCT GTCTACCGAG TGCTGTGATA GAATTAATGC	2760
	CCAAGTTAAC TAACAATTTA TACCAATAG CTTGTTGAAT ATTTGCTTCT AATACAATGT	2820
40	CGATTTGACT GTCTTGACC AAATCTCTAA ATTGTCGAGT TAATGCATTA TCTTGATAC	2880
	GTAATTGATA ATCTCTGAAG TCGGTAACAA CATCGCCTTT CTTTgACCA CTTATATAGA	2940
	CAACTGCTTG GCATACGTTT TTAAATGAAA TATGTTGAGT TTGACCATAA CCATTTTGGG	3000
45	CTAAATGAT GAGCGTGTC TCATGTGCAA GATGAGTTAA ATGAGGAATA ACTGCATCAA	3060
	GTTGATGTGT TTTGACTGCA ATAATAATAA CATCAAATGT GTTTGTGACA TCTTCATAAC	3120
50	CTTTCACAAC AATATCTTGT GCAGGTGCAT GTGGTACAGT ATAATATGTT ATTGTTTTGG	3180
	CGTGTCTCCC GATAAGCGTT GTATGAGGCA ATGATTGTTG TAATTCATAG GCAATAGTTG	3240

(2) INFORMATION FOR SEQ ID NO: 425:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 796 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 425:

AGAACGAACC TTAAATACGT ACGTGTGCGAC CACGAACTGT CATGACAAAT TGATAAAACG 60
 AATAGTTTTT CATTAGTTCA TTGTCACATC AATCACTTTT GTkTCACCTT TAATCACAGC 120
 ATTTTCATCA TAAATATTAA TTGAAGCTGC TTGATCAGTG TTAGTAATTA TAATTGGTGA 180
 AATTACAGAT TTAGCGTTAT TATTAATATA TTCAAGGTTG AATCTTACTA ATGGATCTCC 240
 GACGTTAACT TCGTCACCAC TAGACACTAA CACTTCAAAT CCTTCACCGT CTAATTGAAC 300
 TGTGTCTAAA CCGATATGAA CTAATAATTC TAATCCGTTA TCTGCTTTTA ACCCAATCGC 360
 ATGCTTAGTT GGAAAGACAT TGTCAACACG TCCTGCAATT GGAGACACAA CTTCTCCTTC 420
 AGTTGGATTA ATACCAAAC CTTGCGCCAT CATTTTTTGT GCGAATACAG GATCTGGAAT 480
 ATCTTCAATT TTCACGAATT CTCCAGTTAA TGGTGCATAA ATTGCGATAT CTTTCTGAAC 540
 TTCITTGCCT TTTCCGAATA ATTTTTTAAA CATACTTTCC ACTCCTACTT ATCAAAATGT 600
 GATATTAAAT CGCCATAACC CAATTCTTCT AACTTTTCAT ATGGAATAAA TTGAATTGCA 660
 GCGGAATTGA TACAGTATCT TAAGCCGCCA CTTTCTTTAG GTCCATCAIT AAAGACATGT 720
 CCTAAATGAC TATTTGATTG TTCTGAACGC ACTTCAGTTC TCAACATACC AAATGATTG 780
 TCGACTAATT CTATAA 796

(2) INFORMATION FOR SEQ ID NO: 426:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3393 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 426:

ATCCAGCAAA TCTTTCTTAT CACGTTCTGT AATTTTCTTA GTATCCATCT TGATTAGCTT 60
 TGATAACTTT TCAGCCGTAT CCAACATTTT CGATTGTGTT GTTTTTCGAC CCCTAGTATA 120
 TGTAAATAGCC ATTTTAGAAG CATTATCAAC TAAAACTTTC CCATTTCTGT CTAAAATACG 180

	ATGTGAGCCT TGTGCGATTT GTAAATAACC TAAACGTAGT ACTAGTACTG CAAAAATAAA	300
	TACAATCACA CCAAAATATAA AGTTAATTCT CTTGTTAATT GTATTTTGAA CGATTTTCATC	360
5	ATTTGATTTT TCTTTTAGTC TTTTAAACAA AACTACCTAC CTCTATTCAA AGTCTTTCAC	420
	TTTAAATCAT ATATGAATTT AGAAATTATT TCTATCTTTT TGACAAAAAA ATAACGGTCT	480
	CATTGAAGAG ACCGAACAAG TAATCATACT TTATTTTGTT GCATTATATA ATTTCGTCAAC	540
10	TTTTTCCCAG TTAACATACAT TCCAAAATGC GCCAATGTAG TCAGGGCGTT TGTTTTGATA	600
	TTTTAGGTAA TAAGCGTGTT CCCATACGTC TAAACCTAAA ATAGGTGTTT TACCCTCAGT	660
	TAATGGATTA TCTTGGTTTG GTGTAGTCAC AATTTCTAAC TGGCCATTGT TTACGACTAA	720
15	CCAAGCCCAA CCTGAACCAA AGCGTGCAGC TGCTTTGTCA GCAAATTCTT TTTTAAATTC	780
	TTCTAAAGAA CCCCATTTGTT CTTTAATTTT TTCTACTACA GTACCTTTTT CTTCTGAGTT	840
	TGGTGAAAGT AACTCCCAGA ATAATGAATG GTTTAAATGT CCACCGCCAT TATTACGTAC	900
20	AGCAGTTTGG ATGTTAGCTG GTACACTGTC TAAATTAGCA ACAATTTCTT CAATAGATTT	960
	AGATTCTAAA TCTGTACCTT CTAAGTCAGC ATTTAATTTT GTAACATACG TGTATGATG	1020
	TCTGTATGG TGAATTTCCA TAGTTTCTTT GTCAAATGT GGTCTAATG CATCAAATGC	1080
25	GSTATGTAAT TTTGGTAATT CAAAAGCCAT AAATAATCAT CCTCCTAAAA TGTCTGTAAG	1140
	TAAATAATAA CAAGCAGTGA CTGGTTCAAC AAAGAATTG CTTAAATTCT ACTACTTATT	1200
30	ATTTTCTCTA CTCATTTAAT ATAACCTCAA TCAAAAATAA TTAAACATTT TGTATATAAA	1260
	AAGTTAACAG ATTTGCCATA AAATCATACG AACGGAGTAT GAAATGAACC TTTATCTTCT	1320
	ATAATTTAAA AAATGAGATT TATGCATACA TCGGACCAAA TGTGCATAAA TCTCATTTCT	1380
35	TTATATTAAT CTTGGCAAGA CTCACATACA CCATAAACTT CAAGTTTGTG TTTGTGAATA	1440
	TTAACACCAG GTAGTGATAA TTTTATCTGA TCTATTGGAC AATAATCTAT TACCTTTGTA	1500
	TCTCCACACT TTTCACAGAT AAAATGATGA TGATGATGGT TTGTACAAGC GATTCTAAAC	1560
40	TTCATTTTAC CATCAAGTTC TGTATTTTCA ATAATTCCTA AATCTTTAAA TAAGTGCAGG	1620
	TTTCTATATA TTGTGTCGAA TGAAATTCCA GGATAATTTT CATCCATAAC TTGTTGTATA	1680
	TACTTTGCGT TTATATACTT ATCTTCTTCG ACAAAAATAT CTAACATATC TTTACGTTTA	1740
45	TCTGTATATT TTAAACCGTT CTCTTTTAAA ATTTTAATAG CATCATTTGT ATTCATTGAT	1800
	ATTAGCTCCC TTTTAAACT TCATTCGCAT TTTCTGATAA GCCATTGTAA TCATAAGTAA	1860
	AATAACAAGT AGAACTACAA TTACACCACC CGGAGAAATG TCCATATAGA AAGCTAGGAC	1920
50	TAAGCCTAAT ATTACTGATA ATTCACCTAA AAATACACTT AGTAATATCA ATTGCTTAAA	1980

55

TATCCCTACA ACACGCATTG AGGCAGAAAT AACCATCGCT ACAATAACAA TAAATAAAAA 2100
 TTGAATCCAT TTAGGAATGC CAATGACTTT ACTATATTCC TCATCAAATG ACAATATAAA 2160
 5 TAATTCTTTA TAAAACAATG TAATAAACAG AACAACTATG ATGGCAATGA CAATAATCGT 2220
 TGTAAATCA CTTATATTCA CTGCGCTTAT TGAGCCAAAT AGCAATCCAA CAATTTCTTG 2280
 ATTGAACCCA TCAGCTAATG AAATGAAGAT TGCACCTCAAG GCGATACCAG CACTCATTAT 2340
 10 AATTGGAATA GCAATTTCTT GGTAAGCAGT GTATGACGTT CTTAATTTTT CAATTAGAAG 2400
 CGCACCTACT ATTGCGAATA AGATTCCAAA CCACATTGGA TTAATAAATA CTAGTGTGG 2460
 CATAATAGTA AGTAAAAACA TACCGAAAGA TATACCACCT AAAGTTACAT GACTTAGAGC 2520
 15 ATCAGCTWTA AGTGATAGTC GTCTAACAAc GATAAAAGCA CCGATTAGAG GCGCAATAAA 2580
 ACCTAtCAAG ATAcCACTAA TTAAAGAGTA CCTCATAAAA TCAAATTCA ATAATGCATC 2640
 TATCAATTGT GACACGCCTT TCCATTTTAA ATAACTCAA ATCTTTATTA ATTACAACAT 2700
 20 TCTCGATTAT GCTGATGATC GACAAAACGT ACAGGATGTC CATAAATTTT TGAAATTTCA 2760
 ACTTCATCAA GTGATTTAAA CTCATCAGTT GTACCATGGA AATGCAAATG CTTATTTAAA 2820
 CATGCTACTT CAGTAGCAGT ATCTGCTACA ACACCGATAT CATGAGTAAC TAAGATAATG 2880
 25 GTGATACCTT CTTGTTTAA TTGATCTAAA GTATTATAAA ATCACTTAC ATGTTTTGCA 2940
 TCAATACCAT TCGTTGGTTC ATCAAGTACT AATACTGcAG GTTCTGAAAT CAATGCTCGA 3000
 30 GCAATCATTa CACGTGTGTG TTGACCACCT GATAATTCTG CTATATTTTT ATGAATTAAA 3060
 TCACTTATAT TCAGTCTTTC TAGTACTTTA ATCATTTTTT CATTATCTTT GCTATTAAAT 3120
 GTTTGGAAAA GACGTTTTGT CTTTGTTAAT CCGCTTAAAA CAACTTCTTT AACACTTGCT 3180
 35 GGGAAACCTG AATTAAAGGC ATTTGCTTTT TGTGATACAT AGCTTAATTT AATTGATGTT 3240
 TcTTATTTTT AAAATCAATA CCTCAACAA AAATCTCACC ACTTTGTAAA GGTAATAACC 3300
 CTAGAATCAA CTTCAATAAT GkTGATTAC CAGCACCATT TGGkCCAACA AwTGctAAAA 3360
 40 ATTCACCTTT ATTTATTTTG AATGnnnTAT ATT 3393

(2) INFORMATION FOR SEQ ID NO: 427:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1123 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 427:

CCCTGTTTCT TAGCCTTTTC GTATATGCGT TCATAAGAAG AATTCTTTTT GATTTCAGAT 120
 TGTAAGCAC TGTTCCTACT TGATTGTTA TCTATTTTAT AATCTAAATC TGCAATCTTT 180
 5 CCTGCGTAT CATACGCATC CATTTTAA AATAGCATAT AAATACTTAA CATAGCAATT 240
 ACAGTAATCA AAGTTATGTA TAAACTTTTT TCAAATTTAG TTAATTGTAC AACCACCTTTT 300
 CTCGAAACAG TCTTCTTTTC GGGCTTAGTT TGTGGTTGTT GCTTCGGTAT ACTATTATAA 360
 10 ACTTGTTTCT CATATGGTTG GTACACTTTT TCTACAGCCA TTATAAATTG CTCCTTATTT 420
 AAGTATTTCA GCTACACGTA ATTCGCGCT TCGTGCTCTG TTATTGTCAT CTAAATCTTC 480
 TTCTGTAGCG GTAATCGGTT TACGATTAAC ACGCTTTAAC TTAGGTGTAT ATGCTTCTGG 540
 15 TATAACTGGT AATCCTCTTG GTACCTCTGG ACCTTTTTCA TATTCTTGGA ACACCTGTTT 600
 ACATAAACGA TCTTCTAAAG AATGGAAAGT GATTACCGAA ATCCTGCCAT CTACTTTCAC 660
 TAATTCAATC GCTTGTTCTA TTGAATCTTC AAAAGCTGAC AATTCATCGT TTAGTGCAAT 720
 20 TCGTAGTGCT TGAAATACTC GTTTTGACAGG ATGTCCGCCT TTTCTTCTTG CTTTTCAGG 780
 AATACCTTCT TTTATAATGT CAACTAATTC TAATGTTGTT GTTATTGGTT GTTGTTCGCG 840
 ATGTGCTTCG ATTCTTCGAG CTATCTGTTT TGAAAATTTT TCCTCGCCAT AGCGATAAAA 900
 AATCTTCACT AACGCTTCAT ATGACCAATT GTTAACAATT TCATATGCTG TTAGTTCCTG 960
 TGTTTGGTCC ATACGCATGT CTAATGTTGC GTCATGGTGA TAACTGAATC CTCGTTCTGG 1020
 30 AATGTCGAGT TGTGGGCTTG AAACACCCAA GTCGTAATAA ATTCCATCTA CTTTTTCAAT 1080
 GTTTAAGTCT TTTAATATTT GAGTTAATTC ACGGAAGTTG CTA 1123

(2) INFORMATION FOR SEQ ID NO: 428:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 838 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 428:

45 TATATCCTCC AaACCATAAA AATGGAATTA TAGTGAAGAC CGCTATGATC CAGGATGACC 60
 ACGTATTATA GAATCTTTCg ATAGAGAAAC GAATCaACGT ATCCgACATC aCGAAATGAA 120
 ATTAGAAGAT TATAAAGATG AGTTAAGAAG AGAATATCTA AAACAATCTG ACAGAATTGA 180
 AGGAGATGAA TAAGCGTGGG ACTTGATTTT AGTGGTTTAC CAGATTAGC AGTATTGGAA 240
 50 CAAATGAAGG AAAAAGAACA GATTAGTGAG GTTATTGCGC CTGAACATGT TCGTATGCAT 300

CATTTCAAAA AATTTGAAGA TGATTTTAAA AATGCGGCAC AAGGGGCTTG GGTGAAAAAT 420
 GCCACAGACG AATTAAAAGA TATTAGTAAT GATTTAGAAA AAATTCAAGA TATTAAAGTA 480
 5 TAAAAAGGTA TTAAGAAACA CAATAAGTAT ATAATCCATT TAATAATAAA TGAATTATAT 540
 AGTTCATAAT TTCGACTATA AGTGGCTATT AGCCATTATT TTTCGGGATC TATGTCAAAT 600
 CGGACTAATG AATTCAATAA TGGAAGTTAA GCAACCAATC TTTGTTTAAAC TTCTTTTTTA 660
 10 TTTTTTGGAA AATAAAGTTT TGAACATAAT AAAATTTGAT TATGTTTAA CGAATTTTGA 720
 TGTTCCTTGA ACTATATATC ATCTAGTCGT CATTTACAAT TGGTAAATAT GACTTCAAAC 780
 15 TGkATGAAGG TGATGGCGAT TAAAAGGCTC ATCCGTAGGT TCTAAAGAAC TAGAnTTT 838

(2) INFORMATION FOR SEQ ID NO: 429:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1150 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 429:

25 nGTGTTCAAA TAATAGAATG GCTAATACCC CAACAATAAA TAGGAATAAA TATGTTGATG 60
 30 TCGTAATTTT AGTAAATAAT GAAATCTAA AATTAGTAAC CCTATTTTGA ATATAAGCTT 120
 TAACTTCTAA TAATACTGGG AAGCCGATTG ATCCTAAAAT TATTAAAAAT ATAACTATCG 180
 CTTGTACAAA ATAATCATGT GCATAAGGAA TTAATGACTT ACCTGTAATA TCTAATCCAC 240
 35 CATTGGTAGT GGCAGAAATA GAAACAAAAA CACCTTGCAT GATTGCATAT TTAAATCTG 300
 GATTATCTCG ATAAAAGTAA AATGCTAATA ACATGGCTCC TACGAGTTCG ATTACAAATA 360
 TTGATTTTAC AATATCAATA ATCAATTTGA CGGTACCACT CATTGTGTTT TTGTTATTAT 420
 40 CTAACATAAT TAACTGACGT TCTCTAATTC CAATATGTTT ACCTAGTACC ACCCATAACA 480
 TCGTACCAAT TGCCATGACG CCAATTCCAC CAATATTTAA TATCAGGAGG ATAATTAATT 540
 GTCCAAATGT AGAATAGGTA TCGACAATAC TTATCGGAGA CAATCCAGTA AACTAATTC 600
 45 CGGATACGGC AACAAATAAT GTGTCAATTG GATTTACTTC TACACCTGGT TTATGaACAT 660
 ACGGTAAGTT TAATAATAAA AATGCAATGA CAATTGCGAC GATATAGTAC AATACAATAC 720
 50 CTTGTTGAGG GCTTGATCTT TTTAAAACT GGCTAAAAAT TGACACGTTG TTCACCTCAA 780
 CATTATTTTA GTTTAATATT AATTTCTTTC TCTTTACCAT CTCTATAAAT CTTGCTGTGA 840
 ATTGATTTCA AGTCATCTTT ATGACTAAAT ATAATCTGCC TAAAGCGTAA ATCATCTTCT 900

55

AAACCAATTGT TGTCAACTTG ATCTACAACA ACACCGTTCT TAACTTTTCC TGGCAATTTA 1020
 ACTGCTTGTC TTTCAAAACT ATTTAGACTG ACAATATTCT TCATTTTAAAC ACCTACATCG 1080
 5 GGATAGTCAA TTTTACCTTT tGTTTCTAAA TCTTTTACAA TCTTTtGTAC TTCATTAAaCA 1140
 GGtATTGCAA 1150

(2) INFORMATION FOR SEQ ID NO: 430:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 797 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 430:

20 AAGAATATTT AGGATTTTAT GTTTCGCAAC ACCCAGTAGA TAAAAAGTTT GTTGCCAAAC 60
 AATATTTAAC GATATTTAAA TTGAGTAACG CGCAGAATTA TAAACCTATA TTAGTACAGT 120
 TTGATAAGT TAAACAAATT CGAACTAAAA ATGGTCAAAA TATGGCATTG GTCACATTAA 180
 25 ATGATGGCAT TGAACTTTA GATGGTGTGA TTTTCCCTAA TCAGTTTAAA AAGTACGAAG 240
 AGTTGTTATC ACATAATGAC TTGTTTATAG TTAGCGGGAA ATTTGACCAT AGAAAGCAAC 300
 AACGTCAACT AATTATAAAT GAGATTCAGA CATTAGCCAC TTTTGAAGAA CAAAAATTAG 360
 30 CATTTGCCAA ACAAATTATA ATTAGAAATA AATCACAAAT AGATATGTTT GAAGAGATGA 420
 TTAAAGCTAC GAAAGAGAAT GCTAATGATG TTGTGTTATC CTTTTATGAT GAAACGATTA 480
 AACAAATGAC TACTTTAGGC TATATTAATC AAAAAGATAG TATGTTTAAT AATTTTATAC 540
 35 AATCCTTTAA CCCTAGTGAT ATTAGGCTTA TATACTTTy ATCAACTAAT AAATTATGAT 600
 ATAGTAACT GATGGTTAGA TATTTyTyAA CCATCAGTTT GCGTktATAT TAGTTTTTTA 660
 40 TGCTTATTAT TTTTATGAGT TTCACTTTAC ATTATTGATT AATCCATAAG AATAATTAGC 720
 ACAAAAAGCA GTATACATAA ATTGAGTAAA GAATTTTGTC GATATAACCG nGCGGAAAAA 780
 TAATAACnTT TTGGATn 797

45

(2) INFORMATION FOR SEQ ID NO: 431:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1466 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

	TTTCCAAGAG AAGTTGCGGA AGTAATTAAT AAAACGCATC ATAATAAATT GGTCATTTTCG	60
	ATGATTTTCGk CACAAATCGA TGCGGATAGA ATGGATTATT TACAACGTGA TCGTATTTTC	120
5	ACAGGTGTAT CATATGGTGC TTTTGATATG GAGCGTATTT TAAGATTAAT GCGACCTTCT	180
	AAAGATGAAG TACTAATCAA AGAAAGTGGT ATGCATGCAG TTGAAAACCT TATTATGAGT	240
	CGTTATCAAA TGTATTGGCA AATTTACTTC CACCCAGTTA GTCGTGGTGG AGAAGTGCTG	300
10	CTTAATAAATT GATTGAAACG CGCAAAACAG CTTTATAATG AAGGCTATGA ATTTAAGTTG	360
	CATCCACATG ATTTTATTCC ATTTTTTGAA GAGACAGTTA CGATTGAACA ATATGTTGAA	420
	CTCGATGAAG CGGTAGTTAC GTATTATTTG GAAAAATGGA CAAAAGAAGA TGATGCTATT	480
15	TTAAGTGATT TAGCAAGTCG ATTTATTAAT CGAGACTTAT TTAAATATAT TCCATTTGAT	540
	GGCTCAATTA TTACAATATC AGAACTGCAA GAACTGTTTG AAGCAGGTGG TATTAATCCA	600
20	GATTATTATT TTGTGAGTGA AGCATTCTTCT GATTGGCCAT ATGACTATGA TCGACCGGGG	660
	TCAAATCGCA AACCGATTCA TTTATTAAGA CAAGATGGTA CGATTAGAGA AATAAGCAAT	720
	CAATCATTAG TCATTCATAG TATTACAGGC ATTAATCGCC AAGACTATAA ATTATATTAT	780
25	CCTAGAGAAA TGTTGCAAA GATTAAAGAT AAGACAATTA GAGAAGCTAT TGAAAATTTG	840
	ATTAATGAGC TTAATTAAAC AGGGCTAAAA TTGTTATCGT TAAATATGGA GGTTATATCA	900
	TTGTCTGAGA AAAAAGGCTT TAATTTTAAT ATCATAAAAA ATGACCCTCT AGATGGTCAT	960
30	AAAGGTACAA ATATTGGTTC AATTAGCTTA GACAATATTG CACCAGTTTT TATCGATGTT	1020
	GCTAACAAAG AAGCATTTAT TGATATTGGA GGCATGCATG CTCGTGCCAA AGTTGAAAAA	1080
35	GGTGTGAAAT GGATTACTGA TAAAGCTGCT GTTGAAGGCG ATGAAGCTAA AGaATATTGG	1140
	TTGTGTTGGG TAACAACAGA ACGTAATGAA CAAGGACCAT ATTACGCTGG TTTAACAGCG	1200
	TGCTATTTAT TAGTGAATAA AGCAATTCGT CGTGGTTATA AAAGTATGCC TGAACATGTT	1260
40	AATATGATGG ATAAATCAAT GAAACATCAT ATTATCATAG ATCAAATTGG TGACGAGAAT	1320
	AAAGCTATTT TAAAAGACTT TTTAATGAAC CATGATGAAG GTATGTGGAA GCATTCTTCT	1380
	GATGCTTTAC ATCAAGCATT TAATTAAATA TTAGAACTA AAATTTCCCA ATTAATCTAT	1440
45	AAAGATATGA TTCATTTCTC AATGAC	1466

(2) INFORMATION FOR SEQ ID NO: 432:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2304 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 432:

	TTGTACGAAT ACAAAGTTGC ATTAATATAG AATCATTTTAT GCTTTTTTAAT ATCACAAAAC	60
5	AAATTGATTA AACTAGCTAA AATAGTCAAA ATTGGAATCA AGATTACATA AATATCATAA	120
	CCTCTAGATA TTAGTATACA TAATATAATT AGCGACCCTA TTATAATTGC TAAATTATAA	180
	GCCAAACTAA TTTTATATTT CATTTCAATT CCCACCTTAA TAGCATTTTAT TAAAATTCTA	240
10	TAATGTTTAC ATAGACTTCT TAACAACTT TAACCCAGCT AAGCTAAGTA TTGAAATAAT	300
	TACACCTCCT AATATAAATA TATTACTTAA GCTTAGTAAA GGTATAATTA TACTTATCAG	360
15	TCCTAAAGAC AATGTATCCG CTGcATAATT CGaTGtAGAT GAGATACTAA ATACTTTCCC	420
	CATCAAATGA TTTGGCGTTT TTATTTGAAT GGCAACTGAT CTTGTTAGTC CCTCTATAGA	480
	TTGTCCAAGT CCCAATAATG TTGCACCTAT ATATAATATC GCCACACTTG GAAACACATT	540
20	AATAATCGAT AAGCCAATTC CCCAACTAA AACACCAATA CTAAATTTAA AGATTAATCG	600
	CTTTTCTGAC AGCAAACCCA TAATCAATGA CATTAAATAA GATGCTATAC CTAAACATGA	660
	TGTAGCTAGT CCATATACAC CAACGCCCTC TTTTAATATA TTGGAAATAA ACAATGGTAA	720
25	TACAACACGC CAAAGGCCAG TATTAATCAA TATGCAGGCA AATTGGATGA TTATAATAAA	780
	TGGAATTTCT TTAGATTGTT TCAAGAATTC CCAAGTTTCA GAAAAATCTT CTTTGTAGTG	840
	TCTATCAATC ATGTTGTTAT TTGTATATTT TAAAAGTGCA TTAAAAATAA ATCCTAAAAA	900
30	TAGCAATATA CTACAAATAA AAAAGACGCC AACATTACCA ACTAGTATTA CAATGACACC	960
	AATTAAAGCA GGTAAAATAA TATTTGAGCC TCTTTGCAAA CTATCGATTA ACGCATTACT	1020
35	TGTTTGCTAAA TGCTCCTCAT CAATAATTC AGGAAGAATT GCCCTAAACG CAGGATCCGT	1080
	ATAGCAGTTA ATAATGGTGA TAGCTGTAGA TATGGTTAGA AGCGTCAGAT AATTTAAATT	1140
	TGATGTTATT GCAAGTAAAG GAATAATTAT TATAATCAAA CTTAGTATAA GATCAGATAG	1200
40	ATAAAGTATT TTCTTTTTAT TATGTTTATC AGAATATGCG CCACCGAAAA TACCAATAT	1260
	AATAAATGGA AGTGTGTTGAC TCATAACCAT CATTGATAAT TTTAAAGATG ATTGGTTTGT	1320
	CAATTCAACA GTAAACCAAA TTATTTGTAA CGAAAACAGC ACAAACAAC TCCGACGTAA	1380
45	GATATTACCA ATCAATAAAT ATGTAAAGTT TCTATTTTTC AAAACTTCTA AATACAACAT	1440
	ATTTATCACC TCTCATAAAA ATAATTGAAT GCATCCACCA GCTTTTTTAG ACCTTCTTCT	1500
50	AAACTCTCTT TATCCAAAGC GCAATTAATT CTAATATAAT TTAGTCAGTT AAATATCAAT	1560
	TATTTGAAA TATACATACT ACTTGAAACA CCATACATAA CCCCCAAAT GACTACTCAG	1620
	AGGTTATATT CTAATAATTA TGATTATATT AAATATGAAA ATATTATCAA AAAAATCAAA	1680

55

TCATTCTCAT CATATTCTAG GTTGTTTTAA ACAAACTAAA TATAGTGAAT GCAAATCAAC 1800
 TATTATTTAA ATTATGAATT ATTTTAATTC TTTCTTCTAC GAGCCAATAA CATTAAATCCA 1860
 5 GCAATTCCAA TTATACTACT AAAGATCAAA CCTTTTTCGCG TGCTTTCTAA ACCTGTTTTT 1920
 GGTAATTCTG CTCGTTTTTT CTCTTGATTA GCTACTGATT CTTTAGCAAT TTTAGATTTT 1980
 TTAACTTTAT CATTTTTATC CATTGAATGA ACTGGGCCAT TTGGTTTTGC TCTGTCTTTC 2040
 10 GATAATCCTG GATTGTTAGG ATTTACTGGG CCACTTGGAT GAGTTGGTCT GCTCGGCTTC 2100
 TCTGGGTTTT CAGGTCCTTT TGGATCTTTT GGTTTCTCTC CACCGAACTC TACAATCTTA 2160
 TCTACTGGTT GTTTtGTGAT CTCTTCTGTT GGTTGACCCT CGCCAACTTT TTCACCTGTT 2220
 15 AATGGGkTCA CTGTGAWTGG TGkTGTGAWT GTCyTACTTC CTGGkTGTCC TTCTTGTTTC 2280
 ACTCGCTCTT CACCAGGTG TAAT 2304

20 (2) INFORMATION FOR SEQ ID NO: 433:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4733 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 433:

30 GTCTTAATAT TTTACAAGAT ACAGCCTGAC AATATTTATA TTACAAACCA ATATGGTAAA 60
 GAGGCTGTGT CTCGATTAAA AAGAATTTTA GAGCAAAGTA TACCTAATAA AGAATATATT 120
 35 GATTTATCAG AGATATCATT AATTACATTT TTAAACCAG AATATAATAA ACAACATGTG 180
 AATGAGGAGT TTAATTCCTT ATCTTCTTCG AAAATAAAGA ATTTAGCCAA AGAAAATGAT 240
 GGTATTATTA TTTATTTAAA TTTAAGGAT GTTAATTATC AACTGaTAaG TGAAGGCAAT 300
 40 ACTTTTTTTTT CTGCAAAACC ATATTTGCAT TGTGTGTTCA ATAAAGATAT TTTAAATATG 360
 GATAAAGTTG ATATAGAAAA TTTTTTCAAA AGTATAGAAG TCAAATACAG TAAATAAGAT 420
 CAATTATTGA ATCCACCGAC CAGCAGATTT TTTTAAAAA TTAATACCCC GTTAATACCC 480
 45 TTTGCTTCAA TTTGATGAAA ATCAATGAAA TTCAAATTG AAGAAATCCT TTAATATCAA 540
 GGTGTACGAC AGTCTATATA ATCATGCGAA ATTCTAAAAT TTTCTGAcGT AAAAAACAA 600
 ATTCyTAAAG CAGCTCGTCG TTCACCTCAA TTCTCAAAAC GTTAATTGTC GGACGATATA 660
 50 TATACAAAAC ACCTCGATGT TATGTGAGG TGTTTTTTTG CGTTTGTGTG GGGAAATATGG 720
 AATAGAGTAT AAATGAATTA ATGTCTCAGG TATAGAATTA ATTCAACTAT GAATTATTGA 780

55

	TTAGCTTGCG	GAGACCGTAC	GATAGTGTTT	TAATAATAGA	TATTCAAGGG	AAAGAGCTAT	900
	CGGATGATAA	CCAGATGATA	ACTCAAGAAC	CATGTTTAA	GAGTGTATC	AAAATAAGGA	960
5	GTGACTTAAT	TGAAAGATT	TATATTTTGG	AAGAAAGCTT	TTATTyCTGT	TTATTTTATT	1020
	GTTGsgATGC	TGGTGTCTT	ACTTTTtagG	TTTTATATTA	AaACAGATAA	CTTTTCTATA	1080
	TATTTAATGA	GTATCTTCTT	AATTTGTTTA	GGAAGTCTT	CTATCATTTA	TAAGTATAAA	1140
10	ACCAATCGAT	AAATATAAGT	ATGAAATGTA	TAAGTCACAG	TTAGATCTAA	GTCTTGCTGT	1200
	GGCTTTTTTAT	ATCTCCATAA	AACGACAAAT	TCAAGCCCCG	CATAAAACAG	CATTTTTCAGC	1260
15	CCACCATAAA	ACGACAATTT	CAGCCCGCCA	TTGACTAAAC	ACCACATCCC	AAAAATATCG	1320
	TAACAATCCT	CTACATCAAT	CAATCCAACA	TCCCTCATAA	TCACAACGCA	CAAAATCTAT	1380
	TCATGCATTT	TTGGAATACT	TAGTATTACA	AATAACGATT	TTTATTTCATC	TAACAAAGGT	1440
20	TATATAATGT	ACTGAAGGCA	ATTTTATGT	ATTACAAATC	TAATCGTACA	TGTAAAATTT	1500
	TGATAAACAT	CATTAATTTT	GCGTAACTAT	CATTAGATTA	CAAATCACAA	AGTAATTACA	1560
	TGTAATACAC	ATCTATACAT	CACATTTGAA	GGGAAATGaa	TATAAATGAC	TGATAAAAAG	1620
25	TACACTGCAG	CCGATATGGT	TATTGATACT	TTGAAAAATA	ATGGGGTAGA	ATATGTTTTT	1680
	GGTATTCCGG	GTGCAAAGAT	TGACTATCTA	TTAATGCTT	TAATTGATGA	TGGTCCTGAA	1740
	CTTATTGTCA	CTCGTCATGA	ACAAAACGCC	GCAATGATGG	CGCAAGGTAT	TGGAAGATTA	1800
30	ACAGGCAAAC	CGGGTGTAGT	ACTTGTTACA	AGTGGTCTCG	GTGTAAGTAA	TTTAACTACT	1860
	GGATTATTAA	CCGCAACATC	TGAAGGGGAT	CCTGTATTAG	CGTTAGGTGG	CCAAGTGAAa	1920
35	CGnAnATGAT	TTATTACGAT	TAACGCATCA	AAGTATTGAT	AATGCTGCGC	TATTAAAATA	1980
	TTCATCAAAA	TATAGTGAAG	AAGTACAAGA	TCCTGAATCA	TTATCAGAAG	TTATGACAAA	2040
	TGCAATTCTGA	ATTGCTACTT	CAGGAAAAAA	TGGCGCAAGT	TTTATTAGTA	TTCCGCAAGA	2100
40	CGTTATTTCT	TCACCAGTTG	AATCTAAAGC	TATATCACTT	TGCCAAAAAA	CAAATTTAGG	2160
	AGTACCGAGT	GAACAAGATA	TTAATGATGT	CATTGAAGCG	ATTAAAAATG	CATCATTTCC	2220
	TGTTTTATTA	GCTGGTATGA	GAAGTTCAAG	TGCAGAAGAA	ACAAATGCCA	TTCCGCAAATT	2280
45	AGTTGAGCGC	ACGAATTTAC	CAGTTGTAGA	AACATTCCAA	GGTGCAGGTG	TAATTAGTCG	2340
	TGAATTAGAA	AATCATTTCT	TCGGTCGTGT	GGGCTTATTC	CGCAATCAAG	TTGGTGATGA	2400
50	ATTATTACGT	AAAAGTGATT	TAGTTGTTAC	AATCGGTTAT	GATCCAATTG	AATACGAAGC	2460
	TAGTAAGTGG	AATAAGAAT	TAGAAACACA	AATTATCAAT	ATTGACGAGT	TCAAGCTGAA	2520
	ATTACTAATT	ATATGCAACC	GAAAAAAGAG	TTGATTGGTA	ATATTGCTAA	AACGATTGAA	2580

55

	CAATTAAGAA CACATATTGA TGAAGAACT GGTATTAAAG CGACGCATGA AGAAGGAATT	2700
	CTACATCCAG TGGAAATTAT TGAATCTATG CAAAAGGTAT TAACTGATGA TACTACTGTA	2760
5	ACAGTTGATG TTGGAAGTCA CTATATTTGG ATGGCACGTA ATTTCAGAAG TTACAATCCA	2820
	AGACATTTAT TATTTAGCAA TGGTATGCAA ACGCTTGGTG TAGCATTACC GTGGGCAATT	2880
	TCAGCTGCAC TTGTGCGCCC TAATACGCAA GTTGTGTCCG TTGCTGGCGA TGGTGGCTTT	2940
10	TTATTTTCAT CACAAGATTT AGAAACGGCC GTACGTAAAA ATTTAAATAT CATCCAGCTT	3000
	ATTTGGAATG ATGGAAAATA TAACATGGTT GAATTCCAAG AAGAAATGAA ATATAAACGT	3060
15	TCGTCAGGTG TAGACTTCGG TCCTGTAGAT TTTGTAAAA ATGCAGAATC ATTTGGCGCG	3120
	AAAGGTTTAC GAGTTACTAA TCAAGAAGAA TTAGAAGCGG CAATTAAAGA GGGCTATGAA	3180
	ACAGATGGTC CAGTATTAAT TGATATACCT GTAAATTACA AAGATAATAT CAAACTTTCA	3240
20	ACAAATATGT TACCTGACGT ATTTAACTAA AATAAAGATA AATGTTAAAG AGGAGTGGGA	3300
	GATTTTATGA CTAATGTTTT ATACCAACAT GGTACATTAG GCACATTAAT GGCAGGATTA	3360
	TTAAAAGGAA CTGCATCAAT AAATGAATTA TTGCAACATG GTGACTTAGG TATCGCTACA	3420
25	CTAACAGGTT CAAACGGTGA GGTAATCTTT TTAGATGGAA AAGCTTACCA TGCAAATGAA	3480
	CATAAAGAAT TTGTAGAATT AAAAGGTGAC GAGTTAACAC CATATGCAAC TGTAATAAA	3540
30	TTTGTAGCAG ATACAAGCTA TGAACGAAA GATAAATCTT CAGAAGCAGT TTTTGCAGAA	3600
	ATTAAGGAAA AGATGTTGAG TGAAAATTTA TTTTCAGCAG TAAAAATTTT AGGCTTATTT	3660
	AAAAAATGC ATGTACGTAT GATGCCGGCT CAAGAACCAC CTTATACACG TTTAATCGAT	3720
35	TCAGCTAGAA GACAACCTGA ACAAAGTAA ACGTATGTCA AAGGTTTCAGT CGTTGGTTTC	3780
	TTTACACCAG AATTATTCCA TGGTATCGGA TCAGCAGGAT TTCATGTACA CTTTGCAGAT	3840
	GATGATCGTA ACTTTGGTGG ACATGTCTTA GATTTTGAAG TAGAAGATGT TAAAGTAGAA	3900
40	ATCCAAAATA TAGAAACATT TGAACAGCAT TTTCCAATTC AAGATAAAGA TTCACTAAA	3960
	GCAAATATTG ACTATAAAGA TATTGCAGAC GAAATTAGAG AAGCTGAATA ATGAATCCAG	4020
45	AAATATAATG ACGGTTTATG AAAATTGACT TCATAATGCG CGATTTAGAA ATGATAGTTT	4080
	GTAAATATGA TTAACCATGA CTACAATAGA ACAAATATAT TTATAATTAC GTCTAAGTAA	4140
	TAAAATAAAT CCCTTCACTA TTAGCAGTAG TGAGGGGATT TATTAGGTTT CAGATATTTG	4200
50	AGATTTGCTG TTATGTTTAG ATTATAAATT GTGGTACACA CTCATATAAA ATTTACTATT	4260
	GTATAGGCCA ATCTGTTACT ACGAGAAGCA AACAACAATA ATTTACAAGT TCAATAACTA	4320
55	AAAAGACAAA CGCCAATTTT TCAGCGCTTG CCCTATAAAA CTATTTTCAA ATTATTATTT	4380

TTTCTTTAGA TTCACTACTT TTTTATTACC ATCATTCAAA GTAAGCGTAT AAGTTGCTGT 4500
 TTGGGCATTA TTAATTTTTT CTGTTGTAAC ACCACGTTGA GAAGCTAATT CATTTTTTAC 4560
 5 TTTACTGTCA ATTTCTTGAT AAAGAACATT TTTATTTTCT GGGAAGATAA AGTAAGTTCTG 4620
 ATGTAATGCA GTAATACCAT CTAAGAAAT TGTGTAAGGA ACAGTGTGAT AACCATCCAC 4680
 10 AGTCATTTCT TTATAGCCGT TATTACTATC TGCAGATGCT TCGTGACTCG GTA 4733

(2) INFORMATION FOR SEQ ID NO: 434:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1284 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 434:

AATAAATGGT AATTTTATAT CATCAACGGC TACAAATGGG GCAGTTTCAT ATATAAATAA 60
 ACACATCTAC TCAACTGnTC TTAATTAATT AATAAATGAT TTACCAGATG ACATATAACA 120
 25 GCCATTATTA GntTAATTCT AATAGTTTAT TTAAATTTTC TTCGGTTGTC GCCCAACTGG 180
 TTGCGAATCT AACAACACGA TGTTGATCAT CGTATTTTTC CCAAACAGCA AATTTAACCT 240
 30 TTTGTTCTAA CTCTGCTATT TTCTCGTTAC TTAAAATAAA AAATTGTTGA TTGGTTGGAG 300
 AATCAAAGTA AAGACGATAG CCTTTATTTT TAAACCCGTC TTTCATCTTA TTTGCCATTT 360
 CGATAGCATG TCTGCTTATA TTAAAATATA AATTGTCCGT AAATAATTCT AAAAATTGTA 420
 35 TGCCTGTTAA CCGTCCTTTT GCTAAAAGGG CACCGTGGAT GCTTGATTCTG AGTGGTAAAT 480
 TGTTCGGTT CATTATTTT CGTAAAAACA ATGGCTTCCC CGCATAATGC ACCTATCTkC 540
 GTACCACCTA TATAAAATAC ATCACAATAT TTAGCGrTGt CTTTAATAGT CATATCTGAT 600
 40 TGGTCACTCA TCAATCCATA CCCTAATCGT GCACCATCCA TAAATAATGG AAGCTGATAT 660
 TGCTTACATA CTTTGATAA CTCTTCCAAT TCTGaTTTAG AGTATAATGT GCCATATTCT 720
 GTAGGATGAG AAATATATAC CATTCTGGG AATACCATAT GGTCTTTTTT AAAATCACTT 780
 45 TTAAATGTCT CCATGTAAGT TTCAACATCT GAAGCACTAA CTTTTCTTTC CTTAGAGGGT 840
 ATAGTAATTA CTTTATGTCC ACTATATTCA ATTGCACCGC CCTCATGCAC AGCAACATGA 900
 50 CCAGTGTCTG CTGAAATGAC CCCTTCGTAA CTTTCTAACA TTGAATTAAT AACAACCTGA 960
 TTGGTTTGTG TTCCACCTaC TAAAAACGA ATTGTAGCAT TTGGaCAGTC AATTGTATCT 1020
 TTAATCTTTT CAATTGCCTG AGCTGTGAAT TGaTCAAAGC CATATCCCGA AGCTTGTACA 1080

TCGAATGAAA TCACTACATT TCCCCCTAAA ACTAATATCA ACATTTTAAAT AAGATAAACC 1200
 AATTTCAAAA CTAGTTCGAT ATTTAAAATG TATTATGGAT GGnTAAAGTT TGTATCGCAT 1260
 5 TATCGCGAAG TTGnATAAAT ATAT 1284

(2) INFORMATION FOR SEQ ID NO: 435:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1072 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 435:

ACAGCTTTTG GTAATGGAGA AAWTCAaTAt GAAACAGTAT AArGCGTATT TaATCGATTT 60
 20 AGATGGCACA ATGTATATGG GAACAGATGA GATTGATGGA GCAAnACAAT TCATCGATTa
 TTtAAATGTA AAAGGCATTC CTCATTTATA CGTAACTAAT AATTCAACAA AAACACCTGA 180
 GCAAGTAACT GAAAAATTAC GTGAAATGCA CATTGATGCT AAACCAGAAG AGGTTGTAAC 240
 25 GTCAGCGTTA GCCACTGCTG ATTATATTTc AGAACAATCA CCAGGAGCAT CAGTATATAT 300
 GTTAGGTGGG AGTGGTTTAA ATACTGCGTT AACCGAAGCG GGACTTGTCA TTAAAAATGA 360
 CGAGCATGTT GATTATGTAG TTATTGGACT TGACGAACAA GTTACATATG AAAAGCTTGC 420
 30 GATTGCAACG TTAGGTGTAA GAAATGGtGC AACATTTATT TCTACAAATC CTGATGTATC 480
 AATTCCTAAA GAGCGTGGTT TATTACCTGG TAATGGTGCT ATTACAAGTG TTGTAAGTGT 540
 35 ATCGACAGGT GTATCGCCAC AATTTATTGG TAAACCAGAA CCGATTATTA TGGTTAAAGC 600
 ATTAGAAATT TTAGGATTAG ATAAATCCGA AGTTGCTATG GTAGGCGATT TGTACGATAC 660
 CGATATTATG TCTGGTATTA ACGTAGGTAT GGATACGATT CATGTACAAA CAGGTGTATC 720
 40 TACGTTAGAA GATGTGCAAA ATAAAAATGT GCCACCAACG TATTCTTTTA AAGATTTAAA 780
 TGAAGCAATA GCTGAATTAG AAAAATAGAT ATAGTCATTT TATAAAGTAG GTGAATTGAT 840
 TTGGTAAAAA TAGTTGTTTC GAGGAAAATT CCAGATAAAT TTTATCAACA ATTAAGTAAA 900
 45 CTTGGTGACG TTGTTATGTG GCAAAAATCA TTAGTGCCTA TGCCTAAAGA TCAATTTGTG 960
 ACaGcCTCG TGACGCAGAT GCTTGTTTTA TTACATTAAG TGAACAGATC GATGCAGAAA 1020
 50 TTTTAGCGCA ATCACCAAAT TTAAGgTAA TTGCGAATAT GGCTGTAGGA TA 1072

(2) INFORMATION FOR SEQ ID NO: 436:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3271 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 436:

	TAAAAACTTA CTTTAACACC ATTCCTTTTT AACTTTTTTC GTGTTTCnTT TnCTTAAGTC	60
10	CATCCATATT TTTAATGATG TCATCTGCTG TTTTATCTTT TAAATCTAAC ACTGaGTGAT	120
	AmCGAATTTG TAGCACAGGA TCAAATCCTT TATGGAATCC AGTATGTTCA AATCCTAAGT	180
	TACTCATTTT ATCAAAGAAC CAATCATTAC CAGCATTACC TGTAATCTCG CCATCATGAT	240
15	TCAAGTATTG ATATGGTAAA TATGGATCGA TATGTAGGTA TAGACAACGA TGTTTTTTTAA	300
	CATATTTTGA TAATTCATTA AAGAAAAAGT GTACGAGTTC TTGATTTTCA TAATCAATCA	360
	CTGGACCGCG ATTTGAATAA AAATACTTGA AACTTTTCAT AACAGGTACA GCAGTAAGTA	420
20	AGCAAGCTGC AATGACCTCG TTATTATTGT TTTTATTTC CACTAAATGT GTTTCATAAC	480
	CTTCAGCAAG CTTTAACTCA TAGTGGCCAA CAGTTTGCCT GaaATGACTG TATGGCATGC	540
25	TATCTGTAAa GGCACCAAAC TCTTTAGCTG TTAAATTTGT AAACCTTCATT ATCATTACTC	600
	CTATTTGTCT CTCGTTAATT AATTCATTT CCGTATTGc AGTTTTTCTA TTTCCCCTCT	660
	GCAAATGgCA AAAATAATAA ATCTAATCTA AATAAGTATA CAATAGTTAA TGTTAAAACT	720
30	AAACATAAA CGCTTTAATT GCGTATACTT TTATAGTAAT ATTTAGATTT TTGAATACAA	780
	TTTCAAAAAA AGTAATATGA ACGTTTGGGT TTGCTCATAT TACTTTTTTT GAAATTGTAT	840
	TCAATTTTAT AATTCACCGT TTTTCACTTT TTCAAACAGT ATTCGCCTAA TTTTTTTTAA	900
35	TCAAGTAAAC TTAATTATTC AATGTTTGTT GGATAGATTG TAAATATTTA ATGATTTTCCT	960
	CACGCGTGTT AGATTTAAAT CGCTTAACGA TTTGCTACC AATGACAATG CCATCTGCAA	1020
40	CCTCTTTTAT ATCTGCAACA TGTTGTGGTG TTCTTATACC AAATCCTGCG ACAACTGGCA	1080
	CATTGGCTAT CGCTTTAATT GACTCAATTT TTCGTTTAA TTCTGGATGA AACGCACCGT	1140
	TTTGCCCTGT TGTCGCATTC ATCGTCACAG TATAAATAAA GCCTTCCGCA TGGGATACGA	1200
45	TATCTTTTAT ACGTTTGTCa TCAGTAGTCA TCGCAACTAA CGATATGATT TTGACGCCAT	1260
	AGTGACTAAA TTGTTGTTTT AAACGCTGCG ATAATTCATA TGGTAAATCA GGAATAATTA	1320
	AGCCGTAGAC ACCAGTATCT CGACATTTTT CAAAAACGC TTGTTCTCCA TAATGACAAA	1380
50	TAATATTATA ATACGTCATT AATACATAGT TACACTTAAT TTGATCACCA TGTTTTTCTA	1440
	ATTGATTGAA AATATAATCT ATCGTGATGC CTTGTTTAAAT CGCTTGTTGA CCTGCTTCCA	1500
55	TGATAACTGG ACCATCAGCA ACCGGATCAG AGAAAGGTAC TCCAATTTCA ATTATATCTG	1560

GTATAAATAA TTTAGTCATT TGCAAGACCT CGCTCTACCA TATATTGTCT AATTGTTTTCC 1680
 ATATCTTTAT CGCCACGTCC AGAAATAGTT ACTACAATAA TATCTTCTTT CGACATCGTA 1740
 5 GCGCTAGTC TTTCAACATA ACTCAGTGCA TGTGCACCTT CAATTGCAGG TATAATACCT 1800
 TCATGTTTTG TAAAGTTGAT TAAAGCATTG ATTGCTTGTG TATCACTAGC ATTTTCAAAA 1860
 GTTACTCTAC CAATGTCGTG GTAATAAGAA TGTTCGGTC CAATACCAGG ATAATCAAGT 1920
 10 CCTGCTGAAA TAGAATGTGC TAGTTGCACT TGCCCATCTT CATCTTGAAT TAAATACATT 1980
 TTAGTACCAT GTAATACGCC AGGTGATCCT TTGCCAATTG CAAGTGCATG TTTATCAGTA 2040
 15 TCATCGCCTT GACCTGCGGC TTCAACACCG TATAATGCAA CATCATCTTT AATAAATGGA 2100
 TAAAATGTAC CGATTGCATT TGAGCCACCA CCGATACATG CTACAATTGC ATCCGGAAGT 2160
 CGACCTTCTT TCTTCAATAT CTGTGATTTT ATTTCTTTAC CAATCACACT CTGAAAATCT 2220
 20 CTAACAATCG TTGGGAACGG GTCTGGACCT AATGCAGAAC CTAATAAATA ATGTGTATCA 2280
 TCTACATGAC TTACCCAATA TTGCAATGCT TTATTAACTG CATCCGATAA AGTCCCTTGA 2340
 CCATCTTCAA CTGCCACAAC CTTTGCACCA AGTAATTCCA TTCTAAATAC ATTAAGTTGT 2400
 25 TGTCTTTTAA TATCTTCACT TCCCATAAAG ACAACAAGTT CCATATCAAA TAATGCAGCA 2460
 ACCGTAGCAC TAGCTACACC ATGTTGACCC GCACCAAGTT CAGCAACAAG CTTCTTCTTG 2520
 CCCATTCTTT TAGCAAGCAA CGCTTGACCT AACGCATTAT TAATTTTATG GCGCGCTGTA 2580
 30 TGATTTAGAT CCTCTCGTTT CAAATATATT TTAGCGCCAC CTAGGCTTTC AGTATATGAT 2640
 GCAGCATATG TAAGTGGTGT CGCGCGTCCT ACATACTCTG ATAAATAGTA TTCCAGTTCT 2700
 35 CTTTGAAACT CTGGGTCTGC TTTTGCCTCT TTATAAGCTT TTTTCAACTC AATAATTGCT 2760
 GGCAPTAATG TTTCTGGAAC ATATTGCCCT CCATATTAC CAAAGAAACC TAATTCATCT 2820
 GCTTCTGTTT GTATTTGTTT ATTCAATTGC TCTATCTCCT TTCACAATAT TTACAATTGC 2880
 40 TGTCAITTTT TCTATATCTT TTCGCCCATT TACTTCTATA CCTGATGCAA GATCATAACC 2940
 TTGATGTGAT AATTTAAGTT GATTAAGTGT TTGAATATTT TCAGAGTTAA TGCCTCCTGC 3000
 TATCAAATAA GGTATGTCTT TTATGTGCTT CAAAATAGTC CAGTCATATG TTTGACCGGT 3060
 45 ACCACCATAC GACACTGAGG GTGTGTCGAT AATAAATAAA TCTACGAACC CTTTATATTT 3120
 ATTTATGTTT TGGATTATGT TTTTCATCTG CAGTAAAGCT TTAGTGATTT TAATGCTTGA 3180
 ATATkTCTTT TTAATTTCTT GTaAAAAaC AATAGATTCT GtGTAAGTGT ATTGTGkTAA 3240
 50 wTGaCGwATg CtTAAwACgT GTGCCAATGG T 3271

(2) INFORMATION FOR SEQ ID NO: 437:

(A) LENGTH: 1553 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 437:

10	ATAATGAAAT AATACTGTGT TTTATCTGCG AAATGTATCA TTTTCTAATc GTTTCACAGT	60
	AAAATGAAAA GATAAAGTGT GTTTTTACTT GAATTTTGAC TAAAATTACT CTATATTTAT	120
	TAATTGAGCT ATGCTTATTA TTACAATTTG ATTACAAATT TTAAATTTGT TAATTGAATG	180
15	ATAATATTAA ATAAAGAAAC TTACACAAGC AAATATGAGT TGTAGCCCAA AATACTTGTT	240
	AAATCAAAGT TGAAAGCTAC AAATAATGAA AATTATAAAC TTGAATCTGA AAGTAATTAC	300
	TATAATTATG ACAATGTTAA CTTTTAAACG CACTTATTAA TTAACACAT AATGTTAATA	360
20	TCTAATTTAT TCAAGTACTT TCGCAAGATT TATTATCTAA ATAACGGGGG AAAGAATCAT	420
	GAGTtCACAA AAAAgAAAAT TAGTCTTTTT GCGTTCTTCT TATTAACCGT AATAACGATT	480
25	ACCTTGAAGA CGTATTTTTT TTATTATGTT GATTTTTCTT TAGGTGTTAA AGGTTTAGTA	540
	CAAACTTAA TATTATTGAT GAATCCTTAT AGTTTAGTAG CACTGGTTTT AAGTGTGTTT	600
	CTATTCTTTA AAGGCAAAAA AGCATTTTGG TTCATGTTCA TAGGCGGCTT CTTATTGACG	660
30	TTCTATTAT ATGCCAATGT TGTGTACTTT AGATTCTTCT CTGATTTTTT AACGTTTAGT	720
	ACTTTAAACC AAGTAGGTAA CGTAGAATCT ATGGGTGGTG CGGTTAGTGC ATCATTCAAA	780
	TGGTATGACT TTGTTTATTT CATTGATACG TTAGTTTACT TATTCATTTT AATATTTAAA	840
35	ACAAAATGGT TAGACACAAA AGCATTTAGT AAGAAATTTG TTCCTGTCGT AATGGCGGCT	900
	TCAGTAGCAT TATTCTTCTT AACTTAGCT TTTGCTGAAA CTGACAGACC AGAATTATTA	960
	ACACGTACAT TTGACCATAA ATATTTAGTG AAATATTTAG GACCTTATAA CTTTACAGTA	1020
40	TACGaTGGTG TTAAAACTAT CGAAAATAAT CAACAAAAAg CGCTAgCATC TGAAGATGAC	1080
	TTAACaAAAG TATTAAAtTA TACGAAAcAA CGTCaAACmG AGCCTAACCC rGawTATTAT	1140
45	GGGGTGGcAA rGAAGAAAAA TATTATTArG ATTCATTTAG AAAGTTTCCA AACCTTCTTA	1200
	ATTAATAAAA AGGTTAATGG TAAAGAaGTA ACACCGTTTT TAAACAAATT ATCAAGTGGG	1260
	AAAGAGCAAT TCACATACTT CCCTAACTTT TtCCATCAAA CAGGTCAAGG TAAAACATCT	1320
50	GACTCTGAAT TTACAATGGA TAACAGTTTA TACGGTTTAC CGCAAGgTTC TGCCTTTTCA	1380
	TtaaaaGGAG ATAATACGTA TCAGTCATTA CCAGCAATTT TAGATCAAAA GCAAGGCTAC	1440
55	AAATCTGATG TCATGCACGG TGACTATAAA ACATTCTGGA ACAGAGACCA AGTATATAAA	1500

(2). INFORMATION FOR SEQ ID NO: 438:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1419 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 438:

ACAAAATCAA TCAGCAAAT GATTGGATTA AGACGCAAAT TGAGCGTTCA ATGGAAGGCG 60
 AAACAGTTGG CATTAAATGAT CAAAATATAG AAATATATAG TGAACGtTgA GATTTATACC 120
 ATACACTCGT ACCTTTAAAT CAAGAATTGC ATAAGTTGCG ACTTAAACT CAAAACCTTAA 180
 CCAATGAAAA TTATAATATT AATGATGTGA AAGTTAAAAA GATTATTGAA GATGAACGTC 240
 AAAGACTAGC ACGAGAACTT CACGATTCTG TTAGTCAGCA ACTTTTTCG GCMAGTATGA 300
 TGCTATCTGC TATCAAAGAA ACGAATAGA ACCACCAITTA GACCAACAAA TTCCTATTTT 360
 AGAGAAAATG GTTCAAGATT CGCAGTTAGA AATGCGTGCT TTGCTGTTAC ATTTAAGACC 420
 GCTTGGTTTA AAAGACAAAT CTTTAGGTGA GGGTATTAAA GATTTAGTTA TTGATTTACA 480
 AAAAAAAGTG CCAATGAAAG TTGTGCATGA AATACAAGAT TTTAAAGTGC CTAAAGGTAT 540
 TGAAGATCAT TTGTTTCAGAA TTACACAGGA AGCAATTTTCG AATACATTGC GTCATTCAAA 600
 CGGTACAAAA GTGACAGTAG AATTGTTTAA TAAAGACGAT TATTTATTGT TGAGAATTCA 660
 AGATAATGGT AAAGGTTTTA ATGTTGATGA AAAATTAGAA CAAAGTTATG GACTTAAAAA 720
 TATGCGTGAA AGAGCTTTGG AAATTGGTGC AACGTTCCAT ATTGTATCAT TGCCAGATTC 780
 AGGTACACGT ATCGAGGTGA AAGCACCTTT AAATAAGGAG GATTCGTATG ACGATTAAAG 840
 TATTGTTTGT GGATGATCAT GAAATGGTAC GTATAGGAAT TTCAAGTTAT CTATCAACGC 900
 AAAGTGATAT TGAAGTAGTT GGTGAAGGCG CTTCTGGTAA AGAAGCAATT GCCAAAGCCC 960
 ATGAGTTGAA GCCAGATTTA ATTTTAATGG ATTTACTTAT GGATGACATG GATGGTGTAG 1020
 AAGCGACGAC TCAGATTAAA AAAGATTTAC CGCAAATTAA AGTATTAAATG TTAAGTAGTT 1080
 TTATTGAAGA TAAAGAGGTA TATCGTGCAT TAGATGCAGG TGTCGATAGT TACATTTTAA 1140
 AAACAACAAG TGCAAAGAT ATCGCCGATG CAGTTCGTAA AactTCTAGA GGAGAATCTG 1200
 TTTTGAACC GGAAGTTTTA GTGAAAATGC GTAACCGTAT GAAAAAGCGC GCAGAGTTAT 1260
 ATGAAATGCT TACAGAACGA GAAATGGAAA TATTATTATT GATTGCGAAA GGTACTCAAA 1320
 ATCAAGAAAT TGCTAGTGCA TCGCATATTA CTATTAAAC GGTAAAGACA CATGTGAGTA 1380

(2) INFORMATION FOR SEQ ID NO: 439:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 608 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 439:

ATGTnGGCGG ACGTAAAGGA CATGTTTATA CTGATGATCG AGCATTAGAT ATTGATATCG 60
 TTCCGCCTGC TCAAGCAGAT GGTAAAGCTA CTAACCCCGA ACAATTATTT GCAGCAGGTT 120
 ATGCATCTTG CTTCAACGGT GCTTTTCGACC TAATTTTAAA GCAAAACAAA GTGCGTGATG 180
 CTCATCCAGA AGTAACACTA ACAGTGAGAC TAGAAGATGA TTCAGACTCA GAAAGTCCTA 240
 AATTAAGTGT TTCAATTGAT GCGACAATTA AAAATGTTAT ATCTCAAGAA GAAGCTGAAA 300
 AATATTTACA AATGGCTCAT GAATTTTGTC CATATTCAA AGCGACTCAA GGAAATATTA 360
 ATGTCGATTT AAATGTAAAT GTTGTAGATT AGCATTAACT TAAAGAGATT ATTCAACGTT 420
 ATTAATAAAA TTCACATAAA ATTCAAATt stCrACCAAA AATTTTGGT TGGyTATTTT 480
 TTCTATTCGT GATTGAAATT TCTGGCAATA TTAACTGAA AATGATTGTA CCTTAGTCAT 540
 CATAAATGTG ACCGGTTCCA ATACTGGCTT GACTTCTTCG CATACCGTCT ACAAATAAAA 600
 GTCCAGTG 608

(2) INFORMATION FOR SEQ ID NO: 440:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 682 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 440:

CCTTTCAAtT TCTCCCAAgA TTTACGCATG TCTGACCAAA TGTCAATACA ATACCTGAAT 60
 AATGATTCTT GGAGTAATTG CGAAAAATAT TTAAAAGTTG CTATAGATCA ATTTTCAAAT 120
 TATAGTATCT CATCTCAAGT TTCTAACTAT CATTTTCACAG TATTACTTGG AGATCGCCAA 180
 AAACCACTTA TGTATCTAAA TAAAAATCGC GGTGGTGATG GTGGCATACC AGGTTATATT 240
 ATGATTTATT TAGTGCCGAG TACAAGTACA ATTAATTCTA TGAAAAGCTT AATTGCACAT 300
 GAAGTAAATC ATAATATGCG CTATCAATAT ATTGATTGGG ATGGCGGAAG TTTGATTGAA 360

ATTGGACCGT GGGTAACAAA TACTAATTGA AGTCGTGATA ACGTAAAAAT TAAAAATACT 480
 ATTTATAATC ATTTACATTT AAAGCATATA TTTGAATCGA tGCCTTATCT CTATGGTGAT 540
 5 GATATTAATA AACTTCAAGG TAGGCCTATC GTTGGCTTAT CTCATGCTGC CGGGTATGCA 600
 TGTGGCTATC ACTTGGTAAA ATACTTTTTA CAAAAACAA ACATACCTAT TGAAGTTGCT 660
 ACAACACTTC CAGCACAAAA AA 682
 10

(2) INFORMATION FOR SEQ ID NO: 441:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2574 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 441:

ATCTTCTTTT AACATGACAA ATTGCAACAA AATATTAATG ACGCCAAAAG ACATTTTTTC 60
 ACGTTCAATT AATTCTTCAA CCATTGTCTT TTGCGATATA GTTGGTTCTG ATTCAGACCA 120
 25 AGAAGCTAAC ATATCAATTG GACTCGTTTG TTCAAGTAAC TCAAACCATT CATCACTTTG 180
 TGGCTTTGGA TTCACTTCTG AAGATTTGCC CGCCGAAGAT GATGTAGCAG GAGATTTAC 240
 CTGTAATTTA GGCATTTGAT TTTCGTGTTT CATTAAAGTAA TACGAGCGTG CTTGTTTACG 300
 30 CATTTCTTCA AAGGATAACT GTTGTCCTT TGTAATTGAA TTTAAAATAA CATGCTTCAT 360
 GCCATCTGCT GTTAAACCAT ATAAAGTCGC GAGTTGTGTA ATTAAACGCT TTGCATCTTT 420
 GGTAACAATG TCTTGACTAA TAAAATGTTT ACCTAACATT TGTCTCAACA TTTCAAAGTC 480
 35 AAAAGATTCA TTTGATAAAT CGATACCTTG GTACGGTTCA TTAATCGGaa TATCACTTGT 540
 ATCGATATCT ATTTTGTAG ACGGCACTTT AAAACATCA GTAAATTGTC TTGTTACCTG 600
 40 TTTAAATTCA CTCAAATCAA TTTGTTGATA CTCAAAGTAT TTCTTCAACT CATGAAATCG 660
 ACGATGCTCG ACTTCACTAT ATAAAAGAT TGACAACATT GGATCATTAA AAAATAAATG 720
 TGCTGAAGGG GtTGaATTAA TTGGTAAACA AATTGTGTTT CTTGTTCATC ATGTTTGACA 780
 45 AACGCCTTTA ACAATCCAAT CGCTTCAAGT AAGTCCATTT GTTGTCTAAA CTCTAGTAAA 840
 TTAATTTTAA GTTCATTCAT AAAAAATAA TGAGAAAGAA TCAATGTTTC ATTATGACTT 900
 TCTTTAACGA ATTGAGTCAT AAAATGATAT AAACCCACTG CTTGCGTTCC AATTAGCGGT 960
 50 GTATACAGTC GATTCAATAC CTCTAAATGA TTCGTATTTA AATCAAAGTG TTGCATAACT 1020
 TTGAATTGAT CCTTTGGTCT TAAGCCGAAT TCGAAGGCTT GTCGTCCCaT TTAAGChATC 1080

55

AAATTCTTTA TAGACTGATG CAAATCTAAC ATATGAAACT TGATCAACAT GCATTAAACAA 1200
 GTTCATAACG TGTTACACCTA TATCTCGTGA AGACACTTCC GTATGACCTT CATCTCGTAA 1260
 5 TTGCCATTCA ACCTTGTTAG TTATGTCTTC AAGTTGTTGA TATCTAACTG GTCGTTTCTC 1320
 ACAAGAACGC ACAAGTCCAT TAAGTATCTT TTCTCTTGAA AACTGCTCTC TTGTGCCATC 1380
 TTTTTCACA ACTATAAGCT GACTAACTTC GATATGTTCA AATGTAGTGA AACGTGTTCC 1440
 10 ACAATTTTCA CATTCTCTTC GTCTTCGAAT GGCATTTAAT TCATCGGCAT GCCTTGAATC 1500
 TACAACCTTA GATTGTGTAG AATTACATTT CGGGCATTTT ATTACATCAC CCTCTTTATT 1560
 TTGATTATGC CTAATTATAC TATAAATCTA GAGATGAAAA AAGAATCCCT CAATTTAATT 1620
 15 CATTTAACCA AATAATGAAA CAATAAAAAA CATTATATCG TTAATTATTA AGTAATTTGC 1680
 ATGACAATAT TATTGTATTA AAAATAAAAA ACCTAACTCC GAAGTCAGAG TTAGGCTATA 1740
 20 AATTAATTGT ATTAACCTGC ACTTACAGTT TCTTTTGATG TCAAAAGTGC TCCAATTTGC 1800
 TCAGCAACAT CTACAACCTT ATTTGAATAA CCCCATTCAT TATCATACCA AGCAATAACT 1860
 TTTACTTTAT TCCCTGACAT GACCATTGTT GATTTTGCAT CAATAATAGC TGAATTTGGA 1920
 25 TTAGTATTAA AATCAACAGA CACTAGTGGT TGATGTTCTG CTCTATGAT ACCTTCTAAA 1980
 CCTGCATTTT CAAAAGCTTG GTTTACTTCT TCTGCAGTTA CTCTTTTTTC TAAATCAACA 2040
 ACTAAATCAA CGAGCGATAC ATTCTTTGTT GGTACACGTA ATGCCATGCC GTGTAATTTA 2100
 30 CCTTCTAATT CTGGTAATAC TTCTTTTAAA GCTTTCGCCG CACCAGTAGA AGTAGGAATA 2160
 ATGCTTTCAT TACATGAACG TGCACGTCTT AAATCTTTAT GTGGATTATC AATATTTTTT 2220
 TGGTCATTTG TAATAGCGTG AACAGTAGTC ATTAAACCAT TAACTATTCC AAAGTATTA 2280
 35 TTTAAACTT TTGCAACTGG ACCAATGCAA TTAGTAGTAC ATGAAGCATT ACTAAAAATG 2340
 TCAAATGCTT CTATATCTAA TTGGTTATCA TTTACGCCTT TAACTACCAT TTGAACATGT 2400
 40 CCACCTTTTG nAGGACCAGT TAACAAAaT TTTtTGGCAC CTGCTTTAAT ATGTGCGATG 2460
 GCTTTATCAC CATGaTTAAA TTTACcAGTT GCATCTATAG CAATATCGAT ATCTAATTCT 2520
 TTCCATGGCA AGTTTTcAGG ATTGCGATCA GCAACCAATT TAATTTTATG ATCT 2574

45 (2) INFORMATION FOR SEQ ID NO: 442:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3326 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - 50 (D) TOPOLOGY: linear

55

	CAAGGTACCG	GTTTAGGTTT	GTTTATTGT	AAAATGATTA	TCGAAGAGCA	TGGTGGTTCC	60
	ATAGATGTTA	AAAGCGAATT	AGGGAAAGGC	ACAACATTTA	TTATTAAACT	ACCAAAACCA	120
5	GAATAAAATT	GAATATAGTT	ATTTCAGAAC	GCATGTTATT	GATTAGAGAC	TCTAATTTAT	180
	AGCATGCGTT	TTTTGATTGA	TGTGGGGAAT	TTGTATGTG	GATTAGAACT	TAGGGTTTTT	240
	GCGAATATCA	ACTATTAAAT	ATATTACTAA	TTTATACAAA	AATATAAAGT	TTGATAAAGT	300
10	TATTTATTTG	ATTATAAAAA	TAGGGTAAAA	TATAGATATA	TTGTATTAAT	TAAATTATTC	360
	GAGGTGTCAT	ATGAAAAAAT	TCATTGGATC	AGTTTATAGCT	ACGACATTAA	TTTTAGGGGG	420
15	ATGTTCCACG	ATGGAAAATG	AATCAAAAAA	AGACACAAAA	ACAGAAACAA	AATCTGTACC	480
	AGAAGAAATG	GAAGCTTCAA	AATATGTAGG	CCAAGGCTTC	CAACCGCCTG	CAGAAAAAAA	540
	TGCGATTGAA	TTTGCGAAGA	AGCATCGTAA	AGAATTTGAA	AAAGTAGGTG	AACAATTCTT	600
20	TAAAGATAAC	TTTGGAATAA	AAGTTAAAGC	TACAAATGTT	GTAGGTAAAG	ATGATGGTGT	660
	AGAAGTTTAT	GTGCATTGTG	AAGATCATGG	CATTGTATTT	AATGCAAGTC	TACCTTTGTA	720
	CAAAGATGCC	ATCCATCAAA	AAGGATCAAT	GCGCAGTAAT	GACAATGGTg	ATGATATGAG	780
25	TATGATGGTG	GGTACAGTGC	TGAGTGGCTT	TGAATATCGA	GCGCAAAAAG	AAAAGTATGA	840
	TAACTTATAT	AAATTCTTCA	AAGAAAATGA	AAAGAAATAT	CAATATACAG	GCTTTACAAA	900
30	AGAGGCAATT	AACAAGACAC	AAAATGTCCG	ATATAAAAAT	GAATATTTTT	ATATTACATA	960
	CTCTTCTAGA	AGTTTAAAAG	AATATCGAAA	GTATTATGAA	CCACTGATTC	GAAAAAATGA	1020
	TAAAGAATTT	AAAGAAGGAA	TGGAACGAGC	AAGAAAAGAA	GTGAATTACG	CTGCAATAC	1080
35	AGATGCTGTT	GCTACACTTT	TTTCTACTAA	GAAAAACTTT	ACTAAAGACA	ATACAGTAGA	1140
	TGATGTAATC	GAACATAAGT	ATAAATTATA	TAATTTAAAA	AATAAACCAG	ATAAATCTAC	1200
	AATCACAATA	CAAATAGGGA	AACCCACTAT	TAATACTAAG	AAAGCCTTTT	ATGATGATAA	1260
40	TCGTCCAATA	GAATATGGGG	TGCACAGTAA	AGATGAATAA	AATTAATGAT	AGGGATTTAA	1320
	CAGAATTAAG	TAGTTACTGG	GTTTATCAAA	ATATTGATAT	AAAAAAAGAA	TTTAAAGTTA	1380
	ATGGAAAAAG	GTTTAAACAA	GTAGACAGTT	ATAATGATGA	TAAGAATAGT	AATTTGAATG	1440
45	GTGCTGCTGA	TATTAAAATA	TATGAGTTAT	TAGATGATAA	AAGTAAACCA	ACTGGTCAAC	1500
	AGACAATAAT	TTATCAAGGA	ACATCTAATG	AGGCAATTAA	TCCAAATAAT	CCATTAAAAT	1560
50	CATCGGGGTT	TGGAGATGAT	TGGCTCCAAA	ATGCTAAATT	AATGAATAAT	GATAATGAAA	1620
	GCACAGATTA	TTTAAAGCAA	ACAGATCAAT	TATCAAATCA	ATATAAAATA	AAGTTAGAAG	1680
	ATGCAGATAG	ATTATCAAAT	AGTGATTTTT	TAAAAAATA	TAGAATGGAA	TCAAGTAACT	1740
55							

ATCAAGGAGC GAAACATCCG AATGAAAAAG TTGTTGCTAC TGA CTCAGCA ATGATTCCTT 1860
 ATGCTGCTTG GCAGAAATTT GCTAGACCAC GCTTTGATAA TATGATTAGT TTTAATAGTA 1920
 5 CCAACGATTT ATTAACATGG TTACAAGATC CATTCaTCAA AGATATGCCA GGAAAACGCG 1980
 TTAACATTAA TGATGGTGTG CCCAGGTTAG ATACTTTAAT AGACAGCCAT GTAGGTTATA 2040
 AAAGGAAGTT AAATAGAAAA GATAACACAT ACGATACTGT ACCACTAATC AAAATAAAGT 2100
 10 CGGTAAAAGA TACAGAAATT AAAAATGGAA AAAAAGTAAA AAAGACTATT AACATAACAT 2160
 TAGATATGGA TGGGCGAATT CCAATAAATG TTTGGACAGG AGATTGCGATT GCACGTTCTG 2220
 GAAGAGGAAC TTTAATTAAA CTTAATTTAG AAAATCTTGA TCGGTTGAGT AAAGTATTAA 2280
 15 CTGGTGAAAC AAGTGGTATG TTAGCAGAAT GCGTAATCTT TTTAAATGAA AGTTTTAACA 2340
 TCTCAGAAAA TGAAAATAAA AATTTTGCAG ATAGAAAGAA ACAATTATCA GAAGGATTTA 2400
 20 AGGATAAGAT TAACTTATTT CAGTTAGAAG AAATGGAAAG AACTTTAATT AGTAAAATAA 2460
 ACTCACTTGA AGAAGTTGCA GATGAAACAA TAGAAAGTAT TAGTGCTGTT AAACACTTAT 2520
 TACCTGATTT TGCATTGGAT GCATTAAAAG AAAGAATTAA TGAGTTGTTT AAAGGTATAA 2580
 25 AATCTTTTAT AGAAAAAGTG TATGATAGTA TAGATAATGA AATTTTAGAA ATTTTCAAAA 2640
 ATATAGATCA CGACTTCAGA GATGGAGTAT CTGAAGAAAT GATGAAACAT TTGAAAGTAG 2700
 TGAAACAGAA TATAGACCAA ATAAAAAATC AAAATGATAT TTATGGTAGG CAAATTGCAG 2760
 30 ATATTAGAAG TATTATGAAA CAACAAGATG CAACAATTTT AGATGGAAAT TTTCAAATTA 2820
 ATTGTAGCGG CGAAAATATG GTACAGGGTC TAGTTATACC TTCTAATTAT TTAGGAAGAA 2880
 AAATGAAAT ATTTAAAGAC CATATCGATG ATGGTATTAA AAAAATAGCA GACTATGTTT 2940
 35 AAGGTATATA TGATGAATAT GCATCGAAAA TTGTCGATGT AATAAAATAT TTGATTAATA 3000
 CAATTCCCAA AATACGTAAG AATTTAAGAC ATGCAATTGA AATGTTAAAT GTAAAAAGA 3060
 40 AAGAATTTTT GTCCCTGATT CCTAATGTAA CTTGTAATTA TATTAAACT AAATTAGAAG 3120
 AATTAGATAA TACTTTAGGC AAATGGGAGC CTTTCTTAA TGATTTAAA GCAGTGTCAC 3180
 CAATTTTGA TAACCATTTA GATGATATTG TTAAGAACAT GAAGCCTTTG ATTGTACAAA 3240
 45 TGATATwTGA ACCATCACAT TATGACGATA TGTTTAATTC aAGAAAAGCT TTAACGcCAG 3300
 TGTTCCTCAAG CGTTTTATAA AGGTTG 3326

(2) INFORMATION FOR SEQ ID NO: 443:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5301 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 443:

5	ACCTATAGCT GGCCCGTCAA GTCCTGAATC ATATGTCGCT ATAAACCCAG TTGCATGCCC	60
	AGCAATCTCG GTTTC AATTT CAAAATCATG CTCTTTCAAA CGATCTATTA AAGTTGAGAG	120
	CGCAAATATT TCTTCATTAC CAAGTTCAGG ACGTTCATGA ATTCTATGAC TGATTTTCGAT	180
10	ATAACTATAT TTATTTGTTT CTATATAATC GAGAATTTGT TGTTTTTCAC TCATTTTTTA	240
	CTATCTCCCT TTTACCCTAC ACTCATTTTA ATCATCCACA ATATTTTGTT CTTTCAAAAT	300
	GAATTATTAC TTATTCTATC GGTTTTATCT CATGATGTCA TCTAGTTTTT CTTTATTATC	360
15	AAAATTTTCT AATAACTAAA aGCCTTTCAT AAATTTATAA AACAGTTTCA AATTGTAAAA	420
	CATAAGCCCT ATTGTTACAA TTTAAGTATT GTTAGAAACA CAACTTACAA CAAACTTAAG	480
20	TTTTATAATG CAACAAATCA TAAGCGTTTT ATAGTTTTGA GTAGTTAATA GGAGGAAATT	540
	CAAAATGACA AAAATGAATG TTGAAAGTTT CAATTTAGAT CaTACTAAGG TGGTTGCCCC	600
	ATTTATTAGA TTAGCGGGAA CGATGGAAGG ATTAAACGGA GATGTCATTC ACAAATACGA	660
25	CATTCGTTTC AAACAACCAA ACAAGAACA TATGGATATG cCCGGACTAC ATTCATTAGA	720
	ACATTTAATG GCTGAAAATA TTAGAAATCA TAGTGACAAA GTTGTTGATT TAAGTCCTAT	780
	GGGTTGCCAA ACTGGTTTCT ATGTATCATT TATTAATCAT GATAATTATG ATGATGTATT	840
30	AAATATTGTT GAAGCAACTT TAAATGATGT GCTAAATGCT ACTGAAGTGC CTGCTTGTA	900
	TGAAGTACAA TGTGGCTGGG CAGCAAGTCA TTCATTAGAA GGTGCTAAAA CTATCGCTCA	960
	AGCATTTCTA GACAAACGAA ACGAATGGCA TGATGTTTTT GGTACAGGAA AATAAATCTT	1020
35	AGTCAATCAA GTTAATCAGA AAaGCAGTCG AACAATGATT TTACAATCGC CATTGTCCAA	1080
	CTGCTTTTTA TTATGCTTCA AAGTCAAAAA ATCGAACAAA TGAAAAAGTA AAATCTTTAA	1140
40	CATTTGTCCG ATTTATTTGA GAACCACTAT AATTTCTTAA TTAGTCCCAT TAACACGAAC	1200
	TGCATAGGTA ACCTTAAATA TAGTTGCCAT GTTGGCAATT GTTTATCTCC TAAAGGTAAC	1260
	TTTTTAACTG CCATATAGAT ATTAGCTGGG AATACAGCTA GCAAGAATAG ATTGATTGTA	1320
45	TTTTTCAAGC ATTGAGATGG TCTTTTAATT AAAAGTGCAA GTCCAAATAA TATCTCAAAG	1380
	ATTCCTGTAA CAAGAACCGC TGTTTTTCGA AGTGGCAAAC ATTTCCGGTAT GATATTCTA	1440
	AATTGTCGTT CTCGTGTAAA ATGCAATACA CCTATTACAC TAAAACCTAT TCCTAATAAA	1500
50	TATCTTAGTA TGTTC AATCA GCTTCAACTC CTATTCTGTA ATGATTTTAT GAATTAATGT	1560
	AGGCGATACA ACATGaTCAG CAATTGTTAT GCTTGAATCT AATTTTTTAA CAACATCGTC	1620

55

	AATTTTTTTA	TTTAAAACAA	TACCAACCGC	TAAATCAATA	TCATCCTCTT	TTGTTAAACG	1740
	TCCCGCTCCT	AACATCATCG	AAGCGACACC	TATATCGTTA	GAGACTAATT	CAGTCACATA	1800
5	ACCTGATTTT	TTAGCTTTAT	ATTCAATTTG	ATATTGAGCT	TGTGGCAAAC	GCTCTGGATG	1860
	GTCAATAACA	GTTTCGTCGC	CACCTTGGTT	TTTAATAAAT	GTTTTGAATT	TTTCTAATGC	1920
	TGCACCTGAA	TTAATTGCCT	CAATTAGCAA	CGCTCTCGCT	TCTTCAAGCG	TTTCAGCTTT	1980
10	GTTTGCAAGT	ACAACCATTT	GAGAACCTAA	TGTTAATACA	AGTTCTGTTA	AATCTTTTCG	2040
	ACCTTGTCCT	TTCAACGTAT	CAATTGCTTC	TTGTAACTCA	AGCGCATTGC	CAATCGCAGC	2100
	TCCAAGTGGC	TGATTCATAT	CAGAAATAAT	CGCCATCGTA	TTACGTCCCA	CATTATTACC	2160
15	AATACGTACC	ATTGCGTGCG	CTAATGCTTC	AGCATCTTCT	AATGTTTTCA	TAAATGCACC	2220
	GCTACCAGTT	TTTACATCTA	ATACAATTGC	ATCTGCACCA	GCAGCAATCT	TTTTACTCAT	2280
20	AATTGAAGAG	GCAATTAATG	GTATTGAATT	GACAGTACCA	GTAACATCCC	TTAAGGCATA	2340
	TAATTTTTTG	TCTGCAGGAG	TTAAATTTCC	TGATTGTCCT	ACAACTGCCA	CTTTATTTTC	2400
	ATTAACCAAT	TTACACAAATG	TTGCTTCATC	TATTTCAACA	TGAAAACCAT	CAATTGCTTC	2460
25	TAATTTATCA	ATCGTACCAC	CTGTATGACC	TAATCCACGC	CCACTCATTT	TTGCAACAGG	2520
	AACATCTACA	GCTGCTACTA	ATGGTGCTAA	AACCAATGTA	GTTGTATCTC	CTACACCACC	2580
	TGTTGAGTGC	TTATCTACTT	TGACACCTTT	AATATCACTC	AAATCTATCA	TATCACCAGA	2640
30	ATTAACCATA	GCCATCGTTA	ATGCTGCACG	CTCATCATCA	TTCATATCTT	GGAAATAAAT	2700
	CGCCATTGCT	AAACTTGATG	CTTGGAATC	AGGAATATCC	CCTTTAACAT	AGCCGCCAAT	2760
35	AAAGAAATTA	ATTTCTTCCG	TTGTTAGTGT	ATGACCGTCA	CGCTTTTTCT	CAATAATGTC	2820
	TATCAATTCTC	ATTTTTATCA	TCCTTTTCTT	AAAAAGCTTA	GGACAAAGCA	TCTGCGCTTT	2880
	CTCTAGTCCA	TTTTTAAAAG	CACAAGCGAA	AATTATTATA	GCAAGCTATC	GATTTCAATT	2940
40	TTAATATCAC	AATTTCAATG	CGATGTTATT	ATTCTTAAAT	AGATTGGTTA	TAACGTTAAA	3000
	GTCCCTATTA	AATTATCTTA	GAATCATCAT	GGCATTATG	ATGTCTTAAA	GCTGATATCG	3060
	ACATACTTAT	ATATGGTTAC	GATGTCCCAT	GCTTACATAT	TTTTATAAAA	TTAGTAATCT	3120
45	GAATCTGCTT	CTAAACCTTG	CATAATTTGa	ACGCCTGCGC	TCGCACCAAT	ACGTGTCGCA	3180
	CCTGCTTCAA	CCATTTTATT	GAAATCTTCT	AAATTACGTA	CGCCACCTGA	TGCTTTTACT	3240
50	TCTATATCAG	CACCTACTGT	ATCTTTTCATT	AATTTAACGT	CTTCTGCAGT	CGCACCGCCA	3300
	CCTGCAAAAC	CTGTTGAAGT	TTTAACGAAG	TCCGCACCAG	CCGCTTTTGT	TAATTCACCTC	3360
	GCTTTTACAA	TTTCGTCATG	GTCCAACAAT	ACCGTCTCAA	TAATCACTTT	TACTGTGTGA	3420

55

TTTAATGCGC CGATGTTGAT GACCATGTCA ATTTTCATCTG CACCATTTTG AATCGCATCT 3540
 TCTGTTTCAA ATGCTTTCGT CGCAGTTGTC GATGCACCTA ATGGGAATCC TATTACCGTA 3600
 5 CAAACTAACA CCTCTGAATC AGCTAGTCGC TCTGCTGCAT ATTTAACATG CGTTGGATTTC 3660
 ACACATACAG ATTTAAAATG GTATGCTTTC GCTTCATCGA TGATTTGATC GATTTGCGTA 3720
 CGTGTGACT CAGGCTTCAA TAAAGTGTGA TCAATCAATT TTGCACTATT CATTTTCTAT 3780
 10 CTCCTCCTTT ATGGTTGATT ATAAAAATAC GGTGTGTAAT TAGTTGATTG AGCGTCAGGT 3840
 TCATTTAAAT ATCAGGTTAG ATGTTGCTT TTTATGTAAC CGCATACATA TACTATTACA 3900
 TTAATTCATT TCCCATAAAC AAACAATACA ATTGAACGTG ATATCTTCAT TATGAACGAT 3960
 15 GACTTGACAA CAAGCTAATC AGGATTATAT TTTTATAATT CTTTAATTCT ATAGTACAAA 4020
 AATTCGCAA AAAGGGAAC AAATGTTATC TTAAAATTAT TAATGAATAT TAAGGAGAAG 4080
 20 ATAACAAATG ACAAAGGTA CACCACATAT TCAACCAAAT GGAGTAAAAA TTGCTAAAAC 4140
 AGTATTAATG CCTGGCGATC CGCTACGTGC AAAATATATT GCTGATAATT TTTTAGAAAA 4200
 TGTGAACAA TTTAACGATG TACGTAACAT GTTTGGTTAC ACTGGTACAT ATAAAGGTAA 4260
 25 AGAAGTTTCT GTAATGGGT CTGGTATGGG TATTCCAAGT ATTGGTATTT ACTCATATGA 4320
 GTTATACAAC TTCTTTGATG TAGATACAAT CATTCGTATC GGTCTTTGTG GCGCATTACA 4380
 AGAAAATGTT AACTTATACG ATGTTATTAT TGCACAAGCT GCATCAACTA ATTCAAATTA 4440
 30 TGTAGATCAA TACAATATTC CAGGTCATTT CGCGCCTATC GCTGACTTCG AGTTAGTAAC 4500
 TAAAGCTAAA AATGTCGCTG ACCAAATCGG TGCTACTACA CACGTAGGTA ACGTATTATC 4560
 TTCTGATACA TTTTACAATG CCGATCCAAC ATTCAATGAT GCTTGAAAAA AAATGGGTAT 4620
 35 TTTAGGTATC GAAATGGAAT CAGCTGGTTT ATATTTAAAT GCGATTCATG CTGGTAAAAA 4680
 AGCACTTGGT ATTTTCACAG TAAGTGATCA TATTTTACGT GACGAAGCTA CTACACCTGA 4740
 40 AGAACGTCAA AATTCATTTA CACAAATGAT GGAAATCGCT TTAGAAATCG CAGAGTAACT 4800
 TATTTAAATT GACTTTAATT GCTCTTTAAC AATGCGATTA AACTCAAAAA GCCAACACAT 4860
 TCTGGGCGTA TCCCCATTTA TGTGTTGGCT TTTATTTATA TTATTACTTA TCTGTAGATT 4920
 45 AGCTTAAGTA AGATTTAAAC ATCCAATTAT GTTTATCTAC TGATGTTTGC ATACCTATAA 4980
 ACATATCTTC TGATACATCA TCGCCAGCAT TACCAGCAAT TTCGATTGCG TTTTCTAATT 5040
 GTTTTGAGAT ATTTGTGAAG TCTTGTGATA ATTCTTCAAC CATTTGTTCT GCAGAGTAAC 5100
 50 CTTTCGCAGC TTCTTTAACA ATTGATTGCT CTAAGCATTC AGTTAATGTA CCTACAGGGT 5160
 TTCCTCCTAC CGCTAAAATT CTTTCAGCTA ATTCGTCTAC ATATTGGCTT GCTTCATTAT 5220

55

AATTGTGTAG CTTTGTGTAA G

5301

(2) INFORMATION FOR SEQ ID NO: 444:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11466 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 444:

15	GAAGATGAAA GCTATATTGA AGATGATGTT ACTAAAAAGG AAGCTATTTT AAGTATGCAA	60
	ATTCCTAAAG GTTCTCTCA AAAATTAAAA GAGAACCGTT TAAAAGAAAC GATACAGTTA	120
	TATGGTAGAG ATGACTTTAT AGGTGGTATT GCTGTAGAAA TTGTTAGTAG TTCATTATAT	180
20	GAGCAGCAAA TTCCTAACAT TATTTATGAA CACCTTGAGG ATATGAAACA GCATCAATCC	240
	ATCGATGCTA TCAACAAGTC CTATCATAAA CATACACCTG AATCTAAAAAT CAAATTTGTG	300
	TCGCTTACTA AACAAGCACA ACACTCTATT TCAATTAGCT TAATCTTTGC GGTGATTTTA	360
25	TTTGTTAGCG CTGTTCAAGT AGTACTTCAT TATCGTTTAA ACCAACAAGC AGCATTGCAA	420
	CGATTATCAC AATATCATTT AAGCCGTTTC AACTATATA GTACTTATGT AATGACACAT	480
	ACGATTTTGT TATTGTTGGT ACTATTGGCA GTTAGTCTAT ATTTGTCTCA ACCACTCAGC	540
30	TTAATATTTT ACTTAAAATC ACTGTTACTT ATATTGATTT ATGAGATAGG TATCGTCTTT	600
	ATCTTATTCC ATATTCAAAC AATAAGTCAT CGATTATTCA TGACATTTAT ATATGCACTT	660
35	GCTATGGGTA TCGTATACTT GATTATTTTC ATGTAAAGGA GCGTAACTGA TGATAGAAAT	720
	TAATAACCTT TCAAAGCGTT ACCGTAACAA ACAGATTTTC AATCATTTAA CTATGTCCTT	780
	TGATAGTAAT CGTTTAACCG TATTACTTGG TGATAATGGT GCTGGAAAAT CAACATTACT	840
40	TCGTATGATT GCTGGTATTG AAAAAGCTAA TGATGGAACT ATCAACTATT TCGGCGAAAA	900
	ATGGAATCAA AGACAaATAC aAAATCACAT CGGTTATGTG CCACAAGACA TTGCGTTATT	960
	TGAACACATG ACAGTGGCTG AAAACATTAA ATTTTAA TCACTTTGTA AAAATCCAAT	1020
45	TAACGATACA ACTATCAACG AATATTTACA GCAATTAAAC TTTGATGATA CGTCTGCCAA	1080
	AGTATCTACA TTGTCCGGTG GGAATAAACG TAAAATTAAT ATATTAGTAG GTTTACTAGG	1140
	TCAACCTCGA ATTCTCATTT TAGATGAACC GACAGTTGGT ATTGATTTAA AATCTAGACA	1200
50	TGACATCCAC CAACTACTTA ACATCATGAA ATCTAAATGT TTAATTATAT TAACTACCCA	1260
	TCATTTAGAT GAAGTTGAAG CACTTGCAGA TGATATCAAG TTAATTGGCC AAGATCCTTT	1320

	CCCAAGCTGC GTATGATATC GCAACTTGGG ATTTTCTGTA TTATCTACTT TGCAAGTATG	1440
	ACGTTGGGTC TACTGCATAT TGATTACCGA TGCCACCAGA CATACTTGG AAGTGACGT	1500
5	GAGGCGCTGT TGAATTACCC GTACTACCTG AATATGCAAT TTGGTCACCA GCTTTGACTT	1560
	TATCACCAGC TGAAACAGTT AAACGATTAT TATGCATATA CCATTGGTAG TTATTACTGT	1620
10	TCGCTTCTTT AATCGTTACT TGATTGCCGC CACCATAGTT ACTCCAACCT GCTTGTAATA	1680
	CTGTACCATC AGTTAATGAG TAACTGGTG AATTTTCAGG CATTGCATAG TCGACACCGT	1740
	AATGCGCACC ACCACCGTGA TATTGTCCAT ATGGTTGTAG TTGTTTACGA CTTGTTAACC	1800
15	AGCTTGCGTC TTTCGCATGA CCACTAGcTG TCGCTTTACT TGCTGATCCA CCATTTTGAT	1860
	TAGATGTGCC ATTAGGATAA TTGACCTTTC CATTACCATC ATGGCTGTTA TACGCTTGGT	1920
	TGTTGTTACT ATGTGAATAA TAGCTCGCGT CTGGACCTAC ATTTGATTGA TAACCATATT	1980
20	GATTAATATG CTGTTGGCTT TGA CTGCTG TGTAGTCATT GTTATCTCCT GCTGTTGCTG	2040
	GATTCACATA TGTTTGGCCG CTTCCATTGG CATTTCGATT TTTTGGATAA CAGTTATAAA	2100
	AATAATGCGT ATGTCCTTGA GCATCTACGA ATGTATAGCT ATATTCTTTA TTATCAAACA	2160
25	TTGCTTGATT CCAGTTACCA TCAGGTGTGT GATGATAATC CCCATTAGAA TCAATTGTAT	2220
	AATAAGTACC ATAAGATACG TCTTGTGATT GTGTTGACAT TTGTGTATGT GCTTGTGGG	2280
30	TGTTTGTGCT TTCTGCTGCA TCTGCTTGAT GCGCCATTGT AAATGTAGCG AAsCCATCGT	2340
	TGCAATCGCT GCTGCTGTTA ATTTTTTCAT GTATAAAACA TCCTCCATTA AAGTTAAAGT	2400
	TAGTTTTCAA TTAACTGTA CTGCACATAC TAAAAGAATT AGACAACTGA GTAAAGGATT	2460
35	TAATTCTCAT TTTCCAACTA TTTAATATTC CCGAAATGTT TTACTAACT CATTACATTG	2520
	TCATTACAAA ATAGCCATAC ATTGATATTA AAATGACATC TCTCACTGCA TTCGTTTAAC	2580
	CTTTTATAAA TTTTCAAAAT TAACAACTAA TCGTTCGTCC ATGTTGCGGA TTCAACGCTA	2640
40	ATGCATGATA GTAATCATCC ATCAAATCAT ATCAACCAA TTCCATTATC AATCGCTATT	2700
	GATTGTCATT CAACTTTCTA ATAGTGATAT GCTTCTCAGG CTTAAAAATC GTCATATCTT	2760
	TTCTATTAAT TAAATCATCT GTGAGCTTTA ATGCTACTAA TTCATTGCTG CCATAATACT	2820
45	TAATATATAA CGTTCTTGTA GTTAAATTTA TTACGGTCTG ATACATCGTA TAGTGATTGT	2880
	CAGCATCATG CGGACGTACA ATTCCAATCG GTATATTTAC CGCATCTAAT AAATAAAATG	2940
50	CATTCATTAA ATCCATTCTT TTATCATTGT TTTGAGCAAT GTTTGCTTTC ATAAATGCCA	3000
	TTCTCACAAA GCGCTCAGTT GAAGTAAATC CACCTGGCAA TCCAAATGTA CCTGCTTCAT	3060
	TGCCTAAAGG TTCAATCGTT ACACCTTCCA ATAAATTTGC TGTGCTGGA TAAGGAGAAA	3120

55

	CACCAATAGG ATTATCTTTT ATAACCACTT CACCCTCTTT AAATGAAACT TCGACTGTAT	3240
	GTCCAGTTGC ATCGGAAACA TGATAATGCA ATGGCGGAAC TTCACCGATG TCATTTAAAT	3300
5	ATACAGCTAC AACATGTATT TGGGATGCTT GTTGTTTCAT ATCTTCAATG CTTGTTGTAT	3360
	ATCCCAAAAT CCATGTCACA ATTTCATTTT GCGTAATATT CATCGCGTCC GCTTTGTGTG	3420
	TTGATCCATA TGAACATAA CCTCGGAAAT ATTGTGTGA AATGGCAACG CcATGTtCAT	3480
10	TAACACCATC ACCATAAATA AAACCTTCCA TATCTGTTCC TGTGCCAATA AAGCCATATT	3540
	GCGTTTGGCC TGTGCTGCCA GTGCAAGATT TCCAACGATA ATTTCTAGGC GTCAGTCTG	3600
	GCGAACCATC TAATGGATAA TCATAATCCA TCGTGCGTCC AAGAAGTACT TGATTATTTA	3660
15	AAGTTTGTAT TGTGAATCCT GTGCACATTG TTCTCACTCC TCTGTACCTT CATTTACTTT	3720
	AATCACTTTC AAATAAAGCT GTTTCACCTA AACATACTAT AAAAAATCAA TTATAACAAGC	3780
20	AATTAATTGA TATTCATTCT CAATAACTGT GGTATGATAT GTAAGGAAAT CATGACTTAT	3840
	GTGTGAGTGA ACGATCATCT ATACATCCGT TCACTTCATC TCATGACTTT CTATATTTAA	3900
	TTTTTACAAG GAGTGACATC TGTGAATAAC ACACAATCTT CACCACGCAG TAATATTATT	3960
25	ATTGCGATTA TGTTGTCTGC ATTAACATAT TGGTTGTTTG CACAATCATT TATTAATATA	4020
	GGACCTCTCG TTGGTCAAAC ATATCAAACC TCTCCTGCCG TGTAAATTT ATCTATTAGT	4080
	TTAACTTCCT TCGCCACAGG TATCTTCATG GTGGCTGCAG GTGATATTGC TGATAAATA	4140
30	GGACAACTGA GAATGACATA CATGGGTCTC ATAATCAGTA TGTTTGATC TCTTCTATTA	4200
	ATTATATCGG ACATCACTGC ACTGCTCATC ATCGGTAGAA TTTTACAAGG TCTATCAGCA	4260
	GCTATCTTGT TACCTTCAAC AGTTGGCGTG TTAAATAATC AATTTAAAGG AGAACATTTA	4320
35	AGACGAGCGA TTAGTTATCT AATGATTAGT ACTGTTGGTG GCATCGGCCT AGCTGGTGT	4380
	ATCGGCGGTT TAATTGCCTC AAATTTGCGA TGGCAAACGA ATTTTCATCAT TAGTATAGTC	4440
40	ATTGCTTTCA TTGCCATATT GCTTCTAAAA GGCACACCTG AAAAAGTAAG TCAACATAGC	4500
	CACCGTCATC CATTCGATTA CAAAGGTATG TCGATTTTCG CTGTTATGAT TGGTAGCTTT	4560
	ACATTATTGT TAACACAAGG ATTGGAACAA GGTTGGTTTA GTACATTTTC AATCATTTGT	4620
45	CTGAGCATTT TTATCaTCac TACGTTGATA TTCATCATCA TCGAACGTCG ACATGAAGTA	4680
	CCTTTTATTG ATTTCTCAGT ATTACGCAAC CGTCCGTTCA TTGGTGCAAT TTAAATAAC	4740
	TTTGTTTAA ATAGCGGTCT AGGCGTAACA GTGGTCTTTT TCATATATGC TCAAACACAC	4800
50	CTTGGTTTAc CAgcTGCGCA ATCTGGACTT GTTACATTGC CATATGCCAT TGTGGCAGTT	4860
	GCGATGATTC GTTTAGGTGA AAAAGCAACA TTACGTTTCG GTGGCAAATT GATGCTCATC	4920

55

	TCACAATATG TCATTGCAGT TATCATTGGT TTCGTCATAT GTGCGATAGG TAATGGTTTA	5040
	GTGCGAACAC CTGGACTTAC GATTGCAATT TTCAGTATGC CTAATGAAAA AGTTGGTTTA	5100
5	GCTACAGGAT TATATAAAAT GAGTGGTACA TTAGGTGGCT CCTTTGGTAT AGCACTAAGT	5160
	ACTACAGTTT TCAGTATGTT ACAACTAAAC TATGCACCAA GTGTAGCTGC AACCGTAACA	5220
	TTTATAGTCA GCATTGTATT GATGATCCTT GGCTCATTGT CTGCATACAT GATCATTCCA	5280
10	AAAACAGTTA AATCTTAAAT ATAATAGAAG AATTATGTTT CGAAATATCT TTATCACTTT	5340
	AAAATGATAT ACAAGAAATC CAAGAAAAAT AAGCGAACTG AATAAATAAA GATTCAATTA	5400
15	ACGCATCAGT ATTAGGATTC ACTCTAAAAC GATTAATAGT TTTATAAGAA GGTGTTTGAT	5460
	CTTGAGCTAA CCACATCATT CGAATACTGT CATGAAGTAA TTTCTCTATT CTACGACCAG	5520
	AAAATACAGA TTGAGTATAT GCATATAAGA TGATTTTTAA CATCATCTTT GGATGATAGG	5580
20	ATGTTGCGcC ACGATGATGT CTGAATTCAT CGAATTTGCT ATCAGGTATC GTTTCACAA	5640
	TTTCATTAAAC ATGTCGCGAA ATATCATTTT GAGGAATTCT AACAGAAGTT TTTATTGGTA	5700
	GTGTAAGTTG GGCAAAGTGT CTTATTTTTT TAAAGTATTT CAAAGTAAAA TTACATGTTA	5760
25	ATACGTAGTA TTAATGGCGA GACTCCTGAG GGAGCAGTGC CAGTCGAAGA CCGAGGCTGA	5820
	GACGGCAccc TAGGAAAGCG AAGCCATTCA ATACGAAGTA TTGTATAAAT AGAGAACAGC	5880
30	AGTAAGATAT TTTCTAATTG AAAATTATCT TACTGCTGTT TTTTAGGGAT TTATGTCCCA	5940
	GCCTCTTACT CTAATTATAT TCACTATCAA TTAGACAAAA TGGCCATTTT CAAATATCAC	6000
	GCGTTGTTTC TGACCTTGAA TATATTTATT ATAATTCTCT TTTTGAAAAT CAGTTAACAT	6060
35	TAATTTAGAT GTACCGTATT TTAACACTTT TTGCATTGTT TCTATTCTCA TTTTCTAAA	6120
	TAACCATCCA TCTTTTAAAC CAATACGATT AACAGCATCA TATGATAATT CTACTGTTTC	6180
	TTTAATTTCA AATGTCTTGA ATGAAATAAT CGTGCACATT AAAACGTAT CACCAAAGTA	6240
40	ATAACATCT AAATCATCAC GTTTATGTTG TCCAACTAAC AAACGACCAT ATTCGAACTC	6300
	TTTTCTGGA TATTTCAATT CTAAAAAACT AATAATCTCT TCTTCTTTTA ATTTGAATTG	6360
	CATTTAAAAA CATCCTCTCT TAAGTTTTAA CAAGCCTTAA TTAAAAATTT TTTCAATCAC	6420
45	ATAGTTCAAT ATACATCATT TCGTTATGTT TTTTAATACT TTGTTCAAAA ACAAATATTT	6480
	TATCTTTTAA AATAATGACT TTTGTATTTT TAATCACAAT AAACATTTTA AAATTCTTGT	6540
50	TATCATAATC ATTAAAAGGT ATTAACCTTA ATAATATATT CTCTCGTCTC AACCTTAATC	6600
	GTATACTTCA GACGTCTGTT TGTAGACAAT AAAAGTCATT CACGTCTTCA TATGTCATCA	6660
	AATGTTTATC ATGATATGAT GAATATAATA ATCGGGTATA TAACTGTATG ATTAATTACA	6720

55

	TGGTGTATCA GTTACAACGTG TGTCACATAT TTTAAATCAT AATGATAGTC GTTTTTCCGC	6840
	AACAACGATA AAAAACGTAC ATGCTGTTTT AGAACGTTTA GGCTATGCCC CTAATAAACA	6900
5	TGCAAAACAA TTGCGCGGCA GTAAAATTCA AACTATTGGC GTCATTTTGC CTAGCTTAAC	6960
	AAATCCGTTT TTCTCAGCAC TGATGCAAAG TATTCATGAC CATAAACCAT CTGATGTTGA	7020
	TTTATGCTTT TTAACATCTA CAGCAACTGA TTTGTATGAC AATATTAAAC ATTTAATTGA	7080
10	TCGAGGTATT GACGGATTAA TTATCGCACA ATACATATCA TCCCCGGACG CCCTAAATAA	7140
	CTATCTAAAG AAACATCATG TACCTTATGT CGTACTGGAT CAAAATGACC ATCAAGGCTA	7200
	TACAGATTTT GTTCGGACAA ATGAATATCA AGGTGGACAA CTTGCAGCAC AACATTTAAT	7260
15	AGAACTCGGT CACAACCATA TGATAATTGT TGCACCATAT GACATGATGG CGAATATGTC	7320
	GACTCGTGTC GCTGGATTG TCGATACTTT GCGCGCGAAT CAATTGCCAG AACCACAAAT	7380
20	CGTCCATACT GAATTATCTA AGCGCGGTGG GCTAACCATT GTTGATGACA TCATGGTTCA	7440
	ATCTGCCACT GCAATCTTCG CTATTAACGA TGAACCTCGT ATTGGCATT TACGAGGACT	7500
	AATTGAACAT GGCATCAGTA TCCCGAAAGA TATCTCATT ATAGGTTATG ACGACATTGA	7560
25	TTATGCAGCG TACGTCTCGC CACCTTTAAC TACTGTGGCA CAACCTATAA CTGATATTGG	7620
	CAAAACATCT TTAACCTTAT TACTTCAACG ATTACAGCAC TTAGATAAAT CCATTGATAT	7680
	GATTGAATTA CCAACGACTT TAAAAATTCG TGCAACAAC TGGCTATCATC TTTCAAAC TA	7740
30	ACTACGTATC TTCCGAAATA TACTCATCAT TGTTAGGCCC TTAGCGTTGC TTTAATGCTG	7800
	AGGGTTTTTTA ATCATAATTA TTTTACTAAG AAATTAAAT AATAATGTAT GAATTTTTAA	7860
	ATATGATTTA AACGTTTTCA GTTTTTATGA AAACGCATGC ATTTTACAAA TAAAAATGGT	7920
35	ACGATGGCAC TGGTAAAACG TTTTACTAAA AACAAATCAT GAGGTGTATA ACATGAGCAT	7980
	TGTTGCATTA CTTATCGGGT TAGGCCCCCTT AATTGGCTGG GGCTTCTTCC CAACAGTCGC	8040
40	TTCAAAGTTT GGTGGTAAAC CTGTACATCA AATTATCGGT GCTACTGTAG GTACGTTAAT	8100
	CTTCGCTATT TTATTAGCCG TAGTCACATC AAGTGGCTTC CCTACTGGAA CCAATTTGCT	8160
	ATTCGCCTTA TTATCAGGTG CAGGATGGGG ATTCGGACAA ATCATTACAT TTAAAGCGTT	8220
45	CGAATTAGTC GGCTCATCTC GTGCCATGCC AGTCACAACA GCATTCCAAT TATTAGGCGC	8280
	ATCTTTATGG GGTGTCTTTG CATTAGGAAA TTGGCCAGGC ATTGGTCATA AAATCATTGG	8340
	ATTTACAGCT TTAGTCGTTA TTCTAATTGG AGCGCGTATG ACAGTTTGGA GTGAACGCAA	8400
50	AGAAGCAAGT AACGCCAAAA ATTTACGTCG TGCAGTGGTA CTTCTGTTAA TTGGTGAATT	8460
	TGGATACTGG TTATATTCAG CTGCACCGCA AGCAACTTCT ATTGATGGCC TAACTGCCTT	8520
55		

	AGCAGAGAAT CCATTCCGTA ATAAAATTAC GTGGTTACAA ATTATTTTCAG GTTCTTCTT	8640
	TGCATTTGGT GCTTTAACAT ATCTTATTTT AGCACAACT AATATGAATG GTTTAGCAAC	8700
5	TGGATTTATT CTTTCTCAAA CATCCGTTGT GCTTGCTACA TTAAGTGGTA TTTATTTCTT	8760
	AAAACAACAT AAAACGTCAA AAGAAATGGT TATTACAATC ATCGGCTTAG TACTCATTTT	8820
10	AGTAGCCGCT TCTGTTACAG TATTTATAAA ATAAGGAGTG TAGATGTCAT GAAAAAATCA	8880
	GCTGTTTTAA ATGAACATAT TTCAAAAGCA ATCGCGACAA TTGGTCATTT TGATTTATTA	8940
	ACGATTAATG ACGCTGGCAT GCCAATTCCA AATGATCATC GTCGTATCGA CCTAGCTGTA	9000
15	ACTAAAACT TACCACGCTT TATTGATGTC TTAGCTACAG TGTTAGAAGA AATGGAAATC	9060
	CAAAAAATAT ACTTAGCAGA AGAAATAAAA GAACATAACC CTACACAATT GCAACAAATT	9120
	AAACAATTGA TTTTCATCGA AATCGAAATC ATTTTCATTC CTCACGAAGA AATGAAAAGT	9180
20	AACCTAGCTC ACCCATTAAT TAAAGGTAAT ATTCGTACTG GTGAAaCAAC GCCCTACTCT	9240
	AATATTGCAT TAGAATCGAA TGTTACTTTT TAAAAGTTAT AACTTGAAAG GAGCGTACAC	9300
	ATGACCAACA AAGTTGTTAT TTTAGGTTCA ACGAATGTCG ATCAATTTTT AACAGTTGAA	9360
25	AGATATGCAC AACCAGGCGA AACATTACAT GTTGAAGAAG CACAAAAGC ATTCGGCGGA	9420
	GGTAAAGGTG CCAACCAGGC TATTGCCACT GCACGCATGC AAGCAGACAC AACATTTATT	9480
30	ACTAAAATTG GCACTGATGG CGTTGCTGAT TTCATCTTAG AAGATTTTAA AGTAGCTCAT	9540
	ATTGATACAT CATATATTAT CAAAACAGCT GAAGCAAAA CGGGCCAAGC CTTTATCACT	9600
	GTGAATGCAG AAGGACAAAA CACCATCTAT GTTTATGGTG GTGCGAATAT GACGATGACA	9660
35	CCTGAAGATG TTATTAACGC AAAAGACGCT ATAATCAATG CAGACTTTGT CGcTGCACAA	9720
	TTAGAAGTAC CCATCCCGGC TATTATATCT GCATTTGAAA TTGCCAAGGC ACATGGTGTG	9780
	ACGACAGTAT TAAATCCTGC ACCAGCGAAA GCATTACCTA ATGAATTATT ATCATTAAATC	9840
40	GATATTATTG TGCCAAACGA AACAGAAGCC GAATTGTTAT CTGGGATTAA AGTAACTAAT	9900
	GAACAATCTA TGAAAGACAA TGCCAATTAC TTTTATCTA TAGGCATTAA GACTGTTTTG	9960
45	ATTACGCTAG GTAAGCAAGG TACATATTTT GCTACTAAAA ATCAAAGCCA ACACATCGAA	10020
	GCTTATAAAG TAAATGCGAT TGATACAACT GCTGCAGGCG ACACATTTAT TGGTGCAATT	10080
	GTCAGTCGCT TAAACAAGTC GCAAGATAAC TTAGCAGATG CTATTGATTT TGGTAATAAA	10140
50	GCGAGCTCAC TCACTGTACA AAAACACGGC GCGCAAGCAT CTATTCCTCT ACTAGAAGAA	10200
	GTAAATCAAG TTAAATGAA TCAAACACAG CTATGATATG AAGGTTTAGC ATATAACATG	10260
55	CAACATTCGT ATATCATGGC TGTGCTTTTT TATCTTTATA AAACATCATC TATTAGAAAT	10320

TTTGTAATCT TTTTAACTTC CAAATTaTCG CATATAAATA TGCTATATTA ATGATAATAA 10440
 TTATCAATTA AAAGGAGGTT ATGCTATGTC TAAAGAAGCT GGTACATACAT TTTTAGCTAA 10500
 5 ATTAGGAAAA ACTCGTCTAC GCCCCGGTGG TAAAGAAGCA ACAGATTGGT TAATACAACA 10560
 AGGGGCATTT TCACAAGATA AACAAGTGTT AGAAGTGGCA TGTAATATGT GCACAACATC 10620
 TATTTATCTA GCTCATAACAT ATGGCTGTCA CATTCAAGGC GTTGATATAA ATAAGAAAGC 10680
 10 ATTAGAAAAA GCACAGGAAA ACATTTTCAGC AGCAGGTCTT GAATCATATA TTCAAGTTCA 10740
 ACAAGCGAAT GCTGTAAAT TGCCCTTTGA TGACAATCAA TTCGATATCG TTTTAAATGA 10800
 AGCAATGTTA ACAATGTTAC CCATCGCCAT AAAGGAAAAA GCATTACGCG ACTACTACCG 10860
 15 AGTCTTAAAG CCTGGGGGTA TCTTGTTAAC ACATGATATT GTCATCGTTA ATGAATCACA 10920
 TGCCACACAT GTTGTTAAAT CATTATCTGC AGCAATTAAT GTCAATGTCT CACCGCAGAC 10980
 20 GAAACTTGGC TGGTTAGATT TATATAATCA AGCTGGTTTT AATCATGTGC ATTATCATAC 11040
 TGGTCCAATG AGTTTAAATGA CACCAAAAGG TTTAATTTAT GACGAAGGTA TTGTTGGAAC 11100
 TATAAAGATT ATCAACAATG CTTTGAAAAA AGAAAATCGA CCAATGTTTT GTAAAATGTT 11160
 25 TAAAACGATG ACTAAATTGC GTAAAGATAT GAATTATATT ACTTTTGTCTG CTAAAAAAGA 11220
 GCACTAAATA TAATGCCACT AACTGTACTT TGTATCTATG TTTGACTATC ACTTTAATTT 11280
 CTTTGTGACA CTAATCATCT ACTTAACAAT ATCGTTATCG TTGATTAGTA AGTCATCAAT 11340
 30 TTTGGTTAAA GACTTTCATA AACACTCAAA CATTAACTAT ATACATAGTT AGTGGcATTA 11400
 TTTTTTyCTn AAAATTTTAA CmTCmCGGGr TtGGGAmCrG AAaTGrtAwT TcGCrmAAcT 11460
 35 TAwTcT 11466

(2) INFORMATION FOR SEQ ID NO: 445:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2176 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 445:

TTACATAGTT AACACTAGTT AATCTATTAG TTAACATTAG TTAATAATTA GTTAATTTCC 60
 ATTTGTATTTC TCATGTGATA AATTCTAAAA GCATACaATA AATTTAATAT GTAAAAAGAA 120
 50 AGGGAATACA CATGAAAAAT AAATATATCT CGAAGTTGCT AGTTGGGGCA GCAACAATTA 180
 CGTTAGCTAC AATGATTTCA AATGGGGAAG CAAAAGCGAG TGAAAACACG CAACAACTT 240

	AAGTATTACA TCTAAAAGGT ATCAGAGAAG AACAAACGTAA CCAATACATC AAAACATTAC	360
	GCGAACACCC AGAACGTGCA CAAGAAGTAT TCTCTGAATC ACTTAAAGAC AGCAAGAnCC	420
5	CAGACCGACG TGTTCACAA CAAAACGCTT TTTACAATGT TCTTAAAAAT GATAACTTAA	480
	CTGAACAAGA AAAAAATAAT TACATTGCAC AAATTAAAGA AAACCCTGAT AGaAGCCAAC	540
10	AAGTTTGGGT AGAATCAGTA CAATCTTCTA AAGCTAAAGA ACGTCAAAAT ATTGAAAATG	600
	CGGATAAAGC AATTAAAGAT TTCCAAGATA ACAAAGCACC ACACGATAAA TCAGCAGCAT	660
	ATGAAGCTAA CTCAAAATTA CnTAAAGATT TACGTGATAA AAACAACCGC TTTGTAGAAA	720
15	AAGTTTCAAT TGAAAArGCA ATCGTTTCGTC ATGATGAGCG TGTGAAATCA GCAAATGATG	780
	CAATCTCAAA ATTAAATGAA AAAGATTCAA TTGAAAACAG ACGTTTAGCA CAACGTGAAG	840
	TTAACAAAGC ACCTATGGAT GTAAAAGAGC ATTTACAGAA ACAATTAGAC GCATTAGTTG	900
20	CTCAAAAAGA TGCTGAAAAG AAAGTGGCGC CAAAAGTTGA GGCTCCTCAA ATTCAATCAC	960
	CACAAATTGA AAAACCTAAA GTAGAATCAC CAAAAGTTGA AGTCCCTCAA ATTCAATCAC	1020
	CAAAAGTTGA GGTTCCTCAA TCTAAATTAT TAGGTTACTA CCAATCATTAA AAAGATTTCAT	1080
25	TTAACTATGG TTACAAGTAT TTAACAGATA CTTATAAAAG CTATAAAGaa AAATATGATA	1140
	CAGCAAAGTA CTAATAATAT ACGTACTATA AATACCAAGG TGCGATTGAT CAAACAGTAT	1200
30	TAACAGTACT AGGTAGTGGT TCTAAATCTT ACATCCAACC ATTGAAAGTT GATGATAAAA	1260
	ACGGCTACTT AGCTAAATCA TATGCACAAG TAAGAAACTA TGTAAGTGA TCAATCAATA	1320
	CTGGTAAAGT ATTATATACT TTCTACCAAA ACCCAACATT AGTAAAAACA GCTATTAAAG	1380
35	CTCAAGAAAC TGCATCATCA ATCAAAAATA CATTAAAGTAA TTTATTATCA TTCTGGAAAT	1440
	AATCAATCAA AAATATCTTC TCTAGTTTTA CATCATTTTT TAAATAATTT TCGTAACAAA	1500
	CCGTGATTAA AAAGAACCGT TGATTCTCAA TCGAATCTAC GGTTCTTTTT TCATTTTCCA	1560
40	TCAATTAAAT GCTTCTTCGC TATTTGTCAG CCCACTTTTT TACCTGCAAC TTGTTAAATA	1620
	ATCCTTACAT CGTTAACGAA TAGTTCATCA TTTAGTTGAA TCAGCTCAAC TTTATTAACT	1680
	TCATATTTTC ACAAACCTATT GCGCAATCCA TTCCTTTTCC ACTACAAGCA CCATAATTAA	1740
45	ACAACAATTC AATAAAATAA GACTTGCAAA GCATAGTTAT GTAGCTATAT AAACGCCTGC	1800
	GACCAATAAA TCTTTTAAAC ATAACATAAT GCAAAAACAT CATTTAACAA TGCTAAAAAT	1860
50	GTCTCTTCAA TACATGTTGA TAGTAATTAA CTTTTAACGA ACAGTTAATT CGAAAACGCT	1920
	TACAAATGGA TTATTATATA TATGAACTTA AAATTAAATA GAAAGAAAGT GATTTCTATG	1980
55	ATTAAAAAtA AAATATTAAC AGCAACTTTA GCAGTTGGTT TAATAGCCCC TTTAGCCAAT	2040

CmTyCAAArG AcACaGACAT TACTAGCCAA CGATTtTAGCT ATnACTCCAA ACCTTCCATT 2160
GGATTtGGTA AAGGnT 2176

(2) INFORMATION FOR SEQ ID NO: 446:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1557 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 446:

AAAAGCATGG CTTAAATGAA GTACGCTATA ACAAATTACA AGAACATGCT ATTGTTATGC 60
ATCCGGCACC TGTGAATAGA GGAGTAGAAA TACAAAGCGA TTTAGTAGAA GCTTCAAAAT 120
CAAGAATTTT TAAGCAAATG GAAAATGGCG TTTACTTAAG AATGGCAGTC ATTGATGAAT 180
TATTAAATAA GGTAAGGGGA CGAAAATGAT GAAATTAATT AAAAACGGTA AAGTATTACA 240
AAATGGCGAA TTACAACAAG CAGATATTTT AATTGATGGT AAGGTAATTA AACAAATTGC 300
ACCTGCAATT GAACCAAGCA ATGGTGTGTA CATCATAGAT GCGAAAGGTC ACTTTGTGTC 360
ACCTGGATTT GTCGATGTTT ATGTTCAATT ACGTGAACCT GGTGGTGAAT ATAAAGAGAC 420
AATTGAAACT GGTACTAAAG CTGCTGCTAG AGGCGGATTT ACAACTGTAT GTCCAATGCC 480
TAACACAAGA CCGGTACCAG ATTCTGTAGA ACATTTTGAA GCTTTACAAA AATTAATCGA 540
TGACAATGCT CAAGTACGTG TATTACCTTA TGCTTCAATT ACAACACGTC AATTAGGTAA 600
AGAATTGGTT GATTTCCAG CACTAGTAAA AGAAGGTGCC TTGCGTTTA CAGATGACGG 660
TGTAGGAGTA CAACTGCAA GCATGATGTA TGAAGGCATG ATTGAAGCTG CAAAAGTAAA 720
CAAAGCCATC GTAGCACACT GTGAAGATAA TTCATTAATC TATGGTGGTG CAATGCATGA 780
AGGGAAACGC AGTAAAGAGT TAGGTATACC AGGTATTCCA AACATTTGTG AATCTGTTCA 840
AATCGCAAGA GATGTACTAT TaGCTGAAGC AGCAGGTTGT CATTATCATG TATGTCATGT 900
TTCTACTAAA GAAAGTGTTA GAGTCATTCTG TGACGCTAAA CGCGCAGGCA TTCATGTTAC 960
AGCTGAAGTT ACACCACACC ATTTATTGTT AACAGAAGAT GATATTCCTG GTAATAATGC 1020
CATTTATAAA ATGAATCCAC CATTGAGAAG TACTGAAGAT AGAGAGGCTT TGTTAGAAGG 1080
GTTACTAGAC GGTACAATTG ACTGTATCGC AACAGAcCAT GCACCACATG CACGTGATGA 1140
AAAAGCACAA CCAATGGAAA AAGCaCCATT CGGAATTGTT GGTAGTGAAA CAGCATTCCC 1200
ATTATTATAT ACGCATTTTG TAAAAAATGG TGATTGGACA TTACAACAAT TAGTAGATTA 1260

TTATGCAGAT TTAACAATCA TTGATTTAGA TAGTGAACAA GAAATTAAAG GAGAAGATTT 1380
 CTTATCAAAA GCAGATAATA CACCATTTAT CGGCTATAAA GTTTATGGAA ATCCGATCTT 1440
 5 AACAAATGGTT GAAGGCGAAG TTAAATTTGa GGGGGATAAa TATATGCAA gCAAACGTTA 1500
 TCTAGTGTTA GAAGACGGTC TTTTACGAGG CTACCGTTAG GTCTGATAAC TTAAGTGA 1557

(2) INFORMATION FOR SEQ ID NO: 447:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1799 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 447:

20 GCTAGAAATm TTGmATGaCA ATACAACCTCT GTTAAAaTGA TGGACGTAGA CAAATATGCG 60
 TATTGACGCT TTATTTTAAA AATTaACATG CTTATAACAT GTTTATAGAA GGAGATTAAC 120
 CTATGAAC TAaAGTTCTT TTATATTATA AATATATGAC GATTGATGAC CtGAACAGTT 180
 25 TGCTCAGGAT CACTTAGCCT TTTGTAAAGC ACACCATTTA AAAGGTAGAA TTCTTGTTTC 240
 TACAGAAGGT ATTAACGGCA CATTATCTGG TACAAAAGAA GAAACCGAAC AATATATGGC 300
 ACATATGCAT GCCGATGAAC GATTCAAAGA TATGGTGTTT AAAATTGATG AAGCTGAAGG 360
 30 ACATGCTTTT AAGAAAATGC ATGTACGTCC TCGAAAAGAA ATCGTTGCTT TAGATTTAGA 420
 AGATGACGTC GATCCAAGAC ACACAACCTGG CCAATATTTA TCACCTGTAG AATTTAGAAA 480
 AGCTCTTGAA GATGATGACA CAGTCATTAT TGATGCACGT AATGATTATG AATTTGATTT 540
 35 AGGTCAATTC CGAGGTGCAA TTCGTCCAAA TATCACACGT TTTAGAGATT TGCCTGACTG 600
 GATTAAAGAG AATAAAGCGT TATTTGCAGA TAAAAAAGTG GTTACGTACT GTACTGGTGG 660
 CATTCGATGC GAAAAATTTT CTGGATGGCT TTTAAAAGAA GGTTCGAAG ATGTAGCTCA 720
 40 ACTTCATGGC GGTATTGCTA CATATGGTAA AGATCCTGAA ACAAAGGTG AATATTGGGA 780
 CGGTAAAATG TACGTATTTG ATGACCGTAT CAGTGTTGAT ATCAACCAAG TTGAAAaAC 840
 AATTATTGGT AAGGATTGGT TTGATGGCAA ACCATGTGAA CGTTATATTA ATTGCGCTAA 900
 45 CCCAGAATGT AATAACAAA TATTAGTTTC TGAAGAAAAC GAACTAAAT ATTTAGGTGC 960
 ATGCTCTTAT GAATGTGCTA AACATGAGCG TAATCGTTAT GTTCAAGCAA ATAATATTAG 1020
 50 TGATAATGAG TGGCAACAAC GTTTAACAAA CTTTGATGAT TTACATCAAC ATGCTTAGTT 1080
 TTAATTAAAT ACCTTTCAA ACACGCTTTG AAAATCCGAT TTATAAAGGT TTTTCAAGGC 1140

TAAATTTTAA TACTGCGGGG TGTCTTAAAA TGCACATTTT AGTAACAGGG TTTGCGCCTT 1260
 TTGACAATCA AAATATCAAT CCCTCATGGG AAGCTGTGAC TCAACTAGAA GATATTATTG 1320
 5 GCACACATAC AATCGATAAA TTAAACTAC CAACCTCTTT TAAGAAAGTA GATAATATTA 1380
 TAAATAAAAC GTTGGCATCT AATCATTATG ATGTTGTACT AGCTATAGGA CAAGCTGGTG 1440
 GTAGAAATGC CATTACCCCA GAACGTGTCT CCATTAAATAT TGATGATGCA CGTATTCCAG 1500
 10 ATAATGATGA TTTTCAACCT ATTGATCAAG CCATTCACCT AGACGGTGCG CCAGCTTATT 1560
 TTTCAAATTT ACCaGTTAAA GCAATGACTC AAAGTATTAT TAATCAAGGA CTTCTGGAG 1620
 CACTTTCAA TAGCGCAGGT ACATTTGTTT GTAATCACAC ACTTTATCAC TTAGGTTATT 1680
 15 TACAAGATAA GCATTACCCT CACCTACGAT TCGGATTTAT TCaTGTGCCA TACATACCAG 1740
 AGCAGGTcAT TGGTAAACCC GATACACCAT CTCATGnCCA TTGAGGAAAA GATnAGTTG 1799

20 (2) INFORMATION FOR SEQ ID NO: 448:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1341 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 448:

30 ACTTGGTTTT TTATTGTTTA TAAATAAAAC TCACTTAATA ATGTTTTTCAT AATCTTCTTC 60
 GACTACTTAA TTCTTTAAGA TATTCGTGAA AAGAGACATT AACTAGTTA ATTTTCAAAC 120
 AATACAAAAA GCGTCTACCT CCTACATATA ATTGTAGCGG AGATAGACGC TTAATATTTA 180
 35 TTTAAAAATT ATTTTAAACC ACCGAATGTC ATAACATCAC GGGCAATCAT ACTTTCTTCA 240
 TCTGTTGGAA TAACGACAAC TTAACTGGT GAATGAGGAT AGTTAATAAA TCCTTCTTTA 300
 CCACGTAGTA AGTTTTTCATT TTCTTAGGA TCCCAGTAAA CACCCATAAA TTCTAAGCCT 360
 40 TCAAGAACTT TCGCACGAAT TTCTACTGAG TTTTCACCGA TACCTGCTGT AAATACGATA 420
 ACATCAACAC CATGCATTCT CGCAGCATAT GATCCAATAT ATTTGTGAAT TTTAGAAGCA 480
 AATACATCTA AAGCCATTG TGAACGTGCT TTACCTGATT CAGCTTCTTC TGATAAGTCA 540
 45 CGTAAATCAC TAGATGTACC TGATAATCCT AATAAACCTG ATTCTTTGTT TAAGATTTCC 600
 AATACTTGTT CAGCAGTTTT ACCTGTTTTT TCCATAATAA ATGGAATTAA AGCAGGGTCA 660
 50 ATATTACCAG AACGAGTACC CATTGTTACA CCAGCAAGTG GTGTGAAGCC aTTGATGTAT 720
 CAATAGATTT ACCGCCATCG ATAGCTGCAA TTGATGCTCC ATTACCAATG TGACATGAAA 780

55

TATGGCTTGT ACCATGGAAA CCATACTTAC GAATGCCATA ATCTTTATAA TAATGATATG 900
 GCAAGCTATA TAGATATGCT TTTTCAGGCA TTGTTTGATG GAATGCTGTA TCAAAAATTG 960
 5 CCACATGAGG GATATTTGGT AATAATTTAC GGAAAGCACG AATACCCATC AAGTTaGCTG 1020
 GGTGTGtaAG CGGTGCTAAT TcGCTTAATT CTTCAATTTT CTTTTCAACC TCATCAGTAA 1080
 TAGCTACTGA TTCAGGGAAT TTTTCACCAC CATGTACAAC ACGGTGACCT GTTCCATCGA 1140
 10 TATCGTTAAT ATCATTAAATA ATATTGTGCG CTTTAAAAGC ATCCAACATG ATATCAACTG 1200
 CCTCAACGTG ATCCTTGATA TCTTGTAAGT TTTTAACTTT TTCCCCGTG ACTTCAATTG 1260
 TAAAAATTGA ATCCTTCAAT CCGATTCTTT CTAATAAACC TTTTGTTACT AATTCCTCTT 1320
 15 CAGGCATTCT AATTAATTGA A 1341

(2) INFORMATION FOR SEQ ID NO: 449:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1529 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 449:

TTTTGAAGAT ACTACCGATG AAAATAGACA AAAGATTTTT CAATATTTAT CACCTGAaG 60
 30 AGTTGCAAAT TTCTTTGATC AATTAGATAT TGATGACGAT GAATATGAGT TGCTATTTGA 120
 TAAGATGAAT GCGACATACG CAAGTCACAT ATTAGAAGAA ATGTCATACG ACAATGCAGT 180
 AGATATTTTA AATGAGTTGA CTAAACCAAA AGTTGCTAGT CTTTAAACAT TGATGAATAA 240
 35 AGATGACGCG AATGAAATCA AAGCATTACT TCACTATGAT GAGGATACGG CCGGCGGTAT 300
 TATGACGACG GAgTATTTAT CACTTAAAGC GCATACGCCT GTTAAAGAAG CATTATTATT 360
 GGTCAAAGCG CAaGCACCAG ACGCAGAAAC AATATATGTT ATATTTGTTCG TTGATGATGA 420
 40 TGGTAAATTA GTAGGTGTTT TATCGCTAAG AGATTTAATT GTAGCTGAAA ATGATGCTTA 480
 TATTGAAGAT ATTATGAATG AACGTGTCAT TAGTGTGAAT GTAGCAGACG ACCAAGAAGA 540
 TGTGCTCAA GTTATGAGAG ACTATGATTT CATGGCTGTA CCTGTTATAG ATTACCAAGA 600
 45 ACATTTGCTT GGTATCATCA CGATTGATGA TATTTTAGAC GTTATGGATG AAGAGGCTAG 660
 TGAAGACTAC TCTCGTTTAC CCGGGGTATC AGATATCGAT TCGACTAATG ATTCAATCAT 720
 50 TAAAACAGCA TTA AACGTT TACCATGGTT GATTATTTTA ACATTTTATG GAATGATTAC 780
 TCGGACAATT TTAGGGAGAT TCGAAAAAAC ATTAGAAAAT GTAGCGCTAC TCGCAGCGTT 840

55

TCGTAACATT ACGACAGGGG AAATTAATGA GCAAAGTAAA TTTAGAATTG CATTAAGAGA 960
 AGCAGGAAGT GGTGTATTAT CGGGTGTGTG ATGTTCAACA ATATTATTTA CAATTATTGT 1020
 5 TCGAATATAT CATCAGCCAC TTTTAGCATT AATCGTTGCA GGAAGTTTAA CTTGTGCGAT 1080
 GACGGTGGGG ACGTTTGTAG GTTCGATGAT TCCATTATTG ATGAATAAAT TAAATATCGA 1140
 TCCAGCAGTG GCTAGTGGAC CATTTATTAC AACAAATTAAT GATATTATTA GTATGTTGAT 1200
 10 TTATTTTGGT TTAGCTACAT CATTTATGGC TTAATTAAAT TAAGGAGGAG TTATGGAGTT 1260
 TTTATCTTTA GTTATTGTTG TTTTAGCAGC GTTTTTAACT CCAATAATTG TCAATCGATT 1320
 AAATATTAAT TTCTTGCCAG TTGTTGTTGC AGAAATTTTG ATGGGGATTG TGATTGGAAA 1380
 15 TTCATTTCTA AATATAGTAG AAAGGGATTG AATTCTAAAT ATTTTATCAA CGTTAGGCTT 1440
 TATCTTTTTA ATGTTTTTAA GTGGTTTAGA AATTGATTTT AAAGCTTTTA AAAAAGATAA 1500
 20 ACGCGCACGT CAAGGACAAA ATGATGATG 1529

(2) INFORMATION FOR SEQ ID NO: 450:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1827 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 450:

TTCTGGAAAC CAAAGTATTG TCATCTTCTA CTAGTAGTAT nGGCATCCAT ATCACCCTAAT 60
 ATCATTTAGT ATATTTTATA TTTTCTCCTG ATTTTAATCG ACTTTGAAAA TCTTTAATCC 120
 35 GGCAGTCAAC TTCAAAGCCA TGAATCATCA TTTTTTGCAT TGGTGCTTGT ATAAAGTAAT 180
 AAATCGGCCA AAATAATCGA GGGATATAAT CGTATAGATG TATATAAAG ACTGCCGACT 240
 CTTTGATTAA TCTAAATTCT AACTTCCCTT GATTAAGTGT ATATTTTTTC ACTAAACTTC 300
 40 CACTCAATAA AATTAAAGTT ATTATTCCAT CAGCTGTTTC TTCTATTTTA AATACTGCGA 360
 GCGGTGCGAC CTTATTCTTA ATATATATCT TAAATTGATC ATGTGATTTT TCTGTTTTCA 420
 CAAAGTTCC TTTAGTGTA CCCATCCATG CAATAAAATG GTTTACAACG TTCTTTAATG 480
 45 TCCATCCCTT TGGTAAACT ACCTTCATCG TTGATCTAAC ATCATCATAC TTTGAACTT 540
 GTAATTCTAC ATTAATAAA GAACGTTTAA AACTAAATT TGTGTTTCT ACAGGTGTAC 600
 50 CATATGCACC TAGGCGTTCA ATTGTTTCAT TATCATAACG ACTCCAGGT ATGTAGATTA 660
 CTTTTTTTAC TTGATTGATA GCCGCTGCTC GACCAAAATT ATCTGCTGCG ATTAATGTTA 720

CAGCAATATC TATTTGATTC ATTGCTGCAA CAACCTGTTC GTAATGAAAT ATATCACACT 840
 GAATCCAAGT CATTTCAACA TCATCTGTTT KTTTATTGTC TGGATATTTT GATATAGCAA 900
 5 AAAGTTCAGC ATCATTTTCA ATCACTTCAC TTAAATACTT ACCAATATAT CCTGTTCCAC 960
 CTGCTAATAA AACTTTAGGT TTCATCTAAA ATACTCCTTT AAAGTGTAA CAAAAACAT 1020
 ATTACTCCAC CTTTTAGTTA CATATATATT ATAATAGTAG CAAATGTTTT AAAATTTCAA 1080
 10 AATACTGGAG GCTTTTTATG GCCCATATTA TACGTAGAGT TAGTATCAAA GATGTAGAAA 1140
 ATTTCATTTT AATGTTAGCG AACATATACG ACGAATCTCC GTATATGTTT TACACACCAG 1200
 GAGAATATGA TCCTAGCGTC ACATCGGCTA GTAAACAATT AGAAGAATAT ATCACTTCTC 1260
 15 CGCATAAAGT CATCTTCGTT GCTGAAAGTG ATGAACAACT CGTTGGCTTT GCCTTTGTTA 1320
 ATACGACACC ATTTCAACGC ATTAAACATG TTGCTAAAAT TGATTTAGGT GTAAAGAAAT 1380
 TATATCAACA TCGTGGAATT GGCCAAGCAC TTCTTGATGC CATTATGGCT TGGTGTTTAA 1440
 20 ACAATCAAAT ACACCGAATT GAAGCAAATG TACCACTCAA TAACCAACCT GCCCTCGAGC 1500
 TTTTAAAAG TGCCGATTTT CAAATCGAAG GCGTTTTAAA AGATAAGTTA TTTATCGATG 1560
 GTAAATATTA TGATGACTAT ATGATGGCTA AAATTCTTAA TTAAAGCTAT TTTATCATAA 1620
 25 TCTTGATCA GAATCGTATA ACAACGAATT TAATGGTTAC CTAATACATT ACTCATACTT 1680
 ATCAATGTTA TCTAATCTCA AATAAATACG TACACTCTTA TTCATTTATC AAATTTAAAT 1740
 30 TCAAAATANA ACACCACTAA TGTGTAATTG ATTAATATC AACTACGATT AGTGGTGCTT 1800
 TATATATGTG GTTAGTTTTT CTTACTA 1827

(2) INFORMATION FOR SEQ ID NO: 451:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 616 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 451:

45 ATATAGATTA ATGTTTGTTA TTTGTACTGT GTTTGCCAGT GACATTTATT ATAAAACATA 60
 TGACGTTATT CTATGKTGTA CTCATTCTCT ATATTGTAGG TTTATTAACG ATTAGAAGTA 120
 TTATTAAAAA GTTGAAATAT CAGGAAACAT TATTACGAGA CTAAAAAACT TCCATTGGCA 180
 50 TGTATGTTGT AAAGGTGCAT GTAATGTTGA ACGCCAAATG ATACGGCGTT CAGATTACAT 240
 TAGCATCTAT ACGTTAACAG CATAACCAAT GGAAGTTTTT TTCGAATCTA TTCTTTTATT 300

55

AATGCGTCTT TTGAAAAATG GTCATTAAAG GCATCAGATT GCTTAAAGTC TTCGTATGCA 420
 TGTCGATCAG CAAATCCGAA ATAAATTTTG TATGTTGTAC CTTTAGCAGG TCTTAACAAA 480
 5 CGATAGCTTT TAAAGCCACC AAAGTTTCTG AAATTATCGT CTACACTAAT CAGTTTCTTT 540
 TCAAGTTGAT ATGCATGATC TTCTGTTGAT GGaATGAaG TTGCACaATA GAAATGaTGT 600
 TCAcTGAATT CACCAA 616

10

(2) INFORMATION FOR SEQ ID NO: 452:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 944 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 452:

GCACGAGTGA TTAAACGGTT AATCAATGAA ACATTTGATG CAAATTACAT TGAAGTTATT 60
 GAGGGAGGAA TTGAAGAAAC GCAAACGTTA ATTCACTTAC CTTTTGACTA TGTCTTCTTT 120
 25 ACAGGAAGTG nAAATGTAGG CAAAATCGTT TATCAAGCTG CCAGCGAAAA TTTAGTTCCT 180
 GTGACATTAG AAATGGGCgG TAAAtCTCCa GTCaTCGtTG ATGAAACAGC GAATATTAAA 240
 gTTGCTAGTG AGCGCATTTG TTTTGGGAAA TTCACTAATG CCgGCCaAAC ATGTGTTGCa 300
 30 CCAGATTACA TTTTAGTACA CGAATCTGTA AAAGATGATT TAATCACAGC CCTATCAAAA 360
 ACGTTGCGTG AATTTTATGG TCAAAATATA CAACAAAGTC CAGATTATGG CCGCATTGTA 420
 AACCTTAAAC ATTATCATCG TCTGACTTCA TTACTIONACA GTGCACAAAT GAATATTGTA 480
 35 TTTGGTGGTC ATAGTGATGA GGATGAACGT TATATAGAAC CAACATTGTT AGATCACGTT 540
 ACAAGTGATT CAGCAATTAT GCAAGAAGAA ATTTTTGGTC CTATCTTACC GATTTTAACG 600
 TATCAGTCAT TGGATGAAGC AATAGCCTTT ATTCACCAA GACCAAAACC TTTGAGTTTA 660
 40 TATTTATTTA GCGAAGATGA AAATGCTACA CAACGTGTAA TAAACGAGCT ATCATTGGC 720
 GGCGGCGCTA TTAATGATAC ATTGATGCAC CTAGCGAATC CTAAATTACC ATTTGGTGGT 780
 45 GTTGGTGCCT CAGGTATGGG ACGCTATCAT GGTAAATATT CATTCGACAC TTTTACACAT 840
 GAAAAAAGCT ACATTTTCAA ATChACACGA TTAGAATCAG GTGTCCATT ACCACCATAT 900
 AAAGGTAAAT TTAAATrCAT CAAACCTTtC tTTAAAnATT AATT 944

50

(2) INFORMATION FOR SEQ ID NO: 453:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4820 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 453:

	TCTCCAGTAG ACCTTGTGTA TGAACAGTTT CTTTCATATG AATGAACATC GTTTTTAAAG	60
10	TTTGTTCAC CTCAGCCTTA TCAGGATAAT ATCTAGAGAC AGTCGTCTCT GGCATCTCCA	120
	TTGTATGATA TTTAACCTTA TGCAGCGACC CATGATTTTC GTTATAAACA AATGTATGAT	180
	TCACTTCATC GAAATCATGA TCTTCTCCTG CAATCCAAA AACTGGTACT ACTTGTTGCT	240
15	TATGTGTATC CGTTAATTCC TTAGATAAAG TAATGATTGa AAATATTTTA TGAATGTAT	300
	ACAATGGTCC CCCGAAAAGC CCTGCTTGTT GTCCACCAAT CACAACTTTT GAACCATTAG	360
	CTAAATGTTG TATGTTTAAT TCTTGTTTAC TTGAAAGCTT GAAAGCTTTA AATCACTCAT	420
20	ATATTCACGA ATAACATTCC CTAATGCCGC TTCTCTTCCA TTATTTTCTT TAGACATCCT	480
	TTTTTCAAAA CTAGTTTGTT GAGCTGCATC ATATTGAAAT AATCCTGTTA TTACAGGGTC	540
	ACTGCTCTTT ATTTTGGTA TAACTGATC TTTTTCATTT AAATAACTA CTTTACAGTC	600
25	CATGTTTTTT CTCCTTAAGT ACGCGATTAC AACTATAGT ATAAAGTCTA TACCGGTGAT	660
	TGACAATTTT ACGGCTTGAA AATCAATTTA ATCATGAAA ATTTATAATA TTCATTGTTT	720
30	TACATTTTCA AATCAATGAA AAACACAAGT GGTTTAATGT ATAATAATAG TAGTAAACAA	780
	ATAAGGGGTA GATAAATATG AGTGAAATCA AACGTCTTGA AATTAAATTAC AAAACTGACG	840
	AATTATTCGA AAACTTTAGA GCGTTTGGCA ACAAAGACTT ATACATGGTC AATGAGTTAA	900
35	ACGGTCAAAT GATTGATGCA AGTTCAGATT CACCATTTTA TGGCATATTT GTCGaGATCA	960
	ATTAGGAGCT AGAATGGCAT TACTAAAAAA AGGTGATGTC GAAGAAATCT ACTTCCCAGA	1020
	TTTTGAAGAT TATATATTAT TATGGAAGTT AGAAGTATTA CCAAATATC AAAACAGAGG	1080
40	GTACGCTTCA GAATTGATTG ATTTTGCAA GAGTTTCAAT ATGCCAATTA AAGCCATTGG	1140
	CAGAAATGAT TCTAAGGATT TCTTTTTACA TCATGGATTT ACAGATGTGG AAGCTAAAAA	1200
	TATAGAGGGA CATGATGTCT TATTGTGGAA ACCATAAGAT AATAATATTC GACACTACGA	1260
45	GCAAGAAAAT GCATCTTTTC GTAGTGCTTT TTTTACAATT ACTTTCTTAA GCTAATATAA	1320
	GTaAATCATT TTCAAATTAT TTGTCTTAAC GTACAATATC ATTTAGTTGT TTCCATGrAT	1380
	TAATTTTATA ATCAGGTATA ATTCCTGGAT TATGATCAAA TCCTCTAAaA TTAAACCAGC	1440
50	AAGTAGCTAT ACCCGCATTG ATTCCACCTA GAATGTCAGA TGTTAnAGAA TCTCCaACTA	1500
	TAATCGAGTG CTGcCTTTCA TCCTCACCAA TATCATTAaA AACATAATTA AAAAATTCCG	1560

55

	ACGGCGTCTG ATTTAACCTT CTCTTTTGCG TTTCGGTTAC ACCATTAGTA ACAATATATA	1680
	AATCATGTCT TTTTCGATAAT TCGACAATTG TTTCTAATGT TTGATCAAAG TATTTAACTT	1740
5	TAGCTTCTGC TAATCCATTT CTAAATAACA CATCTGCACG ATGCCCATCA ACTTCCATTT	1800
	GATGATGTTT GAAGTAATTC ACAAATCGTT CTGATAATAC TTCAGACTTC GkTAATTTAT	1860
	TTTGkTGAAA AGCTTCCCAA TGTGGTGaT TGaTTTTTTT AAATGkTAAA AAATCATCCy	1920
10	TTGTTGCTTT ATGATTAAAA ACATTGCGCA TATAGTGAAn CGCCCATTCT TCTGcATCAT	1980
	AAAAATCAAC AATTGTATCA TCAAAGTCTA TCAAAATATT TTTATATCCC AATTTCGCCA	2040
	TCTCCTATAT TGTCTATGTA TCTAAATCTT AACAGAGGCT CAAATTTCGT CAAATAAAAT	2100
15	AAACTGAGTG CATAACATTA AAGTATGCTC ACCCAGTTTA TTTTAAAGAA TATTAGTTAT	2160
	TATATTAGAA TCCAAATAAT TTACCTAGTA AACCACACACC GTTAGCAACG ATGTCTACGA	2220
20	TACTTGCGCC TAATTCACA CTATCATGTT GTTGTGCAGC TTGCACAGTA TTTGCGATTG	2280
	CTTCTGCTAG TCCAGTCATT TAAATCTCTC CCTCACCTTT GAAATAATAC TGATTACTTA	2340
	CATAACATAT TGAAATTAGA ATCCGAATAA TTTACCTAAT AAACCTACGC CATTTTCAAC	2400
25	GATGCTCACA ATGCTTGTGC CTAATTTTGC GCCATCATT TTAATTGCTG CAGTTACGGT	2460
	ATCTTTAATT GCGTTAAATA AACCTTCCAT TGAAAACACT CCTTAAATT TAAATTTGAA	2520
	GATAACAAAA ACGTGCgTAG yTTTTAAATC ACCGAAATGT TATTCGCTTA ACGTTTTGTT	2580
30	GTTGTTATTT TAAAAATAAT TTGATGCAAT TAGTTTGTtT ATCCGCACAA CATCTTATAA	2640
	TGTACTTAAC TGTATTTTAA AGAGAAAAGA AATACAGTTA GGCATTCAA ACTGTATTTA	2700
	ACACAATTAA GTTGcCTGAA TTCGTATTTA AGTCTTATtG AACCTTTTTA GATAAATAGC	2760
35	TCTATAATAG TGAAAAATAT AAACATTTTT TATTTACAAG GTATTGCTAA TTTAAGTTCA	2820
	TTTAGATATA ATAATCTTG TGTGTtTAA CGTGTCTCGG TAGCTCAGCT GGATAGAGCA	2880
	ATGGCCTTCT AAGCCATCGG TCGGGGGTTC GAATCCCTCC CAGGACGTTT ATAGGTATTT	2940
40	TTATACGCAT TACCAAACAA AAGAGTTCCG TGATTACGGG GCTCTTTTTG TTTTGAATTT	3000
	CAGTAATATA GTATGATGCG TCACCAAAAC GTCCCCCGCA TAAGCCCCGA AAATACAGTA	3060
45	ATTAAAACAA GCATGCTTAT TCGTTATAGA ATTTTTTGAC ACACAATTGA CACGCGTCTG	3120
	ACACTTGTTT ATACATTTTT AATTAAGTAA TTTTGTGCTC AAATTCATC TATACTGCAC	3180
	CTGAACTACA CCAACACTAC ACCAAGATTT TTAACACTCA CCATTGTCAT GCGTAGAGAT	3240
50	TTTTATTATT ATATTATTCC TATAGATTTT GATACTATTC AAAATTTTAG GGACTTTTCA	3300
	GGGGCCCCGAA ATCCTATAAT TATAATTATA TACATCTAAA AAAAATAACC ACGTCCATCG	3360

55

5 ATACTATTGG CAAATTTATA AAGTAGTTCA GCGTTTTTCA ATGACATATT GTCTAATGAT 3480
 CTTTCATTTT TTCTCATTCT GTGTATTGTG CTTTGTGGAA CTCCTGTTTG TTTCGATATA 3540
 10 TGTAAGTGC TCAAATCACT GTCTAATAGT TTTTGAATTT GATTTCTCAT TTGTACACC 3600
 CCCTTGTAAG TCTTCAATCA TCATTTAAAT TAAATAATTA CTTTTCACA TATTCCAATA 3660
 TTTAGGTTGC AAAGCATACC TCAAATATCA TTAATTTTGA GATTTAAATG TCAATTATGT 3720
 15 GTCTATCAAT CCAATATACA TACTCTAATA ACCTAATAGT ACACACTCTT CTTATTAATT 3780
 GGCCATAGCT ATCATGATAT AATTAGTGAA GAAAATCACA TAAGAAAGGT TGTAATCAT 3840
 GAGACTTCAA AAAGCACCTC TAGTAACGTC AGGACTAGTC TTAGGATTAT TAGGCCTGGG 3900
 20 TAATCTATTA AAAGACTTAT CTCTTACTTT AAACGCTGTT TGCGGAATCT TTGCTTTCTT 3960
 GATTTGGATT CACCTTTTAT GTACTATGAT CAAATATTTT AATAATGTGA AAGAACAATT 4020
 AAACAGTCCT CTAGTTTCAT CAGTGTTCAC AACATTTTTC ATGTCTGGCT TTTTAGGTAC 4080
 TACTTATTTA AATACATTTT TTAGTAACAT AACTTTTATC AATAGCTTAA TAACGCCTAT 4140
 TTGGATTTTA TGCCTTGTGG GAATTATGAC GCATATGATT ATTTTTTCAA TAAATATTTT 4200
 25 AAAAGATTTT TCACTTGAAA ATGTTTATCC TTCGTGGACT GTACTTTTTA TTGGTATTGn 4260
 TATCGCAGGA TTGACGGCAC CCGTTAGCGG ATATTTTTTC ATAGGTCAAT TAACAGTAAT 4320
 ATATGGCTTT GTAGCTACTT GTATTGTCTT ACCTATAGTT TTCAAGCGAT TAAAGCATT 4380
 30 TCCATTGCAG ACGTCAATCA AACCGAACAC ATCGACAATT TGTGCACCAT TTTCTTTAGy 4440
 CGCTGCAGCA TATGTTATAG CTTTTCCTAA GCGGAATGCT TTTATCGTAA TTATATTTTT 4500
 ACTATTAGCT CAAATATTTT ATTTTTATAT CATTATACAA TTGCCTAAAT TACTAAAAGA 4560
 35 ACCTTTTTTCG CCCGTATTTT CAGCTTTCAC ATTCCCTTTA GTAATCTCAG CAACTGCTTT 4620
 AAAGAACAGT TTGCCTGTAC TTATGKtCC AGACATTtGG AAAGGKCTTT TGTTTATCGA 4680
 AGTGTATTAA GCCACTGTAA TAGTACTTAG AGTCTTTATA GGATATCTTC ACTTCTTTT 4740
 40 AAAAAAGGAA AAACAAGATA AATTTCTnCG TAATGCGTCT CAGTAACACT ATTACCAAGA 4800
 ATTAACACGT ATATTTAATA 4820

(2) INFORMATION FOR SEQ ID NO: 454:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4358 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

	ATTGATCT TTAATATCAC CAATATTTTT AATATCTTCC GGATTCAATC CATATACITG	60
	TACTGTATCT GAGTATTTAA TTGTGAAATA ATCACCTGAT TTAACCTTGT CATCAACTGT	120
5	AATTTGTGAT TTTAATGATA AATAATCTTG GGCTGGTACG ATTTTATTGT TTTTATCTGC	180
	ATCAACGACA GTAAATGTTG TATTTGATGT GATTAAATCA TTAACATTTT TAGCCTCTGT	240
	TGATGATGGC TGTACTGCTG CTATACGCAT TCTGTATTG AAACGTTTAG GTGCTGTACT	300
10	TTTTGGCAAA ATGATATCTG CATTATTTTC ATTATTTGAA TTACTATTGT TATCAACAAG	360
	AGTTTCATCA TTAATCTTGA TAGCATCACT TTTAACATTT AATGTAGTTG ATTCAGTTTT	420
15	GGCATCTACC TTTTGTGTTT CCTCATTAGT TGGTTGAACA TTTACCACTG ATTTATTCTC	480
	TTGCAAATCA GGTGTGAACG CTTCTTGATT ACTTATAGTT TGTTTAGTGT TTAAATCTTC	540
	ATTCGTAGAT TTTGGTGAAG CTTGCTCATC TGATTTGGCA GTTGAAACTT CAACTTTATT	600
20	TCCAGTGGTA GATTGTACAC TTTCTTTTTT TATTAATTTA TTCCCATTTG AAGTCGTTTT	660
	ATTACCTTGa GATGATACCA TTTCTTTTTG ATTATCATTT TTAGTATTGT CTTCTTGATT	720
	TAGTTGCTGC ATATCAACTT TATCACTCGA TTGATTATCA CTTGCTGAAG TTGTCGCTcG	780
25	TTCAATTCTT TATTAGTACT TTCTGCAGCC TTTGCTTCTT GGTTCGCCAG ACCAAAAATT	840
	AATGTTGTAC CTAATAAAAT TGATGCTGTT CCCACTGTGT ACTTTCTAAT CGAAAAATTA	900
	TTTAATCGAT TGGATACCAT GCCTTTCCTT GTTATTGCCG TTTATTTTTC TCTGTTTAGC	960
30	ATTAGATTAC TCCTAATTCA TCAAATTTTT AAATAATACA ATTGTTTTAA ATACAAAAAT	1020
	GTATATCAAT ATAGTATTAC ATTTTGTAGT AAAGCACAAT ACTTTAATTA TTTTCTTTA	1080
	TCGTAAAACG TTATTTAACA TTTGTGTTTA AATAAAAGTT TTTATGAGTt TTGTAATCTT	1140
35	TATTTAATCA TCATAAAAAA TAGTATTATT TGCCCTTGAA ATTAATATCT TAGCTTTTCT	1200
	AATTCATAGA CAATTACATT TCTGTAACAA ATTAAATTGT ATCTATTCCCT TAAAGATTTT	1260
40	TTGTTTTATA TCTGGGAATT TCTAAACAGA AAAAACCAGG CCACATGGAC CTGGTTAAGT	1320
	TAATCATATT ATTTATTTTG TTTTTTACGA CGACCGAATA ACAATAATGA TCCTAATGCC	1380
	GCGAATAATC CACCGAATAA TGTGCCATTA TTTGAATTAT TATTTTCACT ACCTGTTTCT	1440
45	GGTAATGCTT TAGCTGTTTT ATGCTGATCT TTAACCGTAC TCATTGGTTT AGCCGGAGTA	1500
	TGTTTACCTG CATCTGAATC TGAATCGCTA TCTGAATCTG AGTCGTTGTC TGAGTCCGAA	1560
	TCGCTATCTG AATCTGAGTC GCTGTCTGAA TCTGAATCGC TATCCGAGTC TGAGTCGCTA	1620
50	TCTGAGTCTG AGTCGCTATC TGAATCTGAA TCGCTGTCTG AGTCTGAATC GCTATCTGAG	1680
	TCTGAATCGC TGTCCGAATC TGAGTCGCTA TCTGAATCTG AATCGCTATC TGAATCTGAG	1740

55

	TCTGAATCTG AGTCGCTGTC TGAATCTGAA TCACTGTCTG AGTCTGAGTC GCTGTCTGAG	1860
	TCTGAATCGC TGTCAGAATC TGAGTCGCTA TCTGAGTCTG AATCTGAATC ACTGTCTGAG	1920
5	TCCGAATCGC TATCTGAATC TGAATCGCTA TCTGAGTCTG AGTCGCTATC CGAATCTGAG	1980
	TCGCTATCTG AGTCTGAGTC GCTATCCGAG TCTGAATCGC TGTCTGAGTC TGAGTCGCTG	2040
10	TCTGAATCTG AATCGCTATC TGAGTCTGAG TCGCTGTCTG AATCGCTGTC TGAATCTGAG	2100
	TCGCTATCTG AATCTGAGTC GCTATCTGAG TCTGAATCGC TGTCAGAATC TGAGTCGCTA	2160
	TCTGATGTTT CTTCTTCGTA GTAGCCATTA TCAAGTGTGA AATCATCATG ATCCGTAATT	2220
15	GTTACATCAA CTTCCGCCACC ATCGGCATCT TTATCATCTT CAGTTGTATT TGTACCTGTT	2280
	TGAGTTAAGC CAGCAGGTTT TTCAAAGATA ACTTTGTATT TACCACTATC TAAATTATCA	2340
	AAGCGGTATT TACCATTTC ATCTGTYTCA GTTGTACCAA TTACTTCGCC TTTTTCGTTT	2400
20	TGCAAAGTAA CTTTAACACC TTTAATTCCT TTTTCAGTCG AATCTTGTTT ACCATCTTTA	2460
	TTACTGTCTG ACCAAACATA ATCACCTAAA CTATATTTTG GTGTTTTGTA GAATCCACTA	2520
	TCTAATGTCA TGTTGTCAGC GTCTTTAATG ACACCTGTTG TAGTTAGTCC ATCAGAATCT	2580
25	ACAGCATCAT CTGTACCTAC ATTTGCAGTT GTCGGTGTAT AACCGGCTGG TGTTGAAAAC	2640
	TCTACACTAT AAGTTCCATT GCTTAAACCA GTGAAGTAT ATTTACCATT TTCATCTGTT	2700
30	GTCGTACGAT CTAATTCTTT ACCGTTACTA TCTTTAAGAA TGACATAAAC ACCTTTAATC	2760
	CCTTTTTTCAT TGGCATCTTG TTTACCATCT TTATTTGTAT CTTCCCATAC ATAGTCACCT	2820
	AGATTATATT TCTTTTGTC GCCATTAGCA GTTGATGAGC CATTACATT TGAATAACTA	2880
35	TTTGACCAAC TATATTTAGT TTTGTCAGTG TCTAAAGTAT AATCAATTTT TCCATTATCT	2940
	GTTGAAGTAT TATCTGGATA AGCAACTGTG TGAATGATGT ATTGTTTATT GCTGCTTGTT	3000
	TGGCCTTTCA TTAAATCGAC TGTAGCTGTT TTATTATCAT TACTATAAAT AACATCGAAT	3060
40	TGATCAGTAA CATCTTTAAG TTTTGAAGTA TCAGGGGTGA AACTATCCAC AAATTGATTT	3120
	TGATCTGTCA CTTCGTAAAT TTTGAAGTTT TTTGCATTG GATTAAATTT ATATCCAGTT	3180
	AAATTAGTAA CAAACGTTTG TTTAGTATAT GTATTTTGTAG GTTGATTAC ATATGCAGTC	3240
45	ATATTACGCG ATAAATCTTC ATTGTTAATA TAGTTTGTAC TTGAAATAAG CGGTTGTGCT	3300
	TTTTTATTAC CATAATCGAC AATGATTCT TCGCTATATG TATCATTACC TAAAGTTACT	3360
	TCCATTTTAT AAGCTGTTTT ATCAGTTGTT GCATTTTAC GTTTCGCAA TGCAACTTGT	3420
50	TCAAAGCTAC CTCTAACATT TGTATATTGA TCTACATAGT TCGTAAAAGT ATATGTTGTT	3480
	GTGTTTGTG TACTATCATA AATACCTTTT GCAATAATAT TACCTTGGGC ATTATATAAA	3540

55

	GTAAATGTAT CGCCCTCTTT AACAGAATCA TCGATTGTGT AATTTGCTTT TAATTTTAAA	3660
	ACATCACTTG AAGTTGCCCA AAATTCAGTT TTACCAGTAG TCTGATTAAAC ATGTCCTTTA	3720
5	TCAATCGCAA TGTCAATATT TGAAAAATGT ACTTTATCAT TAACATTTGT TCCTTGTTGT	3780
	GGAGCTGCAA CAGTATTCAC TGCCATGCGA TTAAAGTTC TTGGTTTAAT AGTCGTTGTT	3840
	TTAGGTGTAG TTGAAACATC TTTTGCTTGT GTTAAATTAC TTTTATCAGT TTCATTACTA	3900
10	TATGTAGTTG ATGATTTATC ATTTGTTGTT ACATTGCTAG TTTTGTAGT AGATTGATTA	3960
	GCTGTAGCGT TTTGTGGTGA TTGCATGTTA CTACTAGTTT CTTTAACTGT TGCACATCA	4020
	CTCATTGTCA CTTTAGGCTG ATCTGCAGTT GCAGTTTGGC TATTGTCTTT TAGTTGACGA	4080
15	CTATCAACTT TTTTAGTTGT TTTATTCTCA CTGGGGCTG TCGTTTCATT TTTTGATTGA	4140
	TTTAATTCTC CATTCGTATG TTCTGCCGCT TTAGCTTCAT GACCACTTAA CCCAAAAATC	4200
20	AATGTTGTCC CTACTAAAAT TGAAGCAGTA CCTACAGAAT ACTTCTTAT CGAAAATTTG	4260
	TTTAATCGAT TTGGTATCAT GCCTTTnCTA TTTGTnGCTG TCTTTTTATA ATTCATTAA	4320
	TAATACTCCT TTAAATATC AAAATTTGAT AAATATAA	4358

(2) INFORMATION FOR SEQ ID NO: 455:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 455:

35	TTGACTTCTT AATTCAGCAT TTTCTGCACT TAATGCTTTG TTCTTTTAA TAAGTTGCTT	60
	TCTTGATAA ACTTCGGTAT CTATTTTACT ATTACTATAC CTTTGATTTA AACTAATAT	120
40	ACCAATTAAT GCTACAATGA TAATGATAAG TACAACATAA AAAGACATTT TTCACCAAT	180
	CCTTTTGTGAC TTCTTTAACT TTGTATACAA TAATAATTAA TAAAGATTAA TTGTTATTCA	240
	ATTTCCCACT TTTTATTAG TTGATTTTAG TTCATCATTG TTATAATCAA ATTATAAACT	300
45	GACAGATATT GATGTTCAAT GAATATGACG TGAAAGATTC GTGAATTCAA GTTTATGTCG	360
	AATTTATGTT ATAACGGTCA TTAAATGAC AGAATTAGGT CACTCATAGT ATTTTGAAGA	420
	TTGAATTCAT TAATTTTAAA ATGTATAATG ATATTTGTGA AAGCGCTTGC TTAGGAGGTG	480
50	TATTTGAGAG TGAATGAAAT GAATGCTAAA GAACAATTAG TGGACAATTT AATGAAAACA	540
	TCATCGCAAT TATTTAAATT TCACGGTGAA GTTGCCATGC AGCTTTTCTT AAATGATGAA	600

AAAGTTATTC CGCAATCATA TGC GTTACTA TACATAGATA AGCAAGATCA AGCAATAGCT 720
 AAAGAAGATT TATCACTTTC AAAAATTGCA AAAGTTTATG TGCAATATGA TGATACAACA 780
 5 ATAATGAGTA TTTTCGTTTA TGATGTAGTA AACGATGAAT GGATTTTTAG ATTGGATCCG 840
 AATATACGTA TACCTAAGAG TAACATATAC TTCCATAGTT TAAATTGGGA TGTGGATATA 900
 TTAAACCGGA GtCGTCTAAT GTATGTCTAA TGCaCACCAT CAGaTCATCA TtATCCATTA 960
 10 TAACGrGCAT AGTGCATAAn yACTwCatTT TaTTaaATTG AGaGGgGCAC GATAGGTGCA 1020
 TCAGGACATA ATATAGGAAG CATCAACGCG TGAnCAGGTC 1060

(2) INFORMATION FOR SEQ ID NO: 456:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1262 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 456:

ATAATnACTA AATACnAAAG TTAACTGTC TTACTAATAA TGA CTATGTT ATAATTTTAA 60
 AAGTGATATT TTGGGTAATC GCTATATTAT ATAGAGGAAA GTCCATGCTC ACACAGTCTG 120
 AGATGATTGT AGTGTTCGTG CTTGATGAAA CAATAAATCA AGGCATTAAT TTGACGGCAA 180
 30 TGAAATATCC TAAGTCTTTC GATATGGATA GAGTAATTTG AAAGTGCCAC AGTGACGTAG 240
 CTTTTATAGA AATATAAAAG GTGGAACGCG GTAAACCCCT CGAGTGAGCA ATCCAAATTT 300
 GGTAGGAGCA CTTGTTTAAC GGAATTCAAC GTATAAACGA GACACACTTC GCGAAATGAA 360
 35 GTGGTGTAGA CAGATGGTTA TCACCTGAGT ACCAGTGTGA CTAGTGACG TGATGAGTAC 420
 GATGGAACAG AACATGGCTT ATAGAAATAT CACTACTAGT TTAGCTCTCC TAGATGATGG 480
 AGAGCTTTTT TCATGAAAAG AACACTTAAA ATTAACGCCy TGTCTTGaTA tAATGACaCT 540
 40 GCcTTGTTTT AAAATAGTAA GCGGATGCgt TAATGTATCA GCGATTAAAT TTGTTGGAAA 600
 TGTATAAAAA ACACAAGCTA AGAATAAAAT ACCTGTATAA AAGGAGAATC ATATATGTTT 660
 CAATTACTTG CAGTTTGTCC GATGGGATTA GAAGCTGTTG TTGCTAGGGA AATTCAAGAA 720
 45 TTAGGCTATG AAACAAATGT TGAAAATGGT CGTATATTTT TTGAaGGAGA CGCAAGTGCA 780
 ATTGTAAAGG CAAATTTATG GTTGCGCACA GCAGACCGAA TCAAAaTTGT TGTGACGT 840
 50 TTAAACGCAA CAACGTTTGA CGAATTATTC GAACAAACCA AAGCGCTCCC TTGGGAATCT 900
 ATAATTGATA AAGAGGGTAA CTTCCAGTT CAAGGTAGAA GCGTTAAATC AACACTACAT 960

TATAACGAAA AAGGTTGGTT AAATGAATCA GGTGCCAAAT ACCCTGTTGA AGTTGCCATT 1080
 TTAAGAGATA ATGTATTATT GACTATCGAC ACATCAGGTT CTGGTTTGAA CAGACGTGGT 1140
 5 TATAGATTAG CACAAGGTGA AGCACCAATT AAAGAAACGT TGGCAGCAAG TTTAATCCGT 1200
 CTTGCCAAAC TGGGAAAGGT GATTACACCT TTAATTGGT CCCATTGCG GTTCnGGTTA 1260
 CA 1262

10

(2) INFORMATION FOR SEQ ID NO: 457:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1142 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 457:

CCTGGCTGCT TTATCAGCAT CTACTACTTT AAAACCGAAT ACGGATAAGA GTTCTGATAC 60
 TGTGATTTT CTGAGGCGA TTCCACCTGT TAGACCAATA ACTTTCGGCA TAATTTCACT 120
 25 CTTCTTTTAT TTTTGACATA CTGGACAATA ATGACTATTT CTGTGCGCA TGATTTTGT 180
 TTCAATTTGA CTTCCACACA CTTGCATAC CGGCTGCTTA TATACATTAA GATGCAATTG 240
 CATCTACCA GTTTTTCCAT CAGCATGACG ATAATCTGAA ATACTTGTAC CGCCATATT 300
 30 AATACCTTCT TCTAGTACTT CTCTAACATA ATAAAAAACC ATTTCTTGTT GTTGGTGTGT 360
 TAAGTCTTTT ACTTTTTTAT CTGGTAAAAC ACCTGCACGA AACAACGCTT CACATGCGTA 420
 AATATTTCCA CAACCTGCGA TTACTTTATG ATCCAAAATC ACTTGTTTGA TTGGTTTATT 480
 35 CTTATTAGAC TGTTGATGAA TTCGATTTAA ATAATACGTC AATGCTTCAT TTGAAAAAGG 540
 TTCAGGCGCT ATTTCTAAAA ATGAAGGATA AGATGCTACA GACGCAACAT TTCTAATTC 600
 TCCAAAACGA CGTATATCTG AATAAATTAA CTTTTTGTCA TTTGACAACT CAAAAATAAC 660
 40 ATGCCAATGC TTACGATAAT TAGGTATCAT AATATCTTCA AGTTCATCTA CAATGAAAAA 720
 ACCGCCCGCC ATACCTAAAT GACTAATTAA TGTACGTTGT TCTCGTTTAT TATCTAGCTG 780
 45 AAAAAAGATA TATTTACTTC TTCGTTCTAC ATTTGTAATG GTATAGCCTT CCGATAAAGT 840
 TTTAAAAGTA TCTAATTCAA TTCCTTTTAT AATTGTTTCC TTGCCTTGAG CTTTACCTTC 900
 GATTACTTTA TCCGAAAATA TAACGTGTTT AATTTTTTGA TTTATAACGT AGGGTTCAAT 960
 50 TCCTCTTTTT ACATGTTCTA CTTCTGGTAA TTCGGGCATA CCATTAACT CACTTTATTT 1020
 TGCATCATAC CAGGTTGCAC CATAACTTGA GTCTACTTTT AATGGAACAT CTATTGCAA 1080

55

TT

1142

(2) INFORMATION FOR SEQ ID NO: 458:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1814 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 458:

5	CCTTTAGTAA ACAATCCTTC TTTAGTTTTA GTACGTTGTT CCAATCCAAA TAATTTGTAT	60
15	TTCATTGCCT CGCCCGATTG AGTGCCGCTA AAGTTATCAT CTTTCATGTT AGGCGTGTG	120
	GTAAACATGT GTATATCACT GTTTAAACGG TCTTTATAAG CTTCGGTACC TTGTACATCG	180
20	TATTGCTTAT AAATATAACC ACCATCAACA GAGCCTTCTG TTTCTCTACC TTCGCTATCA	240
	GCATAAACAG TCGTTCTAA AAACAACACG TTAGCTTCCT TTTGTTTTCT AACTTCTACA	300
	GGATCTAAAT TTAAATTACC TTTAATAAGT AACATAGCGT CATTTAAATC ACTCATATAG	360
25	TTAGCAsymy CTGATTCAGC ATTATCATAC AAATCAATTA AAGTGATTAC TTTCTCATAA	420
	TCCCCTTTTC TTCTTTGTT GTTGCTAAAT TCTGTAATAG GCATACGTTT GAAAGAGTGT	480
	GATTCAAAAC CGTTTTACG TGGTGTGAGC TTCAATCCAT TTGTTCTACT GGTAAGATAT	540
30	CTATAAACAC CGTGTGAAGT GAATAAATCA ACTGTAAACA CTTTCATCTT CAGTCTTG	600
	TCTATTGGTT TAGTTCTTAA ATATCTAAGC CCTGCGATAC TATTACGTTT AATTGTATTG	660
	TCGTATATGA CAAAAGTACT CATTGCATCA CTCTTGATA AACGCGTTT ATCATCTTGG	720
35	TTTCTAATCA TTAATCATA AGCTTTGCCA TAAATTGACA AATCTAATCC TAAAGATCTA	780
	TTGTGTGACT CAACATCATT TAAATCATTG AACGCCTCAA TAACTTCTAA TACATCTTTG	840
40	TCATCATCTT GATATTGAAT TGGATTACCC AAGAAATAGC CGTTGATAAA ATCGCTAATA	900
	TAAGATGCGT AATCATGCGC TACACGGTTA TCTGCCATGT ACTCTTCTTT GCGTCGTGTT	960
	AACTCAACTA AGTTCCTAGT TTTACCTTCG TAATAATCAC TTAACACTTT CAATCTAGGT	1020
45	CGTTGGTAAT CCATGTGATG TTCAATGTAT TTACTTACTT CATTAAAGTT TTGTAATAAA	1080
	TCGGATTCCG TCCCGTCATA TGTGTAAACA ACATTGGCTT CATCATTAAA TAAGTAATTT	1140
	ATGTTTCCCC GTAGATCTGT ATCTGTTTCA AATTCGTTTA CTTTAAACAT TTGTTCCCTC	1200
50	CTATAATCCT AGAGATTTTA TTGTGTCAAC TTTGGAAGT AGATTTGTGC GTTTCTAAC	1260
	CGGTCTGTAG AATCGTTCCA CTGAATAACG CAACGAATCG ATACAATGAT TGTATGTATC	1320

55

CTCTTCAATA GTCTTGAAAC AACGTTTCATC AACAAATGATT TCAAATTGCA TTAAGAATTG 1440
 TAACCCCTTGT ACAACCGAGC CCTTCCCTTT TTTGGTTGGT AAAATCCTTT TAAGCCCTAG 1500
 5 ATTCCTTAAT TCAGCTATAC TTTTTTGTTT TGCATATCT GCTGTAATTT CTTCTTTAGC 1560
 ATAACCAAGT TGCTTTATGA CATTAGCTAT TTCATCATT AGCATACCTT GTTTAACATA 1620
 CTCTTCAATG ATGTATAACT TCTTTTTCTT TACATCTATT TTAGAATGTA TAAAAGCACT 1680
 10 AGGATCATT ACGTAGCCAA AGTCCAATCC AAAATAAGAA GGTAATGTC TTAATCATC 1740
 TTTATTTATT AAACGTTTTT CATACTTAGG GAAAACCAAT TTGTCTAGTG TAGCAAATTC 1800
 ACCTAACGCA TAAA 1814

(2) INFORMATION FOR SEQ ID NO: 459:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 686 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 459:

AATTAGATT ATTACCCTCC TTAAAAATA TTTGTTTCAC AATTTTTTAT TACCTATTTA 60
 CTGGTTTCAT GTCTTATGGG GCATTTTACT TGTGGTATT TGAAATGTG CAAAATTTAA 120
 30 TCTTATATGT TTCTGGCTT TTCATGACTA TGCTATTTAT GTTTATGAAT ATGCATTCAA 180
 TTATAGATAA AAAAGTACAT ATATTCTTAA AGTCTAATAA ATAGTTACAA ATTTAGTTAG 240
 TTTTCAATTG TTAATTAGGG GTGGTAAACA GTGCTTTGTG AATCTAGACA AATTTATAAA 300
 35 AATCCTAAAT ATCGAGTTAT TAGATATAAT AATGAATATT TCATGGTCGA TTTAGTAAGT 360
 ACTTGGATTA CTTATTTTTT CCCTATGATT AATTGGTTTT TGCCCAAAA ATACGCAAAA 420
 ATTAGCGAAA ATGAATTTGA AAGGTAAAT ATAGTCGAGC CTGTTAAAA TAATGTTTTT 480
 40 TGGCCGGTTG CAGGAAGTTC AGTTCTATTT GGAATTATAT TGAGAAAGTA CGGTAACCTC 540
 TTTAATGTTT AGTTTGAAAA ACAACTAGCA ATCACTGTAT TTTTATCAT GTTAATAGGG 600
 45 ATGTTAATTT TTTATTTTTT TCTAAATAAA AAATTAACAT TAAAAATTTT TAATACCAAC 660
 GTGGGTAATA AGAATAGGAG TTGTAT 686

(2) INFORMATION FOR SEQ ID NO: 460:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1300 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 460:

5 ATCTGCAATT ATGGGCACAC CCAAGCThAT GCAAGTAACT AAAGGAGAAG TACTTTTAGA 60
 CGGTGTAAAT ATTTTAGAaT TAGAAGTTGA TGAAAGAGCA AAAGCAGGAT TATTCTTGGC 120
 AATGCAATAT CCATCAGAAA TTACAGGTGT TACAAATGCT GATTTCATGC GTTCAGCAAT 180
 10 CAATGCGAAA CGTGAAGAAG GACAAGAAAT CAACTTAATG CAATTATTATA AGAAATTAGA 240
 TAAAAACATG GATTTTCTAG ACATAGATAA AGACATGGCA CAACGTTATT TAAATGAAGG 300
 TTTCTCAGGT GGAGAGAAGA AACGTAACGA AATCTTACAA TTAATGATGT TAGAACCTAA 360
 15 GTTTGCaATC TTAGATGAAA TCGATTCAGG GTTAGACATC GATGCATTAA AAGTTGTATC 420
 TAAAGGTATT AACCAAATGC GTGGGGAAAA CTTTGGTGCA TTAATGATTA CACACTATCA 480
 20 ACGATTATTA AATTACATTA CTCCTGATAA AGTACATGTA ATGTATGCTG GTAAAGTCGT 540
 TAAATCTGGT GGTCCAGAAT TAGCAAAACG TCTTGAAGAA GAAGGATATG AATGGGTTAA 600
 AGAAGAGTTC GGTTCAGCTG AATAATCTTA TTAATACAGT ATCCATGAGA TGTTCATCTA 660
 25 TATATGATGA AAATGAACAT TTATACGAAA TAGTAAATTT CATCAAGTAG GAGGAAAAAG 720
 TTATGACAAC TGATATTTTG rACaTTyCTG AAGAACAAC TGTTGATTAT TCTAAAGCCC 780
 ACAATGAACC TTCTTGATG ACAGAATTAC GTAAAAAAGC TTTGAAATTA ACAGAAACTT 840
 30 TAGAAATGCC AAAACCTGAT AAAACAAAAT TAAGAAAATG GGATTTTGAT TCTTTTAAAC 900
 AACACGATGT AAAAGGTGAT GTTTATCAAT CTTTATCACA ATTACCTGAG TCAGTAAGAG 960
 AAATTATTGA CGTAGATCAT TCTAAAACT TAGTAATTCA ACATAATAAT ACGATTGCGT 1020
 35 ACACACAAGT TGATGATAAT GCATCGAAAG ATGGCGTTAT CGTTGAAGGT TTAGCAGACG 1080
 CTCTTATGAA CCATAGTGAT TTAGTACAAA AGTACTTTAT GAAAGATGCA GTAACAGTAG 1140
 ATGAACATCG TATCACAGCG CTACACACGG CATTAGTTAA TGGTGGCGTA TTTGTTTATG 1200
 40 TTCCTAAAAA TGTAGTTGTA GAACATCCAG TACAATACGT TGTGTTGCAC GACGACGAAA 1260
 ATGCAAGCTT TTATAACCAT GTTATCATCG TTACTIONAAGA 1300

(2) INFORMATION FOR SEQ ID NO: 461:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3135 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	GACAGCAAAT TCAAGATACA TTAAATAAAG ATATTGTCAT AAAGCATATT CTTGTTGAG	60
	ATAAATCTAA AAAGAGACCG CTAAATATTA GCCAATATCA TTTAACTGAA GATGTTAATG	120
5	AAATTTTAAA TGATGATTCA TTAGATATTA TCGTTGAAGT CATGGGAGGA ATTGAACCAA	180
	CTGTAGATTG GTTAAGAACA GCACTTAAAA ATAAAAAACA TGTATTACC GCAAATAAAG	240
	ATTTATTAGC AGTACATCTT AAACTTTATG AAGATTTAGC AGAAGAAAAT GGTGTAGCTT	300
10	TAAAGTTTGA AGCGAGTGTA GCAGGTGGTA TTCCGATCGT AAATGCCATA AATAATGGTT	360
	TGAATGCGAA TAATATTTCA AAATTTATGG GAATTTTAAA TGGTACCTCT AATTTTATTT	420
	TATCTAAAAT GACTAAAGAG CAAACGACAT TTGAGGAAGC ACTTGATGAA GCGAAAAGAC	480
15	TTGGTTTTGC TGAAGCGGAT CCAACTGATG ATGTAGAAGG GGTAGATGCA GCGCGTAAAG	540
	TTGTCATTAC ATCATATTTA TCATTTAACC AAGTCATTAA ATTAAACGAC GTTAAACGAA	600
20	GAGGAATTAG TGGCGTAACT TTAACGATA TTAATGTAGC CGATCAACTG GGGTATAAAA	660
	TTAAATTGAT TGGTAAGGGA ATATATGAAA ATGGCAAAGT TAATGCATCG GTAGAACCAA	720
	CGTTAATTGA TAAAAAGCAT CAATTAGCAG CTGTAGAGGA TGAATATAAC GCGATTTATG	780
25	TTATTGGTGA TGCCGTTGGT GACACGATGT TTTATGAAA AGGAGCAGGC AGTTTAGCAA	840
	CAGGTAGTGC CGTTGTCAGT GATTTATTGA ATGTAGCATT ATTCTTTGAA TCAGATTTAC	900
	ACACATTGCC ACCACATTTT GAATTAAAGA CAGATAAAAC ACGGGAAATG ATGGATTCAG	960
30	ATGCAGAAAT TAATATTAAA GAAAAATCCA ATTTCTTTGT AGTAGTGAAT CATGTCAAAG	1020
	GTTCAATTGA AAATTTTGAA AATGAGTTAA AGGCAATATT ACCATTTACG CGATCATTA	1080
	GAGTTGCAAA TTACGATAAT CAATCATATG CCGCTGTTAT AGTTGGATTG GAATCATCAC	1140
35	CGGAAGAATT AATCACTAAG CATGGATACG AATTGACAAA GTATACCCAG TAGAAGGAGT	1200
	TTAATTATAA TGAGAAGATG GCAAGGATTA GTAGAAGAGT TTAAAGCACA TTTACCAGTA	1260
40	AATGAAAATA CACCAAAATT AACATTGAAC GAGGGAAATA CACCACTCAT TCATTGTGAA	1320
	AATATGTCTA AAATACTAGG CATAGATTTA TATGTGAAGT ATGAAGGTGC CAATCCGACA	1380
	GTTCAATTTA AGATCGCGGT ATGGTAATGG CTGTGACAAA AGCAAAAGAG CAAGGTAAGA	1440
45	AAATTGTAAT ATGCGCTTCG ACTGGAAATA CATCAGCGTC TGCAGCAGCA TATGCAGCGA	1500
	GAGCAGGTTT AAAAGCTATC GTCGTAATAC CAGAAGGTAA AATTGCATTA GGTAAATTGT	1560
	CGCAAGCAGT AATGTATGGT GCAGAAATCG TTTCTATTGA AGGAACTTT GATGAAGCTT	1620
50	TAGAAATTGT AAAAGAAATT GCAAAAAGTG GCGAAATCGA GCTTGTAAC TCTGTCAATC	1680
	CATTTAGAAT CGAAGGACAA AAGACAGGCT CATTTGAAAT TGTACAACAA TTAGACGGTG	1740

55

AAGGCTTTAA AGAATATCAT GAAGCTAAAG GATCACAATT GCCGAAAATG TTTGGCTTCC 1860
 AAGCTGAAGG CGCATCACCA ATTGTTCAAA ATAAAGTCAT TAAAAATCCT GAAACGATTG 1920
 5 CAACTGCTAT TCGAATTGGT AATCCTGCTA GTTGGGATAA GGCGACTAAT GCTCTTAAAG 1980
 AATCAAATGG ATTAATAGAT AGTGTTACTG ATGATGAAAT TCTAGAAGCA TATCAGTTAA 2040
 TGACAACTAA AGAAGGTGTC TTTAGTGAAC CAGCGAGTAA TGCTTCTATT GCAGGTTTAA 2100
 10 TTAAATTGCA TAGACAAGGT AAATTACCTC AAGGTAAAAA AGTAGTTGCT ATTTTAACTG 2160
 GTAATGGATT AAAAGATCCT GATACTGCTA TTTCACTACT AGATAATCCG ATAAAGCCAT 2220
 TGCCAAATGA TAAAGATAGC ATTATCGATT ATATTAAAGG AGCTTTATAA CATGTCGAAT 2280
 15 GTTTTGGAGT TAACAATTCC TGCATCAACA GCCAACCTTG GAGTTGGCTT TGATTCTATA 2340
 GGTATGGCTT TAGATAAATT TTTGCATCTG TCTGTAAAGG AAACATCAGG GACAAAATGG 2400
 20 GAATATATTT TCCATGATGA TGCATCTAAG CAATTGCCTA CTGACGAAAC AAACTTTATT 2460
 TATCATGTAG CACAACAAGT TGCTTCTAAA TATAGTGTG ACTTGCCTAA TTTATGTATC 2520
 GAAATGAGAA GTGATATTCC ATTGGCAAGA GGGTTAGGTT CGTCAGCTTC TGCTTTAGTA 2580
 25 GGAGCTATAT ATATCGCAA TTATTTTGGT GATATCCAAC TGTCTAAACA TGAGGTATTA 2640
 CAATTAGCGA CTGAAATCGA AGGACATCCT GATAATGTTG CGCCGACCAT TTATGGTGGT 2700
 TTAATCGCTG GATATTATAA TGATGTCTCG AAAGAAACGT CaGTTGCACA TATCGACATA 2760
 30 CCAGACGTGG ATGTGATTGT AACGATACCA ACTTATGAAC TAAAAACAGA AGCATCAAGA 2820
 CGTGCTTTAC CACAAAAATT AACACATAGT GAAGCGGTTA AAAGTAGTGC AATTAGTAAT 2880
 ACAATGATTT tGgCATTAGC ACAGCACAAT TATGAATTAG CAGGTAACT CATGCAACAA 2940
 35 GATGGCTTTC ATGAACCGTA TCGTCAGCAT TTAATTGCTG AATTGATGA AGTGAAAACA 3000
 ATTGCTAGTC AACATAATGC CTATGCAACT GTAATTAGTG GTGCTGGACC AACTATTTTA 3060
 ATATTTAGTC GTAAAGAAAA TAGTGGGGAA TTGGTTCGCT CTTTAAATAG TCAGGTAGTA 3120
 40 TCATGCCATT CTGAA 3135

(2) INFORMATION FOR SEQ ID NO: 462:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1209 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 462:

55

AGnTCAATAT TTAGATCAAG CCGTTTAAAG TAATTACGAA CAAGTTTATA TCATTTCATGG 120
 TAAAGGTACA GGTGCACTTC AAAAAGGTGT ACAACAACAT TTGAAAAAGC ATAAAAGTGT 180
 5 TAGTGACTTT AGAGGTGGTA TGCCAAGCGA AGGTGGATTT GGC GTTACCG TTGCAACACT 240
 AAAATAAATT ATAATTGAT AAATTAAATA GCTGCAGTTA AAATAATGTA AAGCAACAAG 300
 AATACATTTC AAACATGTTA TTTGAAATAA GCATAAAAT TGAGCAAATA GAAATACATG 360
 10 AAGCATGTTA TCTGATATAA TTTGAACATC ATAATAATAA TTAAGGAGGA TTGGCATTTA 420
 TGGCAATCGT AAAAGTAACA GATGCAGATT TTGATTCAAA AGTAGAATCT GGTGTACAAC 480
 TAGTAGATTT TTGGGCAACA TGGTGTGGTC CATGTAAAT GATCGCTCCG GTATTAGAAG 540
 15 AATTAGCAGC TGA CTATGAA GGTAAGCTG ACATTTTAAA ATTAGATGTT GATGAAAATC 600
 CATCAACTGC AGCTAAATAT GAAGTGATGA GTATTCCAAC ATTAATCGTC TTTAAAGACG 660
 GTCAACCACTG TGATAAAGTT GTTGGTTTCC AACCAAAAGA AACTTAGCT GAAGTTT TAG 720
 20 ATAAACATTT ATAAGTTACA ACCAATGACG ACTGGGGCAT TTCTTTAATG AATTGCTCCA 780
 GTTTTTGTTT GTTTTTTAA TATAAAAAGT TGAATGATAA GTCATCATAT TGTTTACGAC 840
 25 TTGAGAATGG TGGGATTAAT AAATCTATGA ACGTTAAATG ATAATCTAGC ATGCTGATAG 900
 ATTTGTAGCA GTTGGTTTGA TAAAACCATG TTCAATATTA CATGATGTGC ATGAAAAGTC 960
 ATACTCGAAG ATGTTGATTA TTA ACTAGAA TTAGTGGTGA TAAATTTGAA GCACTTTTGT 1020
 30 AGCATCATTC ATTTTAAAAT TAGAAGGGGG GATATTTTGG GAAGACTATA AGCAACGAAT 1080
 TAAAAATAAA TTAAATGTCG TACCTATGGA ACCAGGATGC TATTTAATGA AAGATCGTAA 1140
 TGATCAAGTG ATATATGTTG GCAAAGCTAA AAAGCTAAGA AATCGATTGC GATCATATTT 1200
 35 TCACGGGTG 1209

(2) INFORMATION FOR SEQ ID NO: 463:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2410 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 463:

AGTTCAACAC GACGAATTTT ACCTGAGTTT GTTTTTGGTA AGTCGTCAAC GAATTCAATC 60
 50 TCTCTCGGAT ATTTATATGG TGCAACTTCA TTTTAAACAA ATTGTTGTAG TTCTTTAACT 120
 AACGTATCAT CACCCGCACT ATGGTCCTGT AAAATAACGA ATGCTTTAAC AATATTTCTT 180

	GCATCTTCAA CTTCAAAAGG CCCAATCGTA TAGCCTGAAC TAATAATAAT GTCATCTCGA	300
	CGTCCTTCAA ACCAGAAATA ACCATCATCA TCTACATGAG CTAAGTCACC AGTGATGTAG	360
5	TATTACCTG TTTGCGCTTT CGCGGTACGT tCTGGCTCTT tATAATACCC TTTGAAAAGT	420
	GCTGGCAAAT CAAGTGGTAC TGCAATATTC CCTTTCGTAT TAGCAGGTAC GCTATTCCCC	480
10	TCATCATCTA CTACAGTGAC CGAACTACCC GGAATGCCTT TACCCATTGA TCCAATCCTC	540
	TGTGGTGTAT CTTTTAAAA GCCTATAAGC AAGGTACTTT CAGTCTGGCC ATATCCATCT	600
	CTTACAGTTA AATTAAAGTA TTTCTTGAAT TGTTCAACTA CTTCTCGATT TAGTGGCTCA	660
15	CCTGCAGAAA CGGCACTATG TAAATGCGTT AAGTCATAAT CATTTAAGTT CTGTAATTTA	720
	GCCATCATAC GATATTCTGT CGGTGTACAA CATAAACAT TAATTTGATA TTTTGAAGC	780
	AATTCTAAGT ATGTTTCAGG ACTGAACCTT CCATTAAATA CAAAAGCAGT TGCACCTGAA	840
20	CCTAATACAG ATAAGAAAGG ACTCCATACC CATTTTGGCC AACCTGGTGC TGCTGTTGCC	900
	CAAACTAAGT CATCTTCATT aATACATaC CAATGTTTTG GTGCCATTTG TaATGTGcA	960
	AATCCCCaTC CATGACAATG TGTAACGGCT TTAGGATTGC CAGTTGTACC AGATGTATAT	1020
25	GACAGAATCG CCATATCATC ACGCGTCGTA TCTGCCATTT CTAGTTTGTT ACTTGCGTTT	1080
	TCTTTTTCAG CTTCAAGTGA AATCCATCCA TCTTTTGGAC CGGCAATAAC AAATTTAGTT	1140
	AACGCATCAT ATTCTTTAAT TTTTCAAAT TCAACTGTGA ATGGCTCTAG TGCAATAACT	1200
30	GCATTAATTT CACCATGTGT GATACGGTAT TGTAATCTT TAGTTCTTAG CATTTAGAA	1260
	CATGGAATGA TTGCAACACC TAATTTTAAA GCAGCAATAT ATAATTCATA CGTCGCAATA	1320
35	GATCGTGGCA TCATAATGAG TACTTTATCG CCTTTAGATA AACCGTGCGA TGCTAAAACA	1380
	TTACCTACTT TATTAGACTG TTCAATGAGC TGTGGSTAAG TGACTGATAT ATCTTCGCCT	1440
	TCAGTATTAT GATATAAAAT TGCCTTTTTA TCTGGTATGT GGCTATATTT TTCGATTTCC	1500
40	GAAATAATGT TATATTTTTC AGGCGCGAAT AGAGCTGACT TTTGCATAAC TAACTTCCTT	1560
	TCATACATCC ACTTTTCCTG TGATGAACAT TGTAATTTTA TAAATGAATT ATATACATCA	1620
	TACGCCTATC TTTACAGAAT TTTCAATTAA ATAGGGTTAA ATACCAAAGT CCTCGACACT	1680
45	ACACTTTGAC ATGACGTAGC ATTCAAGGAC TTTCAAATGA TTGAGGGTTG ATATCTCGGG	1740
	CTAGACCATA TCAGCTAATT CAATACGAAT ATTGTATGAT AATTCACGAT TAATTATTTT	1800
	TACATCTGCA CCTTTCGAAG TGCCACGATG CTTGTGTGTA TGCTTGTACT CAGCTGAATT	1860
50	TTGCCAATGA TAAAATGCTT GCCTATTTTC CCACAGCGTA ATAATGATAT AGTGTCTACC	1920
	AGCTGTTCTA GGTCTTAAAA ACCTTAATGC TTTAAATCCA TCAACGTTTT TTAAATGCTT	1980

55

ATTTAACACA CATAATGAAT CATTGATAA ATCATTATC GCTTCTAGCA CATCGTAATA 2100
 TGCAGTGTCA TTATTTTTTT GTATTGTGAG ACAATCATCC AGTTCTTCTA TTACATAACT 2160
 5 TCTATATTCA TCATAAATTT TCATAATAAA TGCCTTCATT TCATTATAT TTTTGGTCAT 2220
 ATTACTKTAT ATCTATTACT AAtkCATTCC CGTATTTATT AATTACAATC ATAGTTTGGC 2280
 TyCTTTTTTAA AAGATAAGAC TTTGTAAAAA GTATTAATAT TTCATGCAAA TGGGGGACAG 2340
 10 GAGTCGCCCC CTATTTTTGT GTCTTCAATT TCATGATCAT TATTTAACAT TAGTCATGAA 2400
 AATAGCCGAC 2410

(2) INFORMATION FOR SEQ ID NO: 464:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 590 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 464:

25 TTTATTAATT GTAAAAAATT GAGTAAATTA TCTTTACATT CTAAATTAGT CTTAACTACA 60
 ACTAGTATCC TAATAATTAT AGGAGCTATT ACATTCTTTT TATTAGAACA GTTTAATACT 120
 ATGCAACATA TGGGACTAGT TGAAAAAATC GGAAATTCCT TTTTCCAATC AGTAACAACA 180
 30 CGAACAGCGG GTTTTAACAG TATAGATATA GCAAGCATT rCAAATCTAC CGCATtAaTG 240
 TTAATGCTAC TTATGTTTAT TGGTGGTGCC CCTCTCAGTG CAGCTGGAGG AATTAAAATA 300
 ACTACTTTTG CAGTTGCGTT TATTTTTGTA CTAAATTATA KACGTAAAGA AAATAATGTT 360
 35 TCAGTATTCA ATAAAGAAAT ATCTGACAAA CATATAAAC TATCTATTGT TACCATTAAT 420
 ATCTCATTTT TATTTATCAG CATCATTACT TTTATATTAT CGATAATTAA TCCGAACATA 480
 40 TCATTAATCA AGTTATTATT CGAAGTGGTT TCTGCATTTC GAACAGTAGG GTTAAGTATG 540
 AACCTTACCA CAGAATATCA TGGTATTACT AAAATAATTA TTATATTCGT 590

(2) INFORMATION FOR SEQ ID NO: 465:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 905 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 465:

TATTCGATTT GANTCTTTTA AAATTATGTT TAAAAACATG TCTAATGATT CCGCATCATT 120
 TTTTGCACCA TCAATAAGCG TTTCAGcAAA CCCCTTAATT GAAGTAATAG GTGTTTTTAA 180
 5 TTCATGTGAA ACATTTGCTA CAAATTCACG TCTTAGATTT TCAAGTTGTT TCAGATTTGT 240
 TATATCATGC ATCACAaCTA AAATCCCTTG CAAACTTTTT TGAGACCTAG TTAaAATCGG 300
 AACGCATGAA ATATCAAAGT ACTTGGCATG GACTTGGTTT ATTGCAACTT CCAATTGTTC 360
 10 ATAAATAGGT TTTTCAACTT TAAAACTTTC TAAaATTAAT TGCTCAATTT CAGTATTAAC 420
 ATAGCCGTGA TAGCCTACTT GTTCAATATT ATGCGAGATG TTGAAGTGT CATAATACGC 480
 TTTATTTGCA ACAACGATTT TTCCATTTTG ATCTATCATT AAAATAGCAC TTGGAATATT 540
 15 TTCAATCGTT GTTTTTAAAC GGTGGATTG AATTTTTTGC TCATTATTAA GCTTTTGAAG 600
 GCGTCGTGCT AAATCAATTG TAGACACAAA AAGCGCTTTA GTTTCTACAA CATTACTTTC 660
 20 AGGTACACGT ATGTGATAAT AACCATTGTC CAACAATTGT GTTGcATAAG TAACCTCTTG 720
 AATGGGACCG ATTAATGTAC GCTTAAaACT ACGGCTTGCA AAATACAGAC AAATGAGTAC 780
 AACTAAACAT GTCAAAATAA GATATTTCCA CAACGTCCA TGCATTTCTG TAATATCGTT 840
 25 ATTGTAACCT TTAATCCATA CATGATAACC GTTAACCTTC TTATLAAAAA TAAAAACGTC 900
 CCTTT 905

(2) INFORMATION FOR SEQ ID NO: 466:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1016 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 466:

TTTGGTTAGC CCATAAAGA AAAAAACAAG TAGTCATTTT TAAACAACAT ATCAAGTCTA 60
 40 CCCAAGAAAT ACGTTTTGAC AAAGCGAAAG TGCTTGAACA CAAAGATGAA ATAGCAAATT 120
 TTATTTCTTT CGAACCACAA AGTTTTGAAT TTTATTATTT TACAGAATCG GAATTTTCAG 180
 45 AAGAACAATT AAATGAAGTT TCGCCAATTA GAATTAAATT CAATGTTATA AGACACACAA 240
 AAGATTTGaT AAAGCATATG CCGAATATAT TTTTGGcTAG ACTTATTTCA GAAGATAATG 300
 ATAAAAAGAC ATATATGTTT TATAAACGCA AAGTATTAAC CGATAACTTT TTAGATAAAT 360
 50 ATATGCAGAA ATTTTCACCG GCAACATACA CAATAATATT TGTAAATGTC TTAATATGGT 420
 TATGTATGAT TTTATATTTA AATAATTTTT CGGATGTAAA ATTATTAGAT GTTGGCGGGT 480

55

ATTTTAGTTT TGAACATATA CTTATGAATA TGCTTTCATT ATTTATTTTT GGTAAAATAG 600
 TCGAAGCAAT TATTGGTTCA TGGCGGATGT TAACTGTATA CTTTATTGCA GGGTTGTTTG 660
 5 GAAACTTTGT ATCACTATCA TTTAATACGA CTACAATTC AGTTGGGGCT AGTGGTGCTA 720
 TATTTGGTCT GATTGGATCA ATTTTTCGCA TGATGTATGT TTCAAAAACA TTAAACAAAA 780
 AAATGTTAGG ACAGTTATTA ATTGCATTAG TGATATTAGT TGGTGTCTTCT CTGTTTATGT 840
 10 CAAATATAAA TATTGTGGCG CATATTGGAG GATTCAATGG TGGTTTATTA ATAACTTTAA 900
 TTGGCTATTA CTATAAAGTG AATCGtAATA TTTTTTGGat TTactAATTG GTATGCTtGT 960
 tATATTTAwT GCACyTCMAA TTagAtTTTT ACmATTAAAG AAGATAATAw TTATAA 1016
 15

(2) INFORMATION FOR SEQ ID NO: 467:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 406 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 467:

AACTTTAAAT TTAGACATCT TTAAACCTC TCTTAAACCA TGCCTATATC TCAAGATGAT 60
 ATTTCAAATG AACAACTACTA TTGCTTGAGA CCATTAAATGA ATGATCATAA ATATTTCTTT 120
 30 CTATAAAATT AGCTTTCCAA TAACTGTGTT GTTGCAATAT ATCATTCACA AGTACACCAT 180
 TTTCGGAAGT ATGATTATCT TTATCTATAC TTAAACAAT TTGTTTAGTT TTAGCATGGC 240
 TAAATTGTTG AAGACCCTTA CACGATAAAC GTATAGCGTC TGAATTCTCA TTAAACAATG 300
 35 CGGCTGGgCA AACAAATGAC ACATTGTACT TCATGTTTGA ACtTCGTTAC AATCaTCGTG 360
 kCatTTTGat AAATAACAAT CCCTCGTAAT kGATTAAGTA TATTAT 406

(2) INFORMATION FOR SEQ ID NO: 468:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1378 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 468:

AGATAATACA CTTGAAGTTG GAATGGTTTG TGACGGTTAT TTAATGCGAA TTGAAAACCTT 60
 AACACCATCA AATTTCTTCA ACTCAGCAAG TGAAGATACG ATTACTAAAA TTAAATTAAA 120

AGGTACAGCG TTAAACTAA GAGAAGCCAT CAATTATGAT GAAATGGTTA TTGTAGATAG 240
 TATGACGTAG TTCCTAATTA TGCKAAAAGG GATTGATGAA AACTGAAGG GCTTTTCATC 300
 5 AATCCCTTTT ATTTTAGGGG AATTGAATAG ATAGTTTAA ACTATACGAA TTATTAATAT 360
 TTGAGATTTA ATTGAAATAA GTTTTAAAAA TTGGAGGAGA TAGATTAAGC GAAGTCATTT 420
 AAAGGTGAAG TTAAGTGTAT TCACAAAAAa TAGCCACACT CATATGACAT CGGATGAGTG 480
 10 TGGCTTAAGG ATCTATGGGG GGAGGAAnCC ATAGATGTTT ACTTTGATAG GCCAGATTAA 540
 ATATCAAAGT ATGCGATTAT TTATAGCTTG ATGCAAAAGT GGTATGCCTA TTAAAAGTTA 600
 CTGCACATAG CTTTTAATAT TCCGTTCAAA GGAAAGGGGC ATACAATTGA ACAATCTGTA 660
 15 ATAGTACTTT TAACCAGCTA TGCTAAAAGT CTAGTAGGGA GAACAGTTGT CCAATCACAT 720
 AAGAACCTCT AACTTCGTTA GTACGATTAA GAAAAGCTTT TTAGTTAGTA TGTAATACAA 780
 20 TTTATTGACG CGCGTGAATC TCTTTTATAA GAGTGTGTAG GGAATGGCGT TGTATAAATT 840
 GTATTAGAAG AACTTCTAAC GCATCTCTGT GGTAAAAGA GATGAAGGGA ACGACAGTTT 900
 aATTAAACT GCATAAGAAC TTCTAGCTTT TCTCTCTCGT TCAAAGAGAA GCAGcTGTTc 960
 25 GCAGTTTAAT CAAAACCACA TAAAGCTTTT AACTTTACTC TTTGATTAA AGAGTGATAA 1020
 ATGTTTACAG TTTAATTAAA ACTGCATAAG AACTTCTAGC TTTTCTCTT CGTTCAaGAG 1080
 AAGCAGCTGT TCGCagTTA ATCAAACCA CATAAGCTT TTAACTTTAC TCTTTGATTT 1140
 30 AAAGAGTGAC AAATGTTTAC AGTTTAATTA AACTGCATA AGAACTTCTA GCTTTTCTCT 1200
 TTCGTTCAAA GAGAAGTTCT AATACCACCA TATCGTGCga TCGGGAACGG TATATATATT 1260
 AATAGGAGGG TAATATATAT TTAACGCACG ATATGGGACT ATTAGCCTTC GACTTTGTTA 1320
 35 TGTTGATGTG TGGCCTAAAA TATTGGAGAT ACCAATATTT TAGGTTGCAT CAACATCA 1378

(2) INFORMATION FOR SEQ ID NO: 469:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4171 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 469:

TCCCAACCAA TAATCGTGGC AAAAATACGG ATATTGGTAT GGCTTAACAA ATTGCAAATA 60
 50 TCGTTTAATC ATACATCCCC CCTAATCTAT TGCCCTATCC TATTCaTAAG CATAAAAATG 120
 AATAGAGGTT GGATACATAA TTTGTAGATG TAAATCTTC TTACAATTTA CATTTTTAAA 180

55

	TTATATTCTA TTCAATTTAA TCTATGGATA CTGTGTCCCC ACACGACAGC AAAAGTTATC	300
	ATACTTCTTT ACATCACTAA GTCAATATAA ATGATTTAAT CAGTATTTAC ACTTTATTTG	360
5	CTTAATACTG TCTAATTTTT TTGTAACGTT CTTTCCAAAC TTTGATAAAA TCTGGCGCGA	420
	ATGGGCCCTT CTCTGTCTT ATCCATTGTT GAAGAATGTC CACGTTGCGT CTTAAAATAA	480
	TATCAATATC ATGCGGATAA TTCATTTGAT TCATATGTTG CTCATATTCA TCTTCATCTA	540
10	ATAAATGATA CTTTCCGTTT GGATATACTT TAATATCTAA ATCATAGTCT ATATATTTTA	600
	ATGCCTCTTC ATCACAACA AATGGTGATG ACAAATTGCA ATAGTAATAA ATTCCATCTT	660
15	CTCTAAACAT GCAGATAACA TTAAACCAAT ATTCTGAGTG AAAGTAAACA ATTGCCGGTT	720
	CACGTGTTAT CCAAGTTCTT CCGTCACTTT CAGTCACTAA CGTATGATCA TTTCCACCAA	780
	TGACAACATG ATCAGTACCC TTTAATATTG TGTTTTCAGA CCAAACGCGA TGAATCTTAC	840
20	CATCATGTTT ATAACCTGA ATTTAATGT TTTCCCTTC TTTAGGTATG GATTCTCTGA	900
	CCATACTCCA CACCACCTTC TGTAAATTTA ACCATTATAA ATTATAGCAT ATTTCAGAAA	960
	TAGTATTATA TAAATACATA TTTTACGAA ATAAGATTTT ACTACTTAAT AATTAACTC	1020
25	GGTAATATTG CTAAGTACTA CAACAGAGAT TTACATGTCC CATTTAAAGT ATATAAAATC	1080
	ATCACTTTTA TATATCAACA CTTTAACTTT TTGACATTGT TATTCTATGA GATTTAAAGA	1140
	TATCATTTAT ACTTTTTAAA ATTAATGTCA CTATGTTTTT CGATAATATT ACCAATCATC	1200
30	GAATGTTACC CATTTATAAA TTGATAAATs TTTGACATAG GTACAGGGAA TGTATATTGA	1260
	TCTCGATCAC TTAAATCAAA CCAATCATG TCATCTGGTA ATGTTTCAAT GTTAATTGCT	1320
	CCTGAAACGG CGTATACTTT AATCTTCCAT GTTAAATGAG TAAATTGATG CTTCAACTCA	1380
35	AAAATAGGTG TTTCTACTGG TTGAATGTCA TGACCGATTT TTTCAAGTCAT TTTACGTCTA	1440
	GCATGCTCAC TTTCAAACAT AGGAAATTGC CACATACCAT GCAATAATTT TTCGCTACGC	1500
40	TTTTGCAACA GATATTGACC TTGATTATTT CTAATTAAAA AGACGGATTG CTCAATTACT	1560
	TTTTTACTTA CATTTTTAGA TTTAACAGGT AACTTTTCAA ATGTACCTTT ATCAAATGCC	1620
	TCACAGTTTT CTTGAACTGG ACAAATAAG CATAATGGAT TTTTGGTGT ACAAATTAAC	1680
45	GCCCCTAATT CCATCATAGC TTGATTAAAC GTTCCAGCTT CTGTAGTAAC ATACGGTAAT	1740
	AATCTTGTT CGTACGATTT CCTCGTCGAT TGTAATTTAA TATCTCGATA GTCATCATTC	1800
	AATCTAGACC ATACTCGAAA AACATTTCCG TCTACAGTTG CTAGTGGTAC ATTATATGCA	1860
50	ATGCTCATTA CTGCAGCTTG TGTGTATGGG CCAACACCTT TTAACGCTTT AAATTGATCA	1920
	GGATCTTTGG GAACTAAGCC TTCATATTTA TCATGAACCT CTTTAATCGC CGTATGAAAA	1980

55

	GCTTGaCTCA	AAACTTCCAC	AGTTGGAAAT	CGTTCAACAA	AACGATGATA	ATAGTCAATA	2100
	ACTGTTTTAA	CTTGTGTCTG	TTGTAACATG	ACCTCACTTA	ACCAAATATA	GTACGGATTG	2160
5	GTCGTTTGTC	GCCATGGCAT	TTCTCTTTGA	TTTTCAACAA	ACCAGTGTAT	CAAATTTTCT	2220
	TTAAAACTAG	ACTGCTGATA	CATTTATAAA	ACCCTTTCCT	CACCAAAATT	AATTGTCITT	2280
	ACTCATAATG	TTTTTATTGT	ACATTAAAAT	CATGGTTAGT	ATGTAAGTTA	ATTTAGTTAT	2340
10	TTGCGAAATT	GGATTATAAT	AGTATATATA	ATATTATGAA	ATGAGTGAAC	TGATATGGAC	2400
	ACTGCAACAC	ATATCGCAAT	TGGGGTGGGC	CTTACAGCAC	TTGCAACTCA	AGATCCAGCA	2460
15	ATGGCTTCTA	CGTTTGGTGC	AACAGCTACA	ACCCTTATCG	TTGGTTCATT	AATTCCTGAT	2520
	GGGGATACTG	TTCTTAAATT	AAAGGACAAT	GCAACATATA	TTTCGCATCA	TAGAGGTATC	2580
	ACGCATTCCA	TCCCTTTCAC	AATACTATGG	CCAATTTTAA	TTACATTTTT	AATATTCACG	2640
20	TTCTTTAGTG	GAACCAACCC	ATTTTCATGTA	TGGATGTGGG	CTCAGCTCGC	AGTATTTTTA	2700
	CATGTCTTTG	TAGATATATT	CAATTCCTAT	GGTACACAAG	CGCTTAGACC	TATCACAAC	2760
	AAATGGATTG	AATTAAGTGT	GATTAACACA	TTTGACCCTA	TTATTTTCAC	AGTTCTTTGT	2820
25	ATTGGTATTG	TATTATGGGT	TATAGGCTTG	CATCCATTG	CAGTCTTCTT	TCCTATAATC	2880
	GCTTTACTAA	TCATTTATTA	CATGATTCGT	TTTAAATGA	GAGCCGTAAT	TAAGCAACAA	2940
	GCTTTAAAG	CAATTCAACA	AGAGCATCAC	CCTGTTAAAG	TATTTGTTGC	GCCAACAATA	3000
30	AAATTTATGG	AATGGCGTGT	CGCGATACAA	ACTGATGCAC	ATGACTATGT	TGGAAAAGCA	3060
	TATGGTAGGA	ATGTGGTGT	TAGTGATAAA	GTGGAACGTC	AAACATTATC	AACAGACTCC	3120
	ATTTTATGGA	AAGTCAAAGG	TAATAAAGAT	ATACGTACTT	TTTTAACTT	TTTATCAATC	3180
35	TATCGTTGGC	AAACAACAAC	GTTAGCAGAT	GGTCTACTG	AAATTCGTTT	GATTGATTTG	3240
	CGTTATTTAA	AAAATGATCA	TTATTCATTT	GTGGCAATTG	CACATGTAAC	AAACGATAAT	3300
	GTCATAGACC	ACTCTTATAT	TGGCTGGGTA	TTTACAGAAG	ATAAGTTACA	ACGTAAACTG	3360
40	TATGCTAAAT	AATTTCAAGT	TATTATTCAC	TAAAGTTAAT	CTATAAAAAA	TGAACAACCG	3420
	GGCAGAAATG	AAATCAAAAC	GATTTTTACT	CTGTCCGTTT	TTTAAATGTA	AAACTATGAA	3480
45	TGCTTTTACA	AAATCTAAAA	TTTATATTGT	TGCTAACAAA	CTACCTTTAA	TGACTCGAAA	3540
	TATCAAAATC	AGTATAGGAA	AACAATATCT	AGATGATATT	CTAATTGTTT	CTGATTCTCA	3600
	CAGATTAATT	TACACAACAG	GTCAGCTAAA	CATCATGAAG	AAGTATCCGC	CTCGTCTGTA	3660
50	CTATCATTTG	AAACATCCTG	TTGATTATCa	GTTTGTGTCA	CTTTAGATGT	TTTATTaTAA	3720
	ATTGCGTGTG	TCGTATACCT	TGCTAATACG	AAATTTACTG	CAGCTATTAA	ACATAAGAAT	3780

55

TGTGCTATAC CATTAACAAT GTAATACATT GGATTTAGCA TTAGGATGTG ATTGATAAAT 3900
 ACATGATTTG GATTTGGTAT GAAAAATAATT GGTAACAATA AGAAACACAA TACACAAACC 3960
 5 CCATAAAATA TGATATTTAT TTTTTCAGWT AACAGTCGAA TAAGACCAAA AGTAACGGAT 4020
 ATTAATCCTA CAAAAATAGT TGCCATCACA ATAAAATAGA ATAGCGCTAT ATATGATGTT 4080
 TCGAAGTTTA CTGGTTTAAAC CAATGCACTA ATCATCGTCA AAATGACTAG CATAATAAAA 4140
 10 CTTAAATAG ACATAATAAC TACTGGCGTC G 4171

(2) INFORMATION FOR SEQ ID NO: 470:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9821 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 470:

TGGTTGAAGT AGCAGTTAAT TCAAAATCTG CAACAGTTTC AGCAGAATaG GGGCTTTCAA 60
 25 AATAATCAAA GGAGAATAAT TTATGACTAA AACTTTAAAG GTTTATAAAG GAGACGACGT 120
 CGTAGCTTCT GAACAAGGTG AAGGCAAAGT GTCAGTAACT TTATCTAATT TAGAAGCGGA 180
 TACAACTTAT CCAAAGGTA CTTACCAAGT GGCATGGGAA GAAAATGGTA AAGAATCTAG 240
 30 TAAAGTTGAT GTACCTCAAT TCAAAACCAA TCCAATTCTA GTCTCAGGCG TATCATTTAC 300
 ACCAGAACT AAATCAATTA TGGTAAATAC CGATGACAAT GTTGAGCCAA ACATTGCACC 360
 AAGCACAGCA ACGAATAAAA TATTGAAATA TACAAGTGAA CATCCAGAAT TTGTTACTGT 420
 35 AGATGAAAAT ACAGGAGCAA TTCACGGTGT AGCTGAAGGT ACTTCAGTAA TCACTGCTAC 480
 GTCTACTGAT GGAAGCGATA AGTCAGGACA AATTTAGTG ACAGTAACAA ACGGATAGGG 540
 ATTTAAGGCG CAGTATATCT GCGTCTTTTT TATTTGAATA AAAGGAGCTA ATACAATGAT 600
 40 TAAATTTGAA ATTAAAGATC GTAAACAGG AAAACAGAG AGCTATACAA AAGAAGATGT 660
 AACAATGGGC GAACAGAAAA ATGCTATGAG TATTTAGAAT TAGTAAATCA AGAGAATAAA 720
 45 AAAGAAGCAC CTAACGCAAC AAAAATGAGA CAAAAGAGC GACAGTTATT AGTAGATTTA 780
 TTTAAAGATG AAGGATTGAC TGAAGAAGAT GTTCTGAACA AGATGAGTAC TAAAACCTAT 840
 ACAAAGCCT TACAAGATAT ATTTGAGAA ATCAATGGTG AAGATGAAGA AGATTCAGAA 900
 50 ACTGAACCAG AAGAGATGGG AAAGACAGAA GAACAATCTC AATAAAGAC ATTTTATCGA 960
 ACATTAAGAA AATACAACGT TTCTGTATGG AGCAGTATGG GTGGACATTA ACTGAAGTCA 1020

	AAGAAAAACA AAGTGAACAA AAAGTCATTA CAGGTACGGA TTAAAGAAAA CTTTTTGGAA	1140
	GCTAGAAAGG AGGTTAATAT GAATGAAAAA GTAGAAGGCA TGACCTTGGA GCTGAAATTA	1200
5	GACCATTTAG GTGTCCAAGA AGGCATGAAA GGTTTAAAGC GACAATTAGG TGTGTGTAAT	1260
	AGTGAAATGA AAGCTAATCT GTCAGCATTT GATAAGTCTG AAAAATCAAT GGAAAAATAT	1320
10	CAGGCGAGAA TTAAGGGGTT AAATGATAGG CTTAAAGTTC AAAAAAGAT GTATTCTCAA	1380
	GTAGAAGATG AGCTTAAACA AGTTAACGCT AATTACCAAA AAGCTAAATC CAGTGTAATA	1440
	GATGTTGAGA AAGCATATTT AAAGTTAGTA GAAGCCAATA AAAAAGAAAA ATTAGCTCTT	1500
15	GATAATCTA AAGAAGCCTT AAAATCATCG AATACAGAAC TTAAAAAGC TGAAAAATCAA	1560
	TATAAACGTA CAAATCAACG TAAACAAGAT GCGTATCAAA AACTTAAACA GTTGAGAGAT	1620
	GCAGAACAAA AGCTTAAGAA TAGTAACCAA GCTACTACTG CACAATAAA AAGAGCAAGT	1680
20	GACGCTACA GAAGCAGTCC GCTAAGCATA AAGCACTTGT TGAACAATAT AAACAAGAAG	1740
	GCAATCAAGT TCAAAAATA AAAGTGCAAA ATGACAATCT TTCAAAATCA AATGATAAAA	1800
	TTGAAAGTTC TTACGCTAAA ACTAATACTA AATTAAAGCA AACAGAAAA GAATTTAATG	1860
25	ATTTAAACAA TACTATTAAG AATCATAGCG CTAATGTGCG AAAAGCTGAA ACAGCTGTTA	1920
	ATAAGAAAA AGCTGCTTTA AATAATTTGG AGCGTTCAAT AGATAAAGCT TCATCCGAAA	1980
	TGAAGACTTT TAACAAAGAA CAAATGATAG CTCAAAGTCA TTTGGGTAAA CTTGCAAGTC	2040
30	AAGCGGATGT CATGTCAAAG AAATTTAGTT CTATTGGAGA CAAATGACT TCCCTGGGAC	2100
	GTACAATGAC GATGGGCGTA TCTACACCGA TTACTTTAGG TTAGGTGCA GCATTAAAAA	2160
	CGAGTGCAGA CTTTGAAGGG CAAATGTCTC GAGTTGGAGC GATTGCACAA GCAAGCAGTA	2220
35	AAGACTTAAA AAGCATGTCT AATCAAGCGG TTGACTTAGG AGCTAAAACA AGTAAAGTG	2280
	CTAACGaAGT TGCTAAAGGT ATGGAAGAAT TGGCAGCTTT AGGCTTTAAT GCCAAACAAA	2340
40	CAATGGAGGC TATGCCAGGT GTTATCAGCG CAGcAGaAGC AAGTGGTGCA GAAATGGCTA	2400
	CAACTGCAAC TGTAATGGCT TCAGCGATTA ACTCTTTCGG TTTAAAGCA TCTGATGCAA	2460
	ATCATGTTGC TGATTFACTT GCGAGATCAG CAAATGATAG TGCTGCAGAT ATTCAATATA	2520
45	TGGGAGATGC ATTAAATAT GCAGGTACTC CAGCAAAAGC ATTAGGAGTT TCAATAGAGG	2580
	ACACTTCTGC AGCAATTGAA GTTTTATCTA ACTCAGGTTT AGAGGGGTCT CAAGCAGGTA	2640
	CTGCATTAAG AGCTTCGTTT ATTAGGCTAG CTAATCCAAG TAAAAGTACA GCTAAGGAAA	2700
50	TGAAAAAATT AGGTATTCAT TTGTCTGATG CTAAGGTGA GTTTGTTGGA ATGGGCGAAT	2760
	TGATTAGACA GTTCCAAGAT AACATGAAAG GCATGACGAG AGAACAAAA TTAGCAACAG	2820

55

	CAGATAAAAT TAATAGCTAT AGCAAATCAT TGAAGAACTC TAATGGTGAA AGTAAAAAAG	2940
	CAGCTGATTT GATGAAAGAT AACCTCAAAG GTGCTCTGGA ACAATTAGGT GCGCCTTTTG	3000
5	AATCGTTAGC AATTGAAGTT GGTAAAGATT TAACGCCTAT GATTAGAGCA GGTGCGGAAG	3060
	GATTAACAAA ATTAGTTGAT GGATTTACAC ATCTTCCTGG TTGGGTTAGA AAGGCTTCGG	3120
	TAGGCTTAGC AATTTTTGGT GCATCTATTG GTCCTGCTGT TCTTGCTGGT GGCTTATTAA	3180
10	TACGTGCAGT TGGGAGCGCG GCTAAAGGCT ATGCATCATT AAATAGACGC ATTGCTGAAA	3240
	ATACAATTCT TTCTAATACC AATTCAAAAG CAATGAAATC TTTAGGTCTT CAAACATTAT	3300
	TTCTTGTTTC TACAACAGGA AAAACGTCAA AAGGCTTTAA AGGATTAGCC GGAGCTATGT	3360
15	TGTTTAATTT AAAACCTATA AATGTTTTGA AAAATTCTGC AAAGCTAGCA ATTTTACCGT	3420
	TCAAACTTTT GAAAAACGGT TTAGGATTAG CCGCAAAATC CTTATTTGCA GTAAGTGGAG	3480
20	GCGCAAGATT TGCTGGTGTA GCCTTAAAGT TTTTAACAGG ACCTATAGGT GCTACAATAA	3540
	CTGCTATTAC AATTGCATAT AAAGTTTTTA AAACCGCATA TGATCGTGTG GAATGGTTCA	3600
	GAAACGGTAT TAACGGTTTA GGAGAACTA TAAAGTTTTT TGGTGGCAAA ATTATTGGCG	3660
25	GTGCTGTTAG GAAGCTAGGA GAGTTTAAAA ATTATCTTGG AAGTATAGGC AAAAGCTTCA	3720
	AAGAAAAGTT TTCAAAGGAT ATGAAAGATG GTTATAAATC TTTGAGTGAC GATGACCTTC	3780
	TGAAAGTAGG AGTCAACAAG TTTAAAGGAT TTATGCAAAC CATGGGCACA GCTTCTAAAA	3840
30	AAGCATCTGA TACTGTAAAA GTGTTGGGGA AAGGTGTTTC AAAAGAAACA GAAAAAGCTT	3900
	TAGAAAAATA CGTACACTAT TCTGAAGAGA ACAACAGAAT CATGGAAAAA GTACGTTTAA	3960
	ACTCGGGTCA AATAACAGAA GACAAAGCAA AAAAAGTTTT GAAAATTGAA GCGGATTTAT	4020
35	CTAATAACCT TATAGCTGAA ATAGAAAAAA GAAATAAAAA GGAAGTCTGAA AAAAGTCAAG	4080
	AACTTATTGA TAAGTATAGT GCATTCGATG AACAAGAAAA GCAAAACATT TTAAGTAGAA	4140
	CTAAAGAAAA AAATGACTTG CGAATTAAAA AAGAGCAAGA ACTCAATCAG AAAATCAAAG	4200
40	AATTGAAAGA AAAAGCTTTA AGTGATGGTC AGATTTCAGA AAATGAAAGA AAAGAAATTG	4260
	AAAAGCTTGA AAATCAAAGA CGTGACATCA CTGTTAAAGA ATTGAGTAAG ACTGAAAAAG	4320
45	AGCAAGAGCG TATTTTAGTA AGAATGCAAA GAAACAGAAA TGCTTATTCA ATAGACGAAG	4380
	CGAGCAAAGC AATTAAAGAA GCAGAAAAAG CAAGAAAAGC AAGAAAAAAA GAAGTGGATA	4440
	AGCAGTATGA AGATGATGTC ATTGCTATAA AAAATAACGT CAACCTTTCT AAGTCTGAAA	4500
50	AAGATAAATT GTTAGCTATT GCTGATCAAA GACATAAGGA TGAAGTAAGA AAGGCAAAAT	4560
	CTAAAAAAGA TGCTGTAGTA GACGTTGTTA AAAAGCAAAA TAAAGATATT GATAAAGAAA	4620

55

GTTGGTGGTC TAACTTTAGA GAAGACCAAA AGAAGAAAAG TGATAAATAC GCTAAAGAAC 4740
 AAGAAGAAAC AGCTCGTAGA AACAGAGAAA ATATAAAGAA ATGGTTTGGG AATGCTTGGG 4800
 5 ACGGCGTAAA AACTAAAAC TGTGAAGCCT TTAGTAAAAT GGGCAGAAAT GCTAATCATT 4860
 TTGGCGGCGA AATGAAAAA ATGTGGAGTG GAATCAAAGG AATTCCAAGC AAATTAAGTT 4920
 CAAGTTGGAG CTCAGCCAAA AGTTCTGTAG GATATCACAC TAAGGCTATA GCTAATAGTA 4980
 10 CTGGTAAATG GTTTGGAAAA GCTTGGCAAT CTGTTAAATC GACTACAGGA AGTATTTACA 5040
 ATCAAATAA GCAAAAGTAT TCAGATGCCT CAGATAAAGC TTGGGCGCAT TCAAAATCTA 5100
 TTTGGAGAGG CACATCAAAA TGGTTTAGCA ACGCATATAA AAGTGCAAAG GGTGGCTAA 5160
 15 TAGATATGGC TAATAAATCG CGCTCGAAAT GGGATAATAT TTCTAGTACA GCATGGTCGA 5220
 ATGCAAAATC CGTTTGGAAA GGAACATCGA AATGGTTTAG TAACTCATAC AAATCTTTAA 5280
 20 AAGGTTGGAC TGGGGATATG TATTCAAGAG CCCACGATCG TTTTGATGCA ATTTCAAGTT 5340
 CGGCATGGTC TAACGCTAAA TCAGTATTTA ATGGTTTTAG AAAATGGCTA TCAAAAACAT 5400
 ATGATTGGAT TAGAGATATT GGTAAAGACA TGGGAAGAGC TGCGGCTGAT TTAGGTAAAA 5460
 25 ATGTTGCTAA TAAAGCTATT GCGGTTTGA ATAGCATGAT TGGCGGTATT AATAAAATAT 5520
 CTAAAGCCAT TACTGATAAA AATCTCATCA AGCCAATACC TACATTGTCT ACTGGTACTT 5580
 TAGCAGGAAA GGGTGTAGCT ACCGATAATT CAGGAGCATT AACGCAACCG ACATTTGCTG 5640
 30 TATTAAATGA TAGAGGTTCT GGAAACGCCC CAGGTGGTGG AGTTCAAGAA ATAATTCACA 5700
 GGGCTGACGG AACATTCCAT GCACCCCAAG GACGAGATGT GGTGTTCCA CTAGGAGTTG 5760
 GAGATAGTGT AATAAATGCC AATGACACTC TGAAGTTACA GCGGATGGGT GTTTTGCCAA 5820
 35 AATTCCATGG TGGTACGAAA AAGAAAAAAT GGATGGAACA AGTTACTGAA AATCTTGGA 5880
 AAAAAGCAGG GGAATTCGGT TCTAAAGCTA AAAACACAGC TCATAATATC AAAAAAGGTG 5940
 CAGAAGAAAT GGTGAAGCG GCAGGCGATA AAATCAAAGA TGGTGCATCT TGGTTAGGCG 6000
 40 ATAAATCGG CGATGTGTGG GATTATGTAC AACATCCAGG GAAACTAGTA AATAAAGTAA 6060
 TGTGAGGTTT AAATATTAAT TTTGGAGGCG GACTAACGCT ACAGTAAAAA TTGCTAAAGG 6120
 CGCGTACTCA TTGCTCAAAA AGAAATTAGT AGACAAAGTA AAATCGTGGT TTGAAGATTT 6180
 45 TGGTGGCGGA GGCGATGGAA GCTATCTATT TGACCATCCA ATTTGGCAA GGTTTGGGAG 6240
 TTACACAGGT GGAATTAAT TTAATGGCGG TCGTCACTAT GGTATCGACT TTGGTATGCC 6300
 50 TACAGGAACG AACATTTATG CTGTTAAAGG CGGTATAGCT GATAAAGTAT GGAATGATTA 6360
 CGGTGGCGGT AATTCTATAC AAATTAAGAC CGGTGCTAAC GAATGGAAT GGTATATGCA 6420

55

	ATCAGGTGCT ACAGGTAATT TCGTTAGAGG AGCACACTTA CATTTCCAAT TGATGCAAGG	6540
	GTCGCATCCA GGGAAATGATA CAGCTAAAGA TCCAGAAAAA TGGTTGAAGT CACTTAAAGG	6600
5	TAGTGGCGTT CGAAGTGGTT CAGGTGTTAA TAAGGCTGCA TCTGCTTGGG CAGGCGATAT	6660
	ACGTCGTGCA GCAAAACGAA TGGGTGTTAA TGTTACTTCG GCTGACGTAG GAAATATCAT	6720
	TAGCTTGATT CAACACGAAT CAGGAGGAAA TGCAGGTATA ACTCAATCTA GTTCGCTTAG	6780
10	AGACATCAAC GTTTTACAGG GCAATCCAGC AAAAGGATTG CTTCAATATA TCCCACAAAC	6840
	ATTTAGACAT TATGCTGTTA GAGGTCACAA CAATATATAT AGTGGTTACG ATCAGTTATT	6900
15	AGCGTTCCTT AACAAACAGAT ATTGGCGCTC ACAGTTTAAAC CCAAGAGGTG GTTGGTCTCC	6960
	AAGTGGTCCA AGAAGATATG CGAATGGTGG TTTGATTACA AAGCATCAAC TTGCTGAAGT	7020
	GGGTGAAGGA GATAAACAGG AGATGGTTAT CCCTTTAACT AGACGTAAAC GAGCAATTCA	7080
20	ATTAAGTGAA CAGGTTATGC GCATCATCGG TATGGATGGC AAGCCAAATA ACATCACTGT	7140
	AAATAATGAT ACTTCTACAG TTGAAAAATT GTTGAAACAA ATTGTTATGT TAAGTGATAA	7200
	AGGAAATAAA TTAACAGATG CGTTGATTCA AACTGTTTCT TCTCAGGATA ATAACCTAGG	7260
25	TTCTAATGAT GCAATTAGAG GTTTAGAAAA AATATTGTCA AAACAAAGTG GGCATAGAGC	7320
	AAATGCAAAT AATTATATGG GAGGTTTGAC TAATTAATGC AATCTTTTGT AAAAATCATA	7380
	GATGGTTACA AGGAAGAAGT AATAACAGAT TTTAATCAGC TTATATTTTT AGATGCAAGG	7440
30	GCTGAAAGTC CAAACACCAA TGATAACAGT GTAACATTA ACGGAGTAGA TGGTATTTTA	7500
	CCGGGCGCAA TTAGTTTTGC GCCTTTTTCA TTAGTAITAA GGTGTTGGCTA TGATGGTATA	7560
35	GATGTTATAG ATTTAAATTT ATTTGAGCAT TGGTTTAGAT CTGTGTTTAA TCGCAGACAT	7620
	CCTTATTATG TTATTACTTC TCAAATGCCT GGTGTTAAAT ATGCAGTGAA TACAGCTAAT	7680
	GTTACATCTA ATTTAAAGA TGGTTCTTCA ACTGAAATTG AAGTAAGTTT AAATGTTTAT	7740
40	AAAGGGTATT CTGAATCAGT TAATTGGACC GATAGCGAGT TCTTATTCGA CTCTAATTGG	7800
	ATGTTTGAAA ATGGAATTCC TCTTGATTTC ACACCTAAAT ATACTCATAC ATCAAATCAA	7860
	TTTACTATTT GGAACGGTTC TACTGATACG ATAAATCCAC GATTCAAGCA CGATTTGAAA	7920
45	ATATTAAITTA ATTTAAATGC GAGTGGAGGA TTTGAACTGG TTAATTATAC AACAGGTGAT	7980
	ATTTTTAAGT ACAACAAAAG TATAGATAAA AACACTGATT TTGTTTTAGA TGGTGTGTAT	8040
	GCATATCGAG ATATAAACAG AGTGGGAATT GATACAAATA GAGGCATTAT AACATTAGCG	8100
50	CCAGGTAAAA ATGAATTTAA GATTaAAGGA GACGTCAGTG ATATTAAAC TACATTTAAG	8160
	TTTCCTTTTA TTTATAGGTA GGTGATTTAA TGGATTATCA TGATCATTTA TCAGTAATGG	8220

55

	ATTATGAACT GAATGAAGCT AGGTACATCA CCTTTACAGT TTATAGAACT ACTCATAATA	8340
	GTTTTGTTTT TGATTATTG ATTTGTGAAA ACTTCATAAT TTATCATGGT GAAAAATACA	8400
5	CAATTAAGCA GACAGCGCCA AAGGTTGAAG GTGATAAAGT TTTTATTGAA GTTACGGCAT	8460
	ATCACATAAT GTATGAATTT CAAAATCACT CAGTGGAAATC AAATAAGCTT GATGACGACA	8520
	GTAGCGAAAC TGGTAAAACG CCAGAATACT CTTTAGATGA GTACTTAAGA TATGGATTTG	8580
10	CAAATCAAAA AACGTCAGTC AAGATGACCT ATAAAAATAAT TGGAGATTTT AAAAGAAAAA	8640
	TACCAATTGA TGAATTAGGT AATAAAAAATG GCTTAGAATA TTGTAAAGAA GCAGTAGATT	8700
	TGTTTGGTTG TATTATTTAT CCAAATGATA CGGAGATATG TTTTATTCTT CCTGAAACAT	8760
15	TCTATCAAAG AAGCGAAAAA GTAATAAGGT ATCAATATAA TACTGATACT GTGTCTGCTA	8820
	CTGTCAGTAC GTTGAATTA AGAACAGCTA TAAAAGTTTT TGGGAAAAAG TACACAGCCG	8880
20	AGGAAAAGAA AAATTATAAT CCTATTAGAA CAACTGACAT TAAATATTCA AATGGTTTTA	8940
	TAAAAGAAGG TACTTATCGT ACCGCAACAA TTGGGTCTAA AGCTACTATT AACTTTGATT	9000
	GCAAGTATGG TAATGAAACA GTTAGATTTA CAATAAAaAA GGGCTCTCaA GGTGGAATAT	9060
25	ATAAGTTGAT TTTAGACGGC AAGCaAATTA AGCaAATTTT TGTTTTGCT AAGTCGGTTC	9120
	AGTCTGAmAC AATAGATTTA ATaAAAAATA TTGATAAAGG CAAGCACGTT TTAGAAATGA	9180
	TATTTTTtGG AGArGrCCCC AAAAATAGAA TTGATATATC TTCAAATAAA AAAGCTAAGC	9240
30	CTTGATATGTA TGTTGGAAC TAAAAATCAA CAGTCTTAAA TTTAATTGCT GACAACTCAG	9300
	GTCGCAATCA ATACAAAGCA ATTGTTGaCT ACGTCGCAGA TAGTGCAAAG CAGTTTGGGA	9360
	TTCGATATGC TAATACGCAA ACAAATGAAG ATATCGAAAC ACAGGATAAG CTGTTAGAAT	9420
35	TTGCAAAAAA GCAAATAAAT GATACTCCTA AGACTGAATT AGATGTTAAT TATATAGGTT	9480
	ATGAAAAAAT AGAGCCAAGA GATAGCGTAT TCTTTGTTCA TGAATTAATG GGATATAACA	9540
40	CTGAATTAaA GGTGTGTTAA CTTGATAGGT CACATCCATT TGTAACGCA ATAGATGAAG	9600
	TGTCTTTCAG CAATGAAATA AAGGATATGG TACAAATTCA ACAAGCGCTT AACAGACGAG	9660
	TTATTGCACA AGATAATAGA TATAACTATC AAGCAAATCG TATAAATCAT TTATACACTA	9720
45	GTACTTTGAA TTCTCCTTTC GAGACAATGG ATATAGGGAG TGTATTAATA TAATGGCAAC	9780
	AGAAGAAGTT AAAATCAAAG CGCTACTTGA AAACGATAAA C	9821

(2) INFORMATION FOR SEQ ID NO: 471:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1017 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 471:

5 TTATTGTTTT CCAATGGTC TTCTCGGTTA TTAAACCATC GGTTCATC CTACGTGCTT 60
 CACTTTCTAT CtATTAATTC AaTTTCAGCT TGACCACCCG CTGTATAAAG GGTCAAAGTT 120
 GCTAATCGAT AGCGTCTCAT TATAGGACCA ACATCGATAT CAATATTTTG AATACGAAAA 180
 10 TATGGTATTA CCTTTTCATC CAAAAATAAA ATGCCGTTTC GTACACGCAA ATGGTGTGTTT 240
 TCAAATGCAT ATCTGCAGTG CTTATATCGA TAGACTGGCG CTATAACAAG CGTGAAAACA 300
 GCAACAAGTA ATATTATAAT CACACTACTA ACAATGGATA AATGGTTATC TAAAATCTCC 360
 15 CAAAATAGCC AGTTCAAAAT ATTAAATGCG ATTAAAAGTA CAAGCGCTAT GGGTATCCAA 420
 AACAGCACAG CACTTAACCT CATCACTTTT TTAGCGTGTG GTGACATAAA ATTATAATCC 480
 CTCATCATTT TCACCTCTTA AATACCATGA TTTCAATTTA TTTGCATCTT CACTTCTTGC 540
 20 GTATTTTAAG TTAATCGATT GGGCTCCAC ACCTTTAGCA ACAATAAAGC TAAAATTATT 600
 TAAATTGTTG CGTTTAAGTA ATGTATGTTG CCAAGTGTCA AATCCTATAA TGTGATGCGC 660
 TTTAAAATAA TAAATATTTT GTTCAATAG CTCGAAATTC TGGATAGTAA TTTGTTCTTC 720
 25 TGTCATTTTA AAACCCGCAT GTTGACATA AAGATATCCT TTGATCACAA ATAAACCAAT 780
 AATGACTATT GTTATAATCG TAAATAACAA TAATAATTGA TTCCAAAAT AACAGCCTAT 840
 30 ACCTGCCATA GCTATGACAA TAATACTAGG TATTAAAAAG TGTCTGTGGA AACCTGACAA 900
 AGGCATACCT TCATTAACTT GTTGATAAGA TAAATCTGGT ACTAAATTCT GGATAATTTG 960
 ATATGCTTTG TCTCGTTTAA TAAACGGCaA TATCGGCACA CTACCTGAAT CATGTGC 1017

(2) INFORMATION FOR SEQ ID NO: 472:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6806 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 472:

45 TAGGATTGAT TAATCCTTTC TTCAAATGA TGAATGTGTA ATGTTAAATA TATAATTTCA 60
 GACTCACTGA TATTAACATC AAATTGTTTT TGTATCATAG TTAATTTTTT ATATGCTGTG 120
 50 TTATAGCAAA TCGGATAGTG ATTTTAAATC ATAGACACAA AATCATCTTG TGCATGTATA 180
 TATTCTTTTC TTCTTAATCG GCGAATTAAA AATTGTACGT GCCTTATAAA ACCTTGGTAT 240

	TTAATAACAT	TATTnGATCA	AGGTCATCTC	ATGCATAGAT	AAATCTTCTG	TATTaGATGC	360
	AATATGTAAT	GCAATAAATC	CTATCTCATC	TTCAGGaAAA	TGTaCATCTA	ATGCTGCATT	420
5	TAAC TGATCA	ATCACCTGTT	TAGCAATATG	ATATGCATCA	CTATATAACT	GCATAGTTTC	480
	CATAACAAAT	GGATTGCTAA	TAAC TTGATT	TTGTTTTAAG	CGTTTATAAG	CAAATATAAT	540
	ATGATCCGTT	AATGAAACTA	CAAGTTGTTT	TGAATCAACA	TTCATCGCAG	TATTAGAAAT	600
10	AAAATTCAAC	GAATCAATAA	TTACTTGTA	TACATTATCA	TCAGCGATTT	CAACTAAACT	660
	TTTATAATGT	GCTTTTTGTT	GCTCACTTTC	TAATTTATAA	ATTTTCTCTA	TTGTAATAGT	720
	TTGGTCGTTT	AACGCCATT	CCTCTTTTTT	GTTAAAACCA	ATACCTTTAC	CGATTAAAAT	780
15	AACTTCTTGA	TCATTATTAG	TACATACTAC	GACATTGTTG	TTCAATGTTT	TAGTAACAAT	840
	ATATTCTCCC	ATTATCATCA	CCTATTTTTT	TATTATTAAG	ATTATATATC	GGAAATGTCT	900
20	AACTTGAAAG	TATAATAATT	TAAATACTTG	TGTTCTATTA	CACGCCTATC	CTATATGATA	960
	TATCTTAATT	TAATTTGAT	GTCTCTCAAA	GTGGAATAAC	TATAATAAAA	AAATCTGACT	1020
	CCCATATTTA	CAAATAATTC	TATTTATTTA	TATATTATCT	GAATTAATAC	TCAATACAAA	1080
25	CTAAAACGTA	CTATTAAATT	GTGCAAAGCT	AAAACAAATT	TATATTCATC	TATCCAACAA	1140
	TATGTCTTAT	CAATGGTATA	GTCTTTGCAC	ACCAATGGAG	GAAATAAATC	TCAACCTTAC	1200
	TATATTAATA	TATAATCAAA	TCTTAGATTA	ACTAGTGTA	TGATACAGAT	GATAATTGAG	1260
30	TACAAATTTA	AAACCCTGAG	ATTTTCGCTT	TAATTTGAAA	ACCTCAGGGT	TTATTTGATT	1320
	TTTATATAAT	GAATCGTTAC	ATTAAAAATA	TTTATTTATC	AGAGTTCTTA	TATTTGTTAG	1380
	CGCCCCAAGC	ACTAATTCCA	AATAAGTTAA	TTTCTAAGTT	TTCAGGTTTA	AAGACAGGGT	1440
35	TCTTGCCTTC	TTTTTTCTGC	TTTTGATAAT	CTTTCATCAA	TGCAAAAGCT	ACATTGGACA	1500
	GTCCTATAAT	GGAAATAATG	TTTACAATTG	CCATTAAGCC	CATAAATAAG	TCTGCCGTAT	1560
40	TCCATACTGT	TTCTGTTTTT	ACAACTGCAC	CGACAAAGAC	AAGTACTACA	ACAAGACATC	1620
	TAAAGATAAA	TAATATTACA	CGGTTTGTTG	ATAAAAATTC	AATATTAGAT	TGACCGTAAT	1680
	AGTAATTACC	TACAACAGAT	GAAAATGCAA	ACAGTGTAAC	tGCTaTTGTT	AAGAAAATAC	1740
45	CTCCAGCAGA	ACCTAAATGC	TCATTAAGTG	CTGATTGAGT	AACTGCAACA	CCTTGAGGTG	1800
	CGTTATCACC	AAATTTTCAGT	CCTGAATATA	GTAAAATCAT	GATTGCAGTT	GCTGTACAAA	1860
	CCAACATTGT	ATCAAAGAAC	ACACCTAATG	ATTGGATTAA	ACCTTGCTTA	ACAGGGTGTG	1920
50	GTACGGCAGC	AGTTGCCGCT	GCATTCGGCG	CAGAACCCAT	ACCAGCTTCG	TTAGAGAATA	1980
	AACCACGTTT	GATACCTTGA	AGAACCGCAG	CACCTACAGC	GCCACCAGTT	ACTTGTTCTGA	2040

55

	GCAATATTAC TAAAACCATA CCAATGTAAA TGATAGCCAT AATCGGTACA ATTAACGAAG	2160
	ATAACGTAGC AATACTACGT ACACCACCAA ATATAATAAT AGCTGTTACG ATTGCTAAAA	2220
5	TAATACCTGT GATTACTGGA CTAATATTAT ATTGCGTATT TAACGACTCC GCAATTGTAT	2280
	TAGATTGCAC TGTGTTAAAT ACAAATGCAA ATGTAATTGT AATTAAAAATC GCAAATACGA	2340
	TACCTAGCCA TTTTTGATTT AAACCTTTAG TAATATAGTA AGCTGGACCA CCACGGAATC	2400
10	CACCATCTTT ATCATGTACT TTATAAACCT GAGCCAAAGT CGCTTCTATA AATGCACTCG	2460
	CTGCACCTAT AAATGCAATA ACCCACATCC AAAATACTGC ACCTGGACCG CCTAAAACAA	2520
	TCGCAGTCGC AACACCAGCA ATATTACCAG TACCAACTCT CGAACCAGCA CTAATCGCAA	2580
15	ATGCTTGGAA TGGCGAAATA CCCTTCTTAC CATCTTCTAA AGTTTCTGGA CGTTCTACTA	2640
	AAGCTCTAAA CATTTTCAGGT AACATTGTA ATTGAACGAA TTTAGAACTA ATCGTAAAGA	2700
	AGAATCCAGC TGTCAATAAT AGACCAATTA AATATTGAGA CCATATTAAA TCGGTACCAA	2760
20	CATGGACAAA TTCTTTAAAC CATCCAGGTA TTAAACTATC GAAATCTTTC AAAATAAACC	2820
	CCTCGCATCC TCTACATGAA TCATGTACCT TCTATAAAAT TAGACCGAAT TGAACCTTCA	2880
25	GTAAATATAG AGATACATCA TCATTTCTTA TACAATACAA GAGATTTATA TTAGTTTGGT	2940
	CAAAGTATAT CGCTAATTTA ACGATAAGTA CTTGGTCAGC ATTTAATATA AATCCCTTGA	3000
	ATTTAGTCAA AATTTAACAT TACTGTATTT TATCATTTAA TTTCGTGATT GCATATAGTT	3060
30	TTTAGCTAAT ATACATGTCT ATTACTTCAC CAAAATCATC TGTATCTACA ATGAATGAGC	3120
	CATTTGTATA TTGTTGAGAT TTATGAATAT CATTAAATTAA ACCATGTTCT TCATTTGATT	3180
	TTGAATATAA TGTATATTGA CTATGTTTAC CTGTCACTAC ATGTGCAGCT ACAATACGAT	3240
35	GTGGATTTTT CTTTAATTCT TTTAATAAAG TTATTCCaCG TTGTGCTCTT TTAGCAACTT	3300
	GTAAGATTTTT AAAACTAATA CGTTTTAACG AGCCGCGTTG TGTGGCCATC AATATAGTAT	3360
	CATTTTCAGA AACACCTTCT GTCATAACAA CGAAATCTTC AGCTTTAAGA TTTATTGATT	3420
40	TAACACCAGC TGCCCTTAAT CCGGTATCTG ATAGTTCACT TGTATTATAC GTTAATGACA	3480
	TACCTTTATT AGTAATGACG GTAATTAATT GATCTTTTTT AAAGCGCATA AACTAATCA	3540
	AATCATCATT TTCTTTAACT TTAGTAGCAA TTAAAGGTTT ATTAACACGC GTTGTTTTAA	3600
45	ATAGAGGCAC TGTACTTTTC TTAATCATGC CATTTTGAGT CGCAAAAACA TAAATGCAT	3660
	CTGTATTAAA GTCCTTTTCA TTAAAGACAT TAATAACCAC TTCATCTTCT TCGATAGGAA	3720
50	CTATTTGTGA TACATGTTGT CCCAATTCTT TCCAACGAAT ATCTGCTAAT TTATGAACCG	3780
	GTATAAATAG ATAACGACCT TTATTTGTAA ATACTAGTAC GGTATCTTGC GTATTTACTT	3840

55

EP 0 786 519 A2

	TAAAGCTACG	AATAGAAGTA	CGTTTAATAT	ATCCATGACG	TGTCATACTT	AAAATAACTY	3960
	CTTCACTAGG	CACCATAACT	TCTTTGTCAA	TTTAAATTTT	TTCAATTTCT	GCTTCAATTA	4020
5	AAGACAGTCG	TTCAGATTTG	AATTTCTTTT	TAATTTTCATT	CAATTCTTCT	TTTATGACAT	4080
	TCAATAATGC	ATCATGGTTA	TCAAGAATAT	GACGTAATTG	TTTGATTAAT	GCTTCAAGTT	4140
	CTTTATGTTC	ACCTTCAAGC	GCAACTATGT	CAGTATTTGT	TAAACGATAT	AACTGTAACA	4200
10	TTACAATTGC	TTCAGCCTGT	TCTTCTGTGA	ACTCGTATAC	TTGATAAGG	TTTTCTTTAG	4260
	CGTCACGCTT	GTTTTTAGAG	CWACGAATCA	ATTCGATTAC	TTTATCTAAA	ATTGACAACG	4320
	CTTTAATCAA	ACCTTCAACG	ATATGCATAC	GTTTTTCTGC	ATTATCTAAT	TCAAACCTCG	4380
15	TTCTATTTGC	AACAACCTCA	ATTTGGTGAT	TCAAATAACT	ATCTATAATT	TGACGAATAC	4440
	CCATCAATTT	TGGACGACCA	TCACTAATAG	CGACCATGTT	GAAATTATAT	GAAATCTGTA	4500
20	AATCAGAGTT	TTTATAAAGA	TAATTTTTGA	TTGATTCACT	GTTACATCT	TTTTTCAATT	4560
	CAATTGCTAT	TCGTAAACCA	GTTCTATCAG	TTTCATCAG	TACTTCAACG	ATACCATCGA	4620
	CTTTTTTGTC	AGCACGTAAT	TCATCGATAC	GTTTTACTAA	GCTACTTTTG	TTCACTTCAT	4680
25	ATGGAATTTT	AGTAATAATT	AACTGTTTAC	GTCCATTGCG	TAAAGTTTCT	TCTTCAACTT	4740
	TAGAACGAAC	TATAATTCTA	CCTTTACCTG	ATTCATAAGC	TTTTTTAATA	CCATCAATAC	4800
	CTTGAATAAT	ACCACCAGTT	GGAAAATCAG	GACCTTTAAT	ATATTTTCATT	AATTGATTGA	4860
30	CTGTAATATC	CGGATTATCA	ATATATTTAA	GTGTTGCTTG	AATCACTTCA	GCTAAATTAT	4920
	GTGGTGGTAT	ATCTGTCGCG	TAACCTGCAG	ATATACCTGT	AGAACCATTG	ACTAGTAAGT	4980
	TAGGAAATCT	TGATGGCAAT	ACCATTGGTT	CGAGTGTCGT	ATCATCATAG	TTTGGAAATGA	5040
35	AAGAAACTGT	CTCTTTATTA	ATATCACGTA	ATAACTCTTC	AGCTAGTAAG	CTTAACCTAG	5100
	CTTCAGTGTA	ACGCATTGCC	GCTGGCGGAT	CATTATCGAT	ACTACCATTA	TTACCATGCA	5160
	TTTCTATTAA	GACATGTCGT	AACTTCCAGT	CTTGACTTAA	ACGGACCATT	GCTTCGTACA	5220
40	CTGAGGAGTC	TCCATGTGGA	TGATATTGAC	CAATAACATC	ACCGACTGTT	TTGCACTTTT	5280
	TACGGAAATT	TTTATCGTGT	GTATTACCAC	TTGAATACAT	TGCATATAAA	ATACGACGTT	5340
	GTACTGGTTT	TAAACCATCA	CGAACATCTG	GCAATGCACG	CTCTTGAATA	ATATATTTAC	5400
45	TATATCTTCC	AAAGCGATCA	CCTAAAACAT	CTTCAAGTGA	TAAATCTTGA	ATTATTTTCAC	5460
	TCACTAGATT	TCCTCCTCAT	CAAATTGATC	ATTTTCAAGC	ACTTGTAATT	CAGAAATTATC	5520
50	TAAAATACTT	TGGTCCTCTT	GCATACCAA	CTCAACATGC	TTTTCAATCC	ATTCACGTCT	5580
	AGGTTGTAAT	TTGTCACCCA	TTAATGTTGT	TACACGTTTA	GATGAACGCA	CTTCATCTTC	5640

55

AGGGTTCATT TCACCCAAAC CTTTGTAACG TTGTAACGTG AaGCCTTTAC CAAGTTCTTT 5760
 TTGCAATTTA TTAAGCTCTT CGTCTGTCCA AGCGTATTCA ACTCGCTTTG TTTTGCCCTT 5820
 5 ACCTTTTTC C AATTTATAAA GTGGAGGTAA AGCAATAAAT ACACGACCTG CTTGAACAAG 5880
 CGGTTTCATA TATTTGAAGA AGAATGTTAA CAATAGCACT TGAATATGCG CACCATCAGT 5940
 ATCAGCATCA GTCATAATAA TTACACGATT ATAATTACTA TCTTCAATTT TAAAGTCAGT 6000
 10 ACCAACGCCT G C C C C GATTG TGTGGATAAT TGTATTAAAT TCTTCATTTT TAAAAATATC 6060
 TTCTAGACGT GCTTTCTCTG TATTAATTAC CTTACCACGT AATGGTAATA TCGCTTGGAA 6120
 TTTGCGGTCT CGTcCAAGTT TTGCTGAACC TCCCGCAGAA TCACCTTCGA CTAAATACAA 6180
 15 TTCATTTTTT TCAGTGT TTT TACTTTGTGC AGGTGTTAAT TTACCAGATA GCAAAGTGTC 6240
 TTTACGCTG TTTTCTTAC CTGAACGAGC ATCTTCACGA GCTTACGTG CAGCTTCCCT 6300
 20 TGCTTGTTGT GCTTTAATCG CTTTTTTCAC AAGTGATTTA GACAATTGTC CTTTTCTTC 6360
 TAAATAGAAT GGCAATTTGT CTGCAACAAC TGAATCAACA GCACCTCTAG CTTCAGAAGT 6420
 ACCCAATTTA GATTTTCGTT GTCCTTCAAA TTGCAATAAT TCTTCTGGAA TACGAACAGA 6480
 25 CACAACAGCT GTTAAACCTT CACGAATATC ATTACCATCT AAGTTTTTAT CTTTTGTTTT 6540
 AAGTTCATTA ATACGACGTG CATAATCATT AAATACACGT GTCATTGCTG TTTTAAAACC 6600
 AACTTCATGT GTACCACCAT CTTTAGTACG TACATTATTT ACAAACCTTA AAATACTTTC 6660
 30 TGAATATTGA TCATTATATT GGAAAGCTAC GTCTACCTCT ATACCATTG CTTCACTGA 6720
 AAATGTAGCC ACGTCATGCA AACTTCTTT TCCTTCATTG ACATAACTAA CAACTCTTT 6780
 GATTCCTTCT TATAATGGTA TGTCTT 6806

(2) INFORMATION FOR SEQ ID NO: 473:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1716 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 473:

GGGGCAhAAA TTCCAATACA CTCAATTACCA AATATATACA CCACTTCCTG CTACAAGTnn 60
 TTTTACTTGA TCTTGGTCTT TTCCGCAGAA AGAGCATTTT CAAATTTTCT TCATCTTCCA 120
 50 TTGAATTTAA ACATTCTTTT TACACCCCTA TTCGTTAAAG ACTATACTAG ATTGGATGTT 180
 ACAATGCAAC ATATTACAT ACAAACCTTT TGCTTAAAGA ATAGTAGCAG ATACATAAGC 240

5 CcTCAACGA ACTTTGCGTT ATCTCTTAAT AAATCGATAA CTTTTTGGAT ACGAACATCA 360
 TTTTAAATGA TATCAGTATT ACCTAAAGTA TTTTGTATAT CTTCAACTGA GATATTAAAT 420
 10 TGTTTACTCA TTTTTTCTAA TTCTTTATCG ATATCTTCAT CAGTAGCTTC GATTTTTTCA 480
 GCTTCAGCGA TCGCAGTTAA AGTTAAGTTA GTTTTAACAC GTTGTCTGTC ATCGTCTTTC 540
 ATTTGCTCTC TTAATTGAGT TTCATCTTGA CCTGAGATTT GGAAGTACGT TTGTAAATCT 600
 15 AAACCTTGTT GTTGAATTCT TTGTGCAAAT TCAGACACCA TACGATCTAA TTCAGTATTA 660
 ACCATTGCTT CAGGAATATC GATTGTGTGA TTATCAGTAG CTTTTGTAAT CGCTTCTTCT 720
 TTTTCAACAT TTTCAGCATC TGTAGCTTTT GTTTCAGCTA AACGTTTACG TAAGTTTTCT 780
 20 TTGTACTCGT CTAATGTATT TGCTTCTGCA TCTAATTCAT TAGCAATTTC ATCTGTTAAT 840
 TCTGGGACTT CTTTAAATTT AATTCGTTA ACTTTTGTTC TGAAAGTTGC TTCTTTACCG 900
 GCTAATTCCT CAGCATGGTA TTCTTCTGGG AATGTTACGA CAACATCTTT TTCTTCGTCA 960
 ACTTTCATAC CTTCTAATTG CTCTTCGAAA CCAGGTATGA ATGAACCTGA ACCGATTTCT 1020
 AAATCGTAAC CTTCAGCTTG TCCACCTTCG AATTCCTTCc CGTCAACTGA ACCACTAAAG 1080
 25 TCGATGTTAA CTGTGTCGCC ATTTTCAACA ACACCATCTT CTTTAACGAC CATTTTCAGCT 1140
 AAATGTCCTA AGCTGTGGTC AATCGCTTCT TGTAATCAT CATCAGATAA TTCAGTTTCT 1200
 TGTTTTTCAA TTTCAAGACC TTTATAGTCT CCTAATTTAA CTTCTGGCTC AACTGTAACT 1260
 30 GTTGCTTCAA AAATGAAATC TTTACCTTTT TCAATTTGAG TAACACTTAC TTCTGGTTGT 1320
 GCAACTGGTT TAATATCAGT TTCGTCAATT GCTTCACCAT AAGCATCTGG TAATAAAATG 1380
 TCGATAGCAT CTTGATATAA TGCTTCTACA CCAAAGCGTT GTTCAAAAAT TGGACGTGGC 1440
 35 ACTTTACCTT TACGGAATCC AGGTACGTTA ATTTGTTTAA CCACTTTTTT GAATGCTTGA 1500
 TCTAACGCTT TGTTTACTTT TTCTGCAGGA ACAGTAACAG TTAATAAACC TTCGTTACCT 1560
 TCCTTTTTTT CCCAAGTTGC TGTCATGTAT ATATACCTCC ATGATTAACT AATTTATTTT 1620
 40 TTCAACTTCC CTATTATATC ATACGTCTAT TCCCTATACA AACATTGAAA TCACAACGTT 1680
 TATATATTG TAAATCAACT TTTTCGTCA AACTA 1716

(2) INFORMATION FOR SEQ ID NO: 474:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 795 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

TGACCAAGTA CGTTTCGAAG TTGCCATTAA AGCATTAAAC CCATCATTGA AAGCATTTCGC 60
 ACCTGTACGT GAGTGGGCAT GGAGTCGTGA AGAAGAAATC GATTATGCAA TTAAACATAA 120
 5 TATCCCTGTA TCAATCAACC ATGATTCACC TTATTCTATC GATCAAAATC TATGGGGCAG 180
 AGCGAATGAA TGTGGTATTT TAGAAGATCC TTATGCTGCG CCACCAGAGG ATGCGTTCGA 240
 TCTAACAAAT GCTTTAGAAG AAACACCAGA TACTGCTGAT GAAATCATT TACGTTTGA 300
 10 TAAAGGCATC CCAGTTCAAA TTGATGGCAA AACATATGAA TTAGACGATT TAATTTTAAC 360
 GTTGAATGCA TTAGCTGGTA AGCATGGTAT CGGAAGAATT GACCATGTAG AAAATAGACT 420
 TGTAGGTATC AAATCAAGAG AAATTTATGA GGCACCTGCT GCAGAAGTTA TTTTAAAAGC 480
 15 GCATAAAGCA TTAGAAACGA TTACGTTAAC GAAAGATGTC GCACACTTTA AACCAATCAT 540
 TGAGAAGCAA TTTGCTGAAC AACTATACAA TGGACTTTGG TTCTCACCTT TAACTGATAG 600
 20 CTTGAAATTA TTTATTGATA GTACTCAGCA ATACGTAAGT GGTGATGTCA GAATTAAATT 660
 ATTCAAAGGT AATGCCATCG TGAATGGTAG AAAATCACCT TACACATTAT ATGATGAAAA 720
 ATTAGCAACT TATACAAAAG AAGATGCATT TATCAAGAC GCTGCTGTTG GCTTTATCGA 780
 25 TATCTATGGT TTACC 795

(2) INFORMATION FOR SEQ ID NO: 475:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 887 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 475:

CGATTGAAAG AAGACGGTTC AGTTGAAAAG TTTCCAAAGC CAGTAATTAG CCAACAACCA 60
 GAAGGATATA CGAGTCATTT TAGAGATCCT AAAGTTTTTA AATATGATG GAAATATTAT 120
 40 GCAATCaTTG GTGmCaAAA TAATGATCaG CaAGGTCGAT TATTACTTTA TAATACTGAA 180
 GATATAATTA ATTGGCATT TTTAGGTGAA ATAAATACAG AGTTGGATGA TTTTGGATAT 240
 45 ATGTGGGAAT GCCCAGATTA CTTTAATGTA GATAATCAAG ATGTCATACT TATTTGTCCA 300
 CAAGGTATTG AACCAAAAGG CGATCAGTTC AAAAATATTT ATCAAAGTGG TTATATACTT 360
 GGAAAGTTTG ATATTGAAAA GTTAACATAT GAACATGAAA ATTTTGTCTGA GCTTGATAAT 420
 50 GGTTTTGATT TCTATGCACC TCAAACATTT TTAGATGAAA AAGGCCGACG AGTACTAATT 480
 GGATGGATGG GGTACCAGA AATCGAATAT CCTACTGATA ATGAAGGATG GGCCCATTCG 540

GCGTTGGAAA AATTACGTCA CAATAAAGAG ACAGCATTtA GGctACGCAA ATAAATTTAC 660
 TCGAAAATTA CATCCGTATG AAGGTAAACA GTATGAATTA ATCATAGATA TTTTGGATAA 720
 5 TGATGCTACC GAAGTGTA CT TTGAATTACG TACATCTAAG ACTTCTTCAA CATTAAATTGC 780
 TTATAACAAG CGTGAAAATA AAATAACATT AGATCGCAGC GACAGTGGTT TATTGCCGAC 840
 AAATGTTGAA gGTACGACGC GTAGTACGAT ATTAGACACG CCATTAA 887

10 (2) INFORMATION FOR SEQ ID NO: 476:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1183 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 476:

TTGGAAnCAA AAAACCATTG GTAAACCGTG TtnAACCGGA TTTCCGATGG ACCTTTTAAA 60
 AChACCAAAT AGAAAGCTTT GATAAAAGGT AATTATGGTA CTGATAACAA ACAAGTTCAA 120
 25 AAACATCATG ATTTAGTACG TATGCTTTTG ATGGATCAAG ATGGTTTTTTT AACTGAAAT 180
 AATAAAGTTG ATCATTTTCAT TGATGGAAAT GATTTATATG ATCAAGTTTTT AAAAGATATT 240
 AAAAAATGCAA AAGAtATATC CATTTAGAGT ACWATACTTT CGCTTWAGAT GGTTWAGGTA 300
 30 AAAGAATTTT ACATGCTTTA GAAGAAAAAT TGAAACAAGG TCTAGAAGTA AAAATATTAT 360
 ATGATGATGT TGGATCTAAA AATGTTAAGA TGGCAAATTT TGATCATTTT AAATCGTTAG 420
 GTGGAGAAGT TGAAGCATT TTTGCTTCAA AATTACCGTT ATTGAATTTC AGAATGAATA 480
 35 ATAGAAATCA TAGAAAAATC ATCGTAATCG ATGGTCAACT AGGTTATGTC GGAGGATTTA 540
 ACATTGGTGA TGAATATCTm GGATTAGGAA AATTAGGATA TTGGAGAGAT ACGCATTTAC 600
 GTATACAAGG GGATGCGGTT GATGCACTGC AGTTGCGATT TATTTTAGAC TGGAATTCGC 660
 40 AAGCGCACCG TCCACAATTT GAATATGATG TTAAGTATTT CCCTAAAAAG AACGGACCAT 720
 TGGGCAATTC ACCAATTCAA ATAGCTGCAA GTGGCCCGGC TAGTGA CTGG CATCAAATTG 780
 AATACGGTTA TACAAAAATG ATTATGAGTG CAAAGAAATC TGTATATTTA CAATCACCAT 840
 45 ATTTCAATTCC GGATAATTCA TATATaAATG CCATTAAAAT TGCTGCTAAA TCAGGTGTAG 900
 ATGTACATTT AATGATTCCA TGTAAGCCAG ATCATCCATT AGTATATTGG GCGACATTTT 960
 50 CAAATGCCTC TGA CT TATTA TCAAGTGGTG TTA AAATTTA TACGTATGAA AATGGATTTA 1020
 TACATTCTAA AATGTGCTTA ATTGATGATG AAATCGTATC AGTGGGCACA GCAAATATGG 1080

55

CTAAAGATTT AAGGGTGGCT TATGAACATG ATATTACAAA ATC

1183

(2) INFORMATION FOR SEQ ID NO: 477:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2332 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477:

GGTTATATGT TTGGTATTAC TCATTATATT GATTGGAGGT TGTGTCATTA TGACAAAAAC 60
 15 AAATGGTCGA AACGCTCAAA TTAAAGAAAA TTTCAACAAA ACATTAAGTG TATATCTAAC 120
 CAAGAATCTC GATGATTTTT ACGATAAGGA AGGTTTTTCGA GATCAAGAAT TTGATAAAAG 180
 20 AGATAAAGGG ACTTGGATTA TTTATTCTGA AATGGTTATC GAACCAAAAG GGAATAATAT 240
 GGAATCGAGA GGAATGGTGC TCTATATCAA TCGCAATACT AGAACAACGA AGGGTAATTT 300
 TATTGTCACC GAAATAACTG AAGATAGTAA AGGATATTCA CGTAGTAAAG AAAAAAATA 360
 25 TCCTGTCAAG ATGGAAAATA ATCGAATTAT TCCAACAAAG CCTATACCGG ATGACAAGTT 420
 AAAAAAGAG ATTGAAAAC TTAAGTTCTT TGTACAATAT GGAAATTTTA AAGATTTTAA 480
 AGATTATAAA AATGGTGATA TTTCAATATA TCCTAATGTG CCAAGTTATT CTGCAAAGTA 540
 30 TCAATTGAAT AATGATGACT ATAATGTTCA ACAGTTAAGA AAACGATATC ATATTCCAAC 600
 CAAACAAGCG CCCGAATTAA AATTGAAAGG ATCCGGCAAT TTAAAAGGCT CATCCGTAGG 660
 ATCTAAGGAT CTAGAATTTA CGTTTGTAGA AAATCAAGAA GAGAATATCT ATTTTTCAGA 720
 35 TTCGGTCGAA TTTACACCTA GCGAGGATGA TAAATCATGA GTCAAACGGA ATATCAAATA 780
 AAATCTGGCA ATATAAAAGG TAACTCTGAA GAAACAAGTA CAGTATCTAA TATAAGTTAT 840
 GAAATAGAAA ACGCAAATAA CAGTGGTTTA AAACAAAATA AAATTGATAA ACAAATTTAA 900
 40 AAGTTACAAG AAAAAAATAA ATTCCTTAAA AATCTTTTCA ATCTTAAAAG TTATACGGAC 960
 CCCAAAACAG GCACGACTAC AAGCGCCTTT TTAAATAAAG ACACTGGCAA AGTTACTTTA 1020
 45 GGTATGACAG GTACTAATGT ACACAAAGAC GCAATATTAA AACAAACATT TGGTGTTCCT 1080
 TCTTATCAAG GATATATAGA CGTGAGTGAA ACgCTaAAAG ATATTGGGGC CGATGTCAAT 1140
 ATTGGCCTTC ATTCCGTCAC AGATAAAGAT CCACATTATA AAAATACCCA AGACTTTTATC 1200
 50 AAAAATATCA AAAAAGACTA TGATATTGAT ATTATTACCG GACATTCGCT GGGCGGTAGA 1260
 GATGCGATGA TTTTAGGTAT GAGTAATGAT ATTAAACATA TCGTTGTGTA TAATCCAGCT 1320

55

ATTGAAAAGT ACGATGGTCA CATTGTAAGA TTTGTGTCTG ATGAAGACGA ATTAGATGCA 1440
 GGTGTCCGCA ATCATTATA TGAACTGCT GGAGAAAAA TAGTACTTAA AAATGGAGAA 1500
 5 GGCCATGCAA TGAGTGGTAT TTTAATGAGC AGAACACAGG CTATAATCTT AGCTGAATTA 1560
 AACAAAGTTA AAGGCTACCA AGACGAAAAT AATAAAGCAT TAAAATCCGT TCGTAAACAA 1620
 ACGAGGCATA GATTACATAA AGTAGAGACG TTAAGAGCGA ATTGGATTCA AACACGGGT 1680
 10 GGATCACTCT CTTCCTCCCw ACAACAATTA TTAGAAGCTT TAACAGCACT AACCATTGCC 1740
 GAAGGCTTAA ATCAATTAGT GAATGAAGAA AGCCAACATT TGAAAAAATG TATCACGCGA 1800
 TGGCACATAA ATTTGGAGAC AACTGGAAAA AAGCGCAAGA AGTTGGAAAT GAAATTGGTG 1860
 15 AAAAATTAAC CTCTGAAGAG GTTATAGATG TATTAAGAAA AGGTGGCGCG TATGAAAGTa 1920
 AACTTGAAAC AGATCCCAAA AGAAAAATTG ATGATAAGAT AAAGAAATTA AATGATGTTT 1980
 ATAAAAATTG TAATGGCTAT ATCGCAAAAA TTAAACAGAG TATCGAAGCA ATTGTTTCTA 2040
 20 ATGACCAAAT GTTAGCGAGC CAGATTGATG GGATGATGTA ATGTTTACTA CGTATAAnAA 2100
 TATTAATGAA CTTGAAAATG CCTATGATGA AGAAAGAAAA CAATTGAATG ATGCATTCAA 2160
 TCAAATTGAT GAATTAAGAC ATCAAACACG CAAGAmATGT GAACAAATGT ATGATCAITT 2220
 25 CTTATATCTC AACATAAAA TGAATTmymS TGAAGACGCT ATGATCAGGA TGACACGTAT 2280
 TATAGAATCT TTCGATAGAG AAACGAATCA ACGTATCCGA CATCACGAAA TG 2332

30 (2) INFORMATION FOR SEQ ID NO: 478:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 865 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 478:

40 TTTACTACCC AGTATCTCTT TTTATAAATT ATATAGCCAC CACATATGGT GGAAAGTCTT 60
 TTTAATTAGA ATTTTGTTTT TTCAGTTAAG AAAGCTTCTA ACTCTGAGAT TGGCATACGA 120
 ACTTGTTCCA TTGAATCTCT GTCACGTACT GTAAGTTGAT TATCTTCTAA TGAATCAAAG 180
 45 TCGAATGTtA CACAATAAGG TGTaCCGATT TCATCTTGAC GACGGTATCT TTTACCGATA 240
 GATTGTGATT CATCGAAATC GATTGAGAAT TTAGAACTTA ATTGCTCAAA AATCTTAATC 300
 50 GtTCGCCAGA TAATTTCTTA CTTAAAGGTA AAATCGCTGC TTTATATGGT GCTAATGCAG 360
 GATGGAAGTG TAAAAGTGA CGTGCACTCT TACTACCTTC AACGCCTTCT TCATCATATG 420

55

GAATATATTT TTCGTTTCGTT TCTGGATCAT GGTATCTGAA ATCTTCACCA GAGTGTTCAG 540
 CATGTTTACG TAAGTCGAAG TCTGTACGAC TTGCGATACC CCATAACTCA CCCCAACCAA 600
 5 ATGGGAATTT ATATTCAATA TCAGTTGTTG CATTGAGTA ATGAGATAAT TCATCTTCAT 660
 CATGATCACC TAAACGCATA TTTTCACTGC TCATATTTAA GCTTGTTAAC CAGTCACTTG 720
 CAAAAGTTTT CCAATAATTT TGCCATTGCA TTTCTTCTCC AGGTTTACAG AAGAATTCAA 780
 10 GTTCCATTG TTCAAATTCT CTTGTTCTGa AAATGaAGTT ACCTGGAGTG aTTTCaTTAC 840
 GGaATGaTTT ACCAATTTGG ACCGG 865

(2) INFORMATION FOR SEQ ID NO: 479:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1444 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 479:

25 TAATGAGTAA ATAAGTACCA GATAAAATCA TGAATATCAT CCAACATGCG GTTAACTCTA 60
 CTAAATAATT AATAATAGTA TTTTCAGTAA ACAAAGAACT ATGTATACTT CGCATCACAT 120
 TAGAATACGT ATGTTTCGCA TTTTGATCTG CAACAAATTG ATTGTTATGA TCTAGAAAGA 180
 30 CGTAACGTG ATTTCCCTGCC ATATCACTCA GTGTAATTG TTTGTTATAT GGTTCATCAA 240
 GTATGCTAAC TTTACTTACA AAGAATCCTT CATATTGTTG TTCAACTTGA TGTACTGCAT 300
 CATTTAATGT TTGATGCGTT TTTACATCAC TGTCACCAA AACTCATTG TTATAAATAT 360
 35 TATTTTCAAC TTCTGGAAG AACAAGTAAC CAATGCCCCG AATGGTTAA GTGATTAACA 420
 GTGGAGCAAT AAATATTGCT GCATAGAAAT GTAATCTTTG TAATGGATTA AATGTATTTT 480
 TCATATTTCC CTCCCAATTG GCTATTATAC GGTGTCAATT CTGTGATGTG TGTGAACAA 540
 40 CTGTGACAAT ATTTATTTTC TAGAAAAATT TAACGATGAT TTGTGATTTT TAGAAAAATG 600
 AACTTTTAAG TTGGAATGTT TGAAGAAAAT TGATTATTCG TATGTTTTAT CAAGCAGCTA 660
 TGATAAAATT TAAACATAAT ACAATGCGAG CCATTAAACG ATCTATGTTT AAATGGACAT 720
 45 CGATATTGTA TGAATTCGTT GTAACAAGCA AGCATTCTTA TGTGAACGAA CCAAAGGGGA 780
 AAGTAACATG ATTAATAAAG AACAATTAGA TCTTTTATAT AAATTAAAAA AAGAAGTTGA 840
 50 AAAGTCGCGA AATGAAGCAC TTTTACATAC AATTAACCAA GTAATTAAGA AAGTATATTT 900
 GCAGCAATAT ACATGTTTCG TCGTTGGACA TTTTCTGCA GGTAAATCGA CACTGATAAA 960

TATTGTGTCA GTTTCAGACA ATCACGATAT TATTGCTAAT TTGCCGAATC AAACGTATGC 1080
 CAAATTATCT AATTATGATG AAGTAAGGGA AATGAATCGC CAAAATGTCC ACGTTGAATC 1140
 5 TGTAGAAATT AATTTTCAAT CAGCTAAATT TGAAAATGGG TTTACGTTGC AAGATACACC 1200
 AGGTGTTGAT TCAAATGTTG CATCACATCA GTCAATAACA GAACAATATA TGTATACAAG 1260
 TAATATGATA TTTTATACGG TTGACTATAA CCACGTTCAA TCTGAACTTA ACTTTAAGTT 1320
 10 TATGAAGCAT ATAAATGATG TTGGaATACC TGTGTGTGTT ATCATTAAATC AAATTGACAG 1380
 CATCCAAGAC GATGGAATTG TCATTCTCTA CGTnTTAAAT CTCGAGTTGG AAAAATCAAT 1440
 TGGC 1444

(2) INFORMATION FOR SEQ ID NO: 480:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6309 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 480:

GCAGCAGCTT ATCGTGAGTT ATCATTATTA TTACGTAGAC CTCCAGGTCG TGAakCATAC 60
 CCAGGTGACG TATTCTACTT ACATAGTAGA tTATTAGAAA GAGCAGCAAA ATTAAACGAT 120
 30 GACTTAGGTG GCGGTTCAAT TACTGCATTA CCAATTATCG AAACACAAGC TGGTGATATT 180
 TCAGCTTATG TACCAACAAA CGTTATTTCA ATTACAGATG GACAAATCTT CTTACAATCT 240
 GATTTATTCT TCTCAGGTGT AAGACCAGCG ATTAATGCCG GACAATCTGT ATCTCGTGTT 300
 35 GGTGGATCTG CACAAATTAA AGCAATGAAG AAAGTTGCTG GTACGTTACG TCTTGACTTA 360
 GCGTCATACA GAGAACTTGA ATCATTGCA CAATTGCGTT CAGACCTTGA TGAATTACT 420
 GCAAGTAAAT TAGAACGTGG TAAACGTACT GTTGAAGTCT TAAAACAAGA TCAAAACAAA 480
 40 CCATTACCAG TCGAACACCA AGTGTGATT ATTTATGCAT TAACAAAAGG ATATTTAGAT 540
 GATATCCAG TTGTAGATAT CACACGTTTT GAAGACGAGT TAAACCACTG GGCAGAATCA 600
 AATGCTACTG AACTGTTAAA TGAAATCAGA GAACTGGTG GCTTACCAGA TGCTGAGAAG 660
 45 TTTGACACAG CAATTAACGA ATTCAAAAAA AGCTTTAGCA AATCTGAATA ATAAACAAGT 720
 TTAGTATAAG GTGGTGAGAT AGTGGCTTCT CTAAAGAAA TAGATACTCG AATAAAATCA 780
 50 ACCAAAAAAA TGAAGCAGAT TACGAAAGCG ATGAACATGG TATCAAGTTC AAACTTCGT 840
 AGAGCTGAAA AAAATACAAA ACAATCACA CCATATATGG ATAAATGCA AGATGCAATT 900

	ACTAGAAGTG GATATTTAGT TATCACGAGT GATAAAGGTT TAGCAGGTGC ATATAGTGCA	1020
	AACGTGCTTA AAAAATTGAT TACTGATATT GAAGCGAAAC ATCAAGATAG TAGCGAATAC	1080
5	AGTATTGTAG TTTTAGGGCA ACAAGGTGTT GATTTCTTAA AAAATAGAGG TTATGACATT	1140
	GAGTATTCTC AAGTAGACGT ACCTGATCAA CCTTCTTTCA AATCTGTTCA AGCACTAGCT	1200
	AACCATGCTA TAGACTTATA CAGTGAAGAA GAAATTGATG AATTAAATAT ATACTATAGT	1260
10	CATTATGTCA GCGTTCTTGA AAACAAGCCT ACATCTAGAC AAGTATTACC ATTATCTCAA	1320
	GAGGATTCTA GTAAGGGGCA TGGTCATTTG TCTTCTTATG AATTTGAGCC AGATAAAGAA	1380
	TCTATCTTAA GTGTAATCTT GCCTCAATAT GTTGAGAGTT TGATTTACGG AACAATATTA	1440
15	GACGCAAAAG CAAGTGAGCA TGCAACACGT ATGACTGCGA TGAAAAATGC CACTGATAAT	1500
	GCAACTGAAC TTATTGATGA CTTATCATTGA GAATATAACA GAGCGAGACA AGCAGAAATT	1560
20	ACGCAACAAA TTAAGTAAAT TGTTGGTGGT TCCGCAGCGC TTGAATAATA TTTAAAGGAG	1620
	GAAAAATAGCA TGGGAATTGG CCGTGTAACCT CAAGTTATGG GTCCTGTAAT TGATGTTCCGA	1680
	TTTGAACATA ACGAAGTTCC TAAAATTAAT AACGCCTTGG TTATTGATGT GCCTAAAGAA	1740
25	GAAGGTACAA TACAACCTAAC ATTAGAAGTT GCGCTGCAAT TAGGTGACGA CGTTGTTCCGT	1800
	ACAATTGCGA TGGATTCAAC TGATGGTGTC CAAAGAGGCA TGGATGTAAA AGATACAGGC	1860
	AAAGAAATTA GTGTACCTGT TGGTGACGAA ACATTAGGTC GTGTATTTAA TGTACTAGGT	1920
30	GAAACAATTG ACCTTAAAGA AGAAATTAGT GATTCTGTTC GCCGCGATCC TATCCATCGT	1980
	CAAGCACCAG CATTGATGA ACTTTCAACA GAAGTTCAAA TTTTAGAAAC AGGTATTAAA	2040
	GTAGTAGATT TACTAGCACC TTATATTAAA GGTGGTAAAA TCGGATTGTT CGGTGGTGCC	2100
35	GGTGTAGGTA AAACAGTATT AATCCAAGAA TTAATTAACA ACATCGCTCA AGAGCACGGT	2160
	GGTATTTCTG TATTCGCCGG TGTAGGTGAA CGTACTCGTG AAGGTAACGA TTTATACTTC	2220
	GAAATGAGTG ACAGTGGTGT AATTAAGAAA ACAGCCATGG TATTCGGGCA AATGAATGAG	2280
40	CCACCTGGTG CACGTATGCG TGTTGCATTA TCTGGTTTAA CAATGGCTGA ATATTTCCGT	2340
	GACGAACAAG GTCAAGACGT ATTATTATTC ATCGATAACA TTTTCAGATT TACACAAGCT	2400
45	GGTTCTGAGG TATCTGCATT ATTAGTTCGT ATGCCTTCTG CAGTAGGTTA CCAACCAACA	2460
	CTTGCTACTG AAATGGGACA ATTACAAGAA CGTATTACGT CTACAACAAA AGGATCAGTT	2520
	ACTTCTATTG AAGCGGTATT CGTACCTGCC GATGACTATA CTGACCCAGC GCCTGCGACA	2580
50	GCGTTTGCCC ATTTAGATGC AACTACAAAC TTAGAACGTA AATTAAGTGA AATGGGTATT	2640
	TATCCAGCCG TGGATCCATT AGCGTCTACA TCAAGAGCAT TGGAACCATC AATTGTAGGT	2700

55

	CAAGATATCA TTGCTATCTT AGGTATGGAC GAATTATCTG ATGAAGATAA ACAAACAGTT	2820
	GAACGCGCAC GTAGAATTCA ATTCTTCTTA TCTCAAACT TCCACGTAGC GGAACAATTT	2880
5	ACTGGTCAAA AAGGTTCTTA TGTACCTGTT AAGACAACAG TTGCAAATTT TAAAGATATC	2940
	TTAGATGGTA AATATGACCA TATTCCAGAA GATGCATTCC GTTTAGTTGG TAGCATGGAT	3000
	GATGTTATTG CAAAAGCTAA AGATATGGGT GTTGAAGTAT AACAAATTAGG AGGAATGGAT	3060
10	AATGAATACA TTAAACCTAG ATATTGTCAC TCCTAATGGT TCTGTTTACA ATCGTGATAA	3120
	TGTTGAACTC GTTGTATATGC AAACAACAGC TGGTGAGATA GGTGTCATGA GTGGACATAT	3180
	TCCAACGTGA GCTGCTTTAA AAACAGGCTT TGTAAGAGT AAATTTTCAG ATGGAACCTGA	3240
15	ATATATTGCT GTAAGCGATG GCTTTGTTGA AGTTAGAAAA GATAAAGTTT CAATCATTGT	3300
	TCAGACTGCA GAAACTGCAA GAGAAATTGA TGTTGAAAGA GCTAAATTAG CCAAAGCAAG	3360
20	AGCAGAGTCT CACTTGGAAA ATGATGACGA CAATACTGAT ATTCATAGAG CCGAAAGAGC	3420
	TTTAGAGAGA GCAAATAACC GTTTGCGTGT GGCTGAATTA AAATAGTAAA TAAAGGGTCG	3480
	AAGATGTGAT TTCATATCTT CGACCCTTTT TTGAATTATA TTGATTTAAA GATACAAAAC	3540
25	ATGAGAGGGG GGAAGGAATT GATAAAGAAC CATTAAAGAT TTATGATGTA GTGGTTCTTT	3600
	ATCATTAAC ACAGCTAATG TGTATTTAAA AATAGGaayA CATgAGTAAA ACTCATGTAT	3660
	AAGAAATACT AATTTCTAAA GAAAAAGTAT TTCTTTATGT TGGGGCCCCG TCAACTACTG	3720
30	CCAAATACAA CACTATAGAG TCTAGACATT GATTTATGTC CGACTCCCAA GAATAGTTTT	3780
	ACTTTTTTAC AATCACTAAT AGATTGCTAA AATCAAAATT TCCTTCACCA CTATCTACAG	3840
	TCGACATTC ATTTTTTGAA ATTATCTACA TTTTTTCATA CCAAGATATT TTATAGTTAT	3900
35	GATATTTATG TAAAAAGAAT TATATAGTAA GTTAGCTTAA ACTTTACTAA AAACGGGTAT	3960
	TAACTTTGT ATCATTATTT AAATTTTTCA TGTACAATGT AATACAGTAA TCTTATGAGG	4020
40	TGATAAAATG GATTATATCG GACAATATGC AGTTATCCAT TTAGTGTTAC ATGTTGTATG	4080
	TATTTGTATT GCCTATTGGG CTTTACAATC AATTAGATTA GATCAATTTT TTAAAAAAGG	4140
	ATACGCCACT CAATTACAAG TGTGTATGAT ATTTGTTGCT ATTTTATTAG GCACTGCAGT	4200
45	AAGCAATTTT ATTGTAGATT TGTTACAATA CTCGACGCAG GTAAAATATT TAATAAAATA	4260
	AGTCTAACTC TATGATTTGT AATCAAACT AGATATAATT AAATAATGAC TTAAAAATAAT	4320
	TTTAAAAATAG GGAAATGTAA AGTAATAGGA GTTCTAAGTG GAGGATTTAC GATGGATAAA	4380
50	ATAGTAATCA AAGGTGGAAT TAAATTAACG GGTGAAGTTA AAGTAGAAGG TGCTAAAAAT	4440
	GCAGTATTAC CAATATTGAC AGCATCTTTA TTAGCTTCTG ATAAACCGAG CAAATTAGTT	4500

55

	GACGTTACAT ACAAAAAGGA CGAAAATGCT GTTGTGCTTG ATGCAACAAA GACTCTAAAT	4620
	GAAGAGGCAC CATATGAATA TGTTAGTAAA ATGCGTGCAA GTATTTTAGT TATGGGmCCT	4680
5	CTTTTAGCAA GACTAGGACA TGCTATTGTT GCATTGCCTG GTGGTTGTGC AATTGGAAGT	4740
	AGACCGATTG AGCAACACAT TAAAGGTTTT GAAGCTTTAG GCGCAGAAAT TCATCTTGAA	4800
	AATGGTAATA TTTATGCTAA TGCTAAAGAT GGATTAAAAG GTACATCAAT TCATTTAGAT	4860
10	TTTCCAAGTG TAGGAGCAAC ACAAATATT ATTATGGCAG CATCATTAGC TAAGGGTAAG	4920
	ACTTTAATTG AAAATGCAGC TAAAGAACCT GAAATTGTCTG ATTTAGCAAA CTACATTAAT	4980
15	GAAATGGGTG GTAGAATTAC TGGTGCTGGT ACAGACACAA TTACAATCAA TGGTGTAGAA	5040
	TCATTACATG GTGTAGAACA TGCTATCATT CCAGATAGAA TTGAAGCAGG CACATTACTA	5100
	ATCGCTGGTG CTATAACGCG TGGTGATATT TTTGTACGTG GTGCAATCAA AGAACATATG	5160
20	GCGAGTTTAG TCTATAAACT AGAAGAAATG GGCCTTGAAT TGGACTATCA AGAAGATGGT	5220
	ATTCGTGTAC GTGCTGAAGG GGAATTACAA CCTGTAGACA TCAAACTCT ACCACATCCT	5280
	GGATTCCCGA CTGATATGCA ATCACAATG ATGGCATTGT TATTAACGGC AAATGGTCAT	5340
25	AAAGTCGTAA CCGAACTGT TTTTGAAAAC CGTTTTATGC ATGTTGCAGA GTTCAAACGT	5400
	ATGAATGCTA ATATCAATGT AGAAGGTCGT aGTGCTAAAC TTGAAGGTAA AAGTCAATTG	5460
	CAAGGTGCAC AAGTTAAAGC GACTGATTTA AGaGCAGCAG CCGCCTTAAT TTTAGCTGGA	5520
30	TTAGTTGCTG ATGGTAAaAC AAGCGTTACT GAATTAACGC ACCTAGATAG AGGCTATGTT	5580
	GACTTACACG GTAAATTGAA GCAATTAGGT GCAGACATTG AACGTATTAA CGATTAATTC	5640
35	AGTAAATTAA TATAATGGAG GATTTCAACC ATGGAAACAA TTTTGTATTA TAACCAAATT	5700
	AAACAAATTA TACCTCACAG ACAGCCATTT TTATTAATTG ATAAAGTAGT TGAATATGAA	5760
	GAAGGTCAAC GTTGTGTGGC TATTAAACAA GTATCAGGAA ACGAACCATT CTTTCAAGGG	5820
40	CATTTTCCTG AGTATGcGGT AATGCCAGGC GTATTAATTA CTGAAGCGTT ActCAAACAG	5880
	GTGCGGTAGC TATTTTAAAT AGTGAAGAAA ATAAAGGTAA AATCGCTTTA TTTGCTGGTA	5940
	TTGATAAATG TCGTTTTAAA CGTCAAGTAG TACCTGGTGA TACTTTAACG TTGGAAGTAG	6000
45	AAATCACTAA AATTAAAGGA CCAATAGGTA AAGGTAATGC TAAAGCTACT GTCGATGGTC	6060
	AACCTGCTTG TAGTTGTGAA CTTACATTTG CAATTCAAGA TGTAATAATA AACAAAAAA	6120
	ACATTCAAAG ATTTAATGTG TTGGCATAAT CTTTGAATGT TTTTATTTT ACTCTTCTAA	6180
50	TTTTTCATCC TTTAACTTTG GTTTAGACTG CaTCATTGCA TTAAATGATT TTTTAAATTC	6240
	TTCACCAGAT AATCCATCAT CAATAAGTTG GTTCTAATAA ACTTTCAGCA TACTGTTGGA	6300

55

(2) INFORMATION FOR SEQ ID NO: 481:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1717 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 481:

AGTTGCTACA CCAGACATGA TGGGTGAAGT TGGTAAATTA GGTGCTGTAT TAGGACCAAA	60
AGGTTTAATG CCAAACCCTA AAAGTGAAC TGTAACAATG GATGTTAAAA AAGCTGTTGA	120
AGAAATCAAA GCTGGTAAAG TAGAATATCG TGCTGAAAAA GCTGGTATCG TACATGCATC	180
AATTGGTAAA GTTTCATTTA CTGATGAACA ATTAATtGAA AACyTCaATA CyTTACAAGA	240
TGTATTAGCT AAAGCTAAAC CATCATCTGC TAAAGGTACA TACTTCAAAT CTGTTGCTGT	300
AACTACAACA ATGGGTCCTG GAGTTAAAAT TGATACTGCA AGTTTCAAAT AATAAATGAT	360
ATAACAATT ACAGGCTGAA AGAAATATCT TTCAGTCTGT AAAAATATAT TGACAATAAG	420
TAATTTCCAA GTTATATTAC TTATTGTGAT TATTTTACCT AAGACAGTAG GAGTTATTTA	480
TAACTTAAAA TTTATCCTGC CGAGGCTAAA ATTGACTTGA ACGTGATGAT CTATGATCTT	540
TCAAGCACTT TTTGCCGTGG GTAGAAAGTG CTTTTTTTAT TAATTTTAAA AAAAGCACCA	600
AAAATTTAAA TGGAGGTGTC TGAATGTCTG CTATCATTGA AGCTAAAAAA CAACTAGTTG	660
ATGAAATTGC TGAGGTACTA TCAAATTCAG TTTCAACAGT AATCGTTGAC TACCGTGGAT	720
TAACAGTAGC TGAAGTTACT GACTTACGTT CACAATTACG TGAAGCTGGT GTTGAGTATA	780
AAGTATACAA AAACACTATG GTACGTCGTG CAGCTGAAAA AGCTGGTATC GAAGGCTTAG	840
ATGAATTCTT AACAGGTCCT ACTGCTATTG CAACTTCAAG TGAAGATGCT GTAGCTGCAG	900
CGAAAGTAAT TTCTGGATT TCTAAAGATC ATGAAGCATT AGAAATTAAA TCAGGCGTTA	960
TGGAAGGCAA TGTTATTACA GCAGAAGAAG TTAAACTGT TGGTTCATTA CCTTCACACG	1020
ATGGTCTTGT ATCTATGCTT TTATCAGTAT TACAAGCTCC TGTACGCAAC TTCGCTTATG	1080
CGGTAAAGC TATTGGAGAA CAAAAGAAG AAAACGCTGA ATAATTTTTA GCGTAAAAAA	1140
ATTAAAAATA ATGGAGGAAT TATAAAATGG CTAATCATGA ACAAATCATT GAAGCGATTA	1200
AAGAAATGTC AGTATTAGAA TTAAACGACT TAGTAAAAGC AATTGAAGAA GAATTTGGTG	1260
TACtGcAGCT GCTCCAGTAG CAGTAGCAGG TGCAGCTGGT GGCCTGACG CTGCAGCAGA	1320
AAAAACTGAA TTTGACGTTG AGTTAACTTC AGCTGTTCA TCTAAATCA AAGTTGTTAA	1380

TCCTAAAGTA ATCAAAGAAG CTTTACCTAA AGAAGAAGCT GAAAAACTTA AAGAACAATT 1500
 AGAAGAAGTT GGAGCTACTG TAGAATTAAA ATAATTCAAG TATCTTAAAC TTAATAATCA 1560
 5 AAGTTTTATA GCAAGTATTG CTATAATATA ATGATTCTTT GAGAAGTTAA AACCCCGTTA 1620
 TTTTGATAAC GGGgTtTTAT TCaTTTAAAG ACTGAGTGAA ATGTTATAAT TATAATGACG 1680
 AGTTACAAAG TGAAGATGAG GTGGGAATAA TGAGTCA 1717

10 (2) INFORMATION FOR SEQ ID NO: 482:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1279 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 482:

GTAAATCTGT TACTCGAAAT GTAACGATTA AAGAAAAGGG CTCATCTCAA ACATATATTT 60
 TGTTAGGCTA TCCAACAAAA GCACAGAAGA ATAGTCATAG CAAATATAGT GGAGTCTTTA 120
 25 TATATAAAGA CTTGAAATCA ATCGAAGATA CAAATAATGC TATTACGATT ATCACCATAA 180
 TTACGGCTGT TATTTTCTTA ACAATTACAA CAGTCTTTGC GTTTTCTTTC TCGTCAAGAA 240
 TTACAAAACC TTTAAGACGT TTAAGAGACC AAGCTACACG TGTATCTGAA GGGGATTACT 300
 30 CTTATAAACC TTCTGTCACA ACGAAAGATG AAATTGGTCA ATTATCGCAG GCATTTAATC 360
 AGATGAGTAC AGAAATCGAA GAGCATGTCG ACGCATTATC CACATCTAAA AATATTAGAG 420
 ACAGCTTAAT TAACTCTATG GTAGAAGGTG TCCTAGGTAT TAATGAGAGT CGACaAATTA 480
 35 TCTTATCTAA TAAGATGGCG AATGATATTA TGGACAATAT TGATGAAGAT GCTAAAGCTT 540
 TCTTATTAAG ACAAATAGAA GATACTTTTA AATCAAAACA AACTGAAATG CGCGATTTAG 600
 AAaTGAATGC ACGATTCTTT GTTGTGACCA CAAGCTATAT CGACAAGATT GAACAGGGAG 660
 40 GTaaAAGTGG TGTGTGTGTG ACAGTTCGTG ATATGACTAA TGAGCACAAT CTAGATCAAA 720
 TGAAGAAAGA TTTCAATGCT AATGTATCAC ATGAATTACG TACACCGATA TCATTACTTC 780
 45 AAGGTTATAC TGAATCAATT GTAGATGGTA TTGTTACAGA ACCGGATGAA ATAAAAGAAT 840
 CGCTTGCCAT TGTCTTGAT GAATCGAAAC GTTTAAATCG TTTAGTTAAT GAATTGTTAA 900
 AIGTCGCACG CATGGATGCT GAAGGGTTAT CCGTAAATAA AGAAGTTCAG CCTATTGCAG 960
 50 CGTTACTAGA TAAGATGAAA ATTAAGTATC GCCAACAAGC TGATGATTTA GGTCTAAATA 1020
 TGACTTTTAA TTAYTGTAAG AAGCGTGTTC GGAGTTATGA TAwGGATCGC ATGGACCAAG 1080

TTACTTGTGA TGAAAATGAA AGCGAAGATA TTTTATACAT TAAAGATACA GGTACAGGCA 1200
 TTGCACCAGA ACATTTACma CAAGTATTTg ATCGTTTTTA TAAAGTTGAT GCAGCGnAnA 1260
 5 ACCCCGnGGT AACCAngTA 1279

(2) INFORMATION FOR SEQ ID NO: 483:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1144 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 483:

GAGCTGTTGT TACTTTGATG CCTGCAGCTT TATTACGGCT GACTTGGTAA TGATAAGTTT 60
 20 CAGCATATTG CTCAATATAT GCTATATCAT ATTGAATGGT ACGAGGTGAT ACACCAAGTT 120
 GATTAGCAAT GGTATTGATT GGAATAAACG TTTGCTCATG AATTAAAAGA TACAAAATTT 180
 CGATTTGTCT ATAACCTAAC AACGTAATAT CCTCCTATTT GTAATTGTAA GCGATTTCTT 240
 25 AAAAACGTAG ATATGCAATC TCTTTCATAT TTTAATCCGA AAAATTGCAT ATCAAAATGT 300
 TTATGGCGCA AGATTTTATA GGAACCTTTTA AAATAAATTA TATATTCATG TTGACAATTT 360
 AAAAATGTCG CAGTATATTT AGTTAGACAT CTAACGAAAT GGTGGTGCAA TAAATGGAAT 420
 30 TCACTTATTC GTATTTATTT AGAATGATTA GTCATGAGAT GAAACAAAAG GCTGATCAAA 480
 AGTTAGAGCA ATTTGATATT ACAAATGAGC AAGGTCATAC GTTAGGTTAT CTTTATGCAC 540
 ATCAACAAGA TGGACTGACA CAAATGATA TTGcTAAAGC ATTACAACGA ACAGGTCCAA 600
 35 CTGTCAGTAA TTTATTAAGG AACCTTGAAC GTAAAAAGCT GATCTATCGC TATGTCGATG 660
 CACAAGATAC GAGAAGAAAG AATATAGGGC TGACTACCTC TGGGATTAAA CTCGTAGAAG 720
 CATTCACTTC GATATTTGAT GAAATGGAAC AAACACTCGT ATCGCAGTTA TCTGAAGAAG 780
 40 AAAATGAACA AATGAAAGCA AACTTAACTA AAATGTTATC TAGTTTACAA TAAATGATAA 840
 GTGTGACTGG TAGAAATCAG TCACTTTGTC TTTAATATTA TAGTTAGATA TCTAATTGTT 900
 45 AGTAAGCTAA TTATTGAAA AGACAAGGAG TATTGAACAA TGAAAGACGA ACAATTATAT 960
 TATTTTGAGA AATCGCCAGT ATTAAAGCG ATGATGCATT TCTCATTGCC AATGATGATA 1020
 GGGACTTTAT TAAGCGTTAT TTATGGCATA TTAAATATTT ACTTTATAGG ATTTtTyAGAm 1080
 50 GAYAGCCACA TGATTTCTGC tAatCTCTCT AACACTGCCA GTATTTGCTA TCTTAATGGG 1140
 GTTA 1144

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1158 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 484:

10 ACAACCAATT TTAATAACA TGGTrrTCAT gCATTITCAA TATTTAGATG TCGACCGCAT 60
 AATCGAAGAA TCGCCGACAA TAGTACTTAT CGATGAGTTA GCACATACGA ATATTCTAG 120
 15 AGATCGTCAT GAGAAACGAT ATATGGATAT TGAAGAAATT TTAAATCATG GTATCGATGT 180
 TCATACCACT TTGaACATTC aaCATATTGA AAGTTTAAGT AGTCAAATTG AACTGATGAC 240
 CGGTGTACAT GTTAAAGAAC GTGTACCCGA CTATTTTATA ATGAGCGCCG ATGTATTAGA 300
 20 AGTCGTAGAT ATCTCACCTG AACAAATTAAT TAAACGCTTA AAAGCTGGCA AGGTATATAa 360
 AAAGGATAGG CTAGATGTAG CATTTAGTAA TTTCTTTACG TATGCCACC TAAGCGAaTG 420
 CGTACATTGA CGTTAAGAAC AGTTGCCGAC TTGATGAGTG ATAAAGAAAA AGTCCGACAC 480
 25 AACCATAAAA CGTCACTCAA ACCTCATATT GCTGTGGCAA TTAGTGGGAG CATTTATAAT 540
 GAAGCAGTAA TTAAGAGGC ATTCCATATT GCTCAAAAAG AACATGCGAa GTTCACTGCT 600
 ATTTATATAG ATGTATTCGA AAAAAACAGG CAATATAAAG ATAGTCAAAA GCAAGTGCAT 660
 30 CAACATCTCA TGCTTGCAA ATCATTAGGA GCAAAAGTAA AAGTAGTTTA TAGCCAAACC 720
 GTTGCAATTAG GATTAGACGA ATGGTGTAAG AATCAAGATG TAACCAAATT AATTATCGGA 780
 CAACATATTA GAAATAAGTG GCGAGACTTT TTCAATACAC CTTTAATTGA CCATTTAATG 840
 35 TCCTTTGAAC ATAGCTATAA AATCGAAATC GTTCCAATCA AACAAATACC TGTGAATTG 900
 AAAATGAACA AATCACCTTA TCGTCCTAAA GGCAAACGTT TCGCCATAGA TATGTTAAAA 960
 ATGATTTTGA TTCAAATAAT TTGTGTAATG ATGGGACTGT GGATTTATCA ACTTGATAAG 1020
 40 CATGAGTCTA GTACGATTAT TTTAATGATT TTTCTCATCG GCATCATTTT ATTATCCATT 1080
 TGGACGCGGT CCTTCATCAT TGGCTTTTaG CAGCAATTAt TAACGTATTT GTgTkTAATT 1140
 45 ATkTTTTtAC GGAACCTA 1158

(2) INFORMATION FOR SEQ ID NO: 485:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2224 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485:

	ATCATATGGT CGATTTAACA GATCCAACGT ACTGCTAAAT AATTACATGA CGTTTAAACA	60
5	TAGCATTGAT TATAACTATT TCTAAGTCTT CGCATTATTT GCGATGATGT GCGAATAGTT	120
	ATTTTTATTT AAAAATATAA AAAAATAGAT GCAGCAAAAT TTAAAGCAT TTTATTTTGA	180
	ACATATTAAA AGGGAGCGTA TCATAATGGA ATGTAATGTT TATATCGTAT GCATTACGGA	240
10	TAAATAATAT ATAAATCATT CTTGAGGAGT GAAAGAATAA TGAGAGACTA CACAAAGCAA	300
	TACATTAATG GCGAATGGGT AGAAAGTAAT AGTAATGAAA CGATAGAAGT TATAAATCCA	360
	GCAACCGAAG AAGTAATCGG GAAAGTTGCT AAAGGTAATA AAGCTGATGT TGATAAGCC	420
15	GTCGAGGCGG CAGACGATGT TTATTTAGAG TTCCGTCATA CATCTGTGAA AGAAAGACAA	480
	GCGTTATTAG ATAAAATTGT AAAAGATTAT GAAAACAGAA AAGACGATAT TGTACAAGCT	540
20	ATTACGGATG AATTAGGTGC TCCTTTATCA TTATCTGAGC GTGTCCATTA TCAAATGGGA	600
	CTAAACCATT TTGTTGCAGC GAGAGACGCA TTAGATAACT ACGAATTTGA AGAACGCCGC	660
	GGAGATGATT TAGTTGTTAA AGAAGCAATC GGTGTATCTG GATTAAATTAC ACCGTGGAAC	720
25	TTCCCTACAA ACCAAACATC ATTAAATTA GCAGCAGCAT TTGCGGCTGG TAGTCCAGTT	780
	GTACTIONAAC CATCTGAAGA AACACCATTT GCAGCTGTTA TTTTAGCTGA GATTTTTGAT	840
	AAAGTCGGTG TTCCTAAAGG TGTATTAAAC CTTGTTAATG GTGATGGTGC TGGTGTGGG	900
30	AATCCTTTAT CTGAACATCC TAAAGTACGC ATGATGTCAT TTACAGGATC AGGCCCTACT	960
	GGTTCTAAAA TTATGGAAAA AGCCGCTAAA GATTTTAAAA AGGTATCATT AGAGCTTGGT	1020
	GGCAAATCAC CATATATCGT CCTAGATGAC GTAGATATTA AAGAAGCGGc TAAAGCAACa	1080
35	aCAGGCAAAG TTGTTAATAA TACTGGTCAA GTATGTACAG CTGGTACACG TGTTTTAGTG	1140
	CCTAACAAAA TTAAAGATGC ATTCTTAGCT GAATTAAAAG AACAAATTTAG CCAAGTGCCT	1200
	GTCGGTAATC CAAGAGAAGA TGGTACACAA GTAGGCCCTA TCATTAGTAA AAAACAATTT	1260
40	GATCAAGTAC AAAATTATAT TAATAAAGGT ATTGAAGAAG GTGCTGAATT ATTTTATGGT	1320
	GGTCCTGGTA AACCAGAAGG ACTTGAAAAA GGATACTTTG CACGTCCGAC AATTTTTATT	1380
45	AATGTAGATA ATCAATGAC GATAGCACAA GAWGAAATTT TTGGGCCAGT AATGTCAGTT	1440
	ATCACTTATA ACGATTTAGA TGAAGCGATT CAAATTGCAA ATGATACAAA ATATGGTTTG	1500
	GCAGGATATG TTATTGGTAA GGACAAAGAA ACATTGCATA AAGTAGCTCG TTCTATTGAA	1560
50	GCAGGTACAG TAGAAATAAA CGAAGCAGGT AGAAAGCCAG ATTTACCATT TGGTGGCTAT	1620
	AAACAATCTG GTTTAGGTCTG TGAATGGGGC GATTATGGTA TTGAAGAGTT CTTAGAAGTG	1680

55

AGTGCACATG ACTAATTAAG TTTTGTGTAC TGTTTTAATT TTGCAATTTT TATAAATAGA 1800
 TTTTGTAAAT AAAATAAAAA TTTGCTATAG TTATTCATGT ATTTAAAAGG TTGGGGATTA 1860
 5 GCATAATGGG ATTGTGCTAG CACAGTTATT TATGCATTGT CATGCCTATC TATTACTTAC 1920
 TAACTAAAAA ATAATGAAAT GGGTGTAAC TATATGCCTG AAAGAGAACG TACATCTCCT 1980
 CAGTATGAAT CATTCCACGA ATTGTACAAG AACTATACTA CCAAGGAACT CACTCAnAAA 2040
 10 GCTAAACTC TTAAGTTGAC GAACTATAGT AAATTAAATn AAAAAGAACT TGTTCCTAGCT 2100
 ATTATGGAAG CACAAATGGA nAAAGATGGT AACTATTATA TGGAAGGTAT CTTAGATGAT 2160
 ATACAACCAG ATGTTTATGG TTTTTTAAGA ACAGTGAAC ATTCTAAAGG GGAAAAAGAT 2220
 15 ATTT 2224

(2) INFORMATION FOR SEQ ID NO: 486:

(i) SEQUENCE CHARACTERISTICS:
 20 (A) LENGTH: 1690 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 486:

ACATTACnTT GAATCGAAAG TTTcATAAC GACTTGAATG CCAGTTTCTT TTTCAAATTT 60
 30 CTTAATTAAC TCTGGATCAA TATATTGCGC CCAATTGTAT ACGTAAATTT TTTGATTTGT 120
 ATGCACTTGT TCTTTAGATT TAAACCAATG ACTTAAAGTA AGACAAAGCA TACCCACAAC 180
 TAATGCACCT ATAATGAGTT GTAAAAATCG TTTcATTATT TTACACCTCG CTTGATTAGT 240
 35 TTTTCTTAT TTATCACGTA TTGAATCAAA TAATATCCTA GTATTCCTAA TACAATAACA 300
 GCAAACAATA ATGTTGAAAT CGCATTAAAT TCCATACTAA TTCCTTTTCT CGCCATAGCA 360
 TAACTTCAA CTGATAACAC ACTAAAGCCA TTACCAGTAA CGAAGAACT TACTGTGAAA 420
 40 TCGTCTAGTG AATAAGTTAA AGCCATAAAG AATCCTCCTA TAATAGAAGG TAAAATATTA 480
 GGAATAATAA TGTGCTTAA TAATTGTGGT TCAGTCGCTC CTAAATCTCT TGCAGCATTT 540
 AACATATTAT TATTCATyTC ATACAGTTGT GGTAAGACGA TAATCACAAC TATAGGTATG 600
 45 CAAAATGCAA TATGAGATAT TAGAACTGTC CaaAAKCCTA AACCAAGACC AGTAAAATGG 660
 CCAATCGTTG TAAACATAAT TAAGAATGAT GCACCTATGA CAACGTCGGA TGATACCATC 720
 AAGACATTAT TCAATGTTAG TAAAGTTACT TTAACTTTT TATTTCTTAA ATAATAAATA 780
 50 GCAATGGCAC CAAATGTACC AATAACTGTA GAAATTGAGG CTGCTAAAAG TGCTACAGCT 840

55

AATGTAAAAT GTTCAAAGTG AATCATATTA CCAGCCGAAT TGAATGAATA GAACATTAAA 960
 AAGAATATTG GGATGTATAA AATCGCTAAA AGTATCCCGA TATACAGCTT TCCATACCAT 1020
 5 TTCATATGAT TCACCCTCTC CCATTAGATG ATTTTGTAAAT GATTAAAATG AATGCCATAA 1080
 ATACAATTAA GAATATAGCT ATAGTTGATC CCATACCATA ATTTTGAATT GTTAAAAATT 1140
 GTTCCTCTAT TGCCGTACCT ATATTTATGA CTTTATTACC TGCAATTAAT CTTGTAATCA 1200
 10 TAAATAATGA AAGTGATGGA ATAAAGGTTA CTTGAATCCC AGTCATAACA CCTTCTTTTG 1260
 TTAACGGCAT GATTACTTTT CTAAAAGTAT AGAAAGGACT GGCACCTAAA TCACTTGAGG 1320
 CCTGCAATAA ATTATTAGGA ATTGCTTTCA TGCTATTAAA TATAGGTAAA ATCATAAATG 1380
 15 GTATATAAAT GTAACCTGCC ACTACTAAAA ACGCACCAGT TGTAAATAAC AAATTGAATG 1440
 ATGGTAAATT AAATAAGTGG AAAATTGATT AATCAGCCA TCATGACTTA ATAAACCTAT 1500
 AAAAGCATAT GTCTTTAACA ATAAATTAT CCATGTTGGA ATAATCATT TCAATTAATA 1560
 20 GATATTTTGA AATTTGGAAC GAGTAATATA ATAGGCAGT GGATAACTGA TAGTCAAGGT 1620
 AATAATTGTT ATTGAAGCGG CATATAAAAT TGAATATGCA AACATTTTCA AATATTTTGT 1680
 25 AGTAAAAATT 1690

(2) INFORMATION FOR SEQ ID NO: 487:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2112 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 487:

ACGAAAAGAA ATATTATGAT GAACAAAAAG AAAGAATAAC GATTATATG AAGTACAATG 60
 TGAAAGGTTA TAAAAATATA AGCTTCGCTA ATTTTAAAGA AAACCCAATG GATGGTTATT 120
 40 CTATTAGTGG TTATATAAAT AatGaTAAAA AGTTATCATT TACAGCTGGT ATAAGATCTG 180
 TTGATGATTT TCAATTTGAT ACCGATATTT CTTATACAGA TGAATTGGGT AGAAAATTTA 240
 ATAAAAATCC TAAGTCAGTT TCTGAAATAA AAAAAGAGCA AAATACGTCC AATAAATAAT 300
 45 TGTTCAATATT GTGATGAAAC AAAAATATAA GTCATTAGAT GAGTTTAACT ATGTTATAAA 360
 TATTTGTAGT ATCTATAAAA ATCTCGACAC TATTAAAATG ATAAAGTGCC GAGGTTTTCT 420
 TACTTATTTA GTTAATTCAA AGTTTATGCC AGATTATAA GAATTTGTGA CACTTTTAAT 480
 50 AGTGTAACCAT TGATTATTAC AATTTATCAA ATGGTCCTTT AGAAGGTATA AATAACAAAA 540

	TATTATTATG TTCAAACTT TACGCTCCAA AAAGTAAAAA GGAAGTTAAG CAATGTTTAG	660
	TTGCTTAaCT TCGGaTATTG AACGCATCAG TCCAATTTGA CATAGAGCCT TTTTtagTTC	720
5	TTGATGTTTC TCTTTAAaAC CTTGCATATT TTACAAAAAG AAAAATTAGC AGTATAATTA	780
	AGACAACGAA AATAAGTATT TACTTATACA CCAATCCCCT CACTATTTGC GGTAGTGAGG	840
	GGATTTTTAT TGGTGCGGCT ATATGTCACC TATTTTGTAT TCGGTCTACT TAGCCAATAA	900
10	GAAAAAACG CAATGGCACA GCCACTGATG ACTGGTGCTA TGATGTGAAC GaAAATAAGC	960
	ATCACCTTAT ACACCTCCTC TCTGCGTCTA AATTGAAGsC TGAGaGrTAG GcGACTCTAC	1020
	TATTATATCA TCGGCAaATA TACAAGCACA GTCACCTTGCT TCTGATAAGT TATATGATTC	1080
15	TAGCTGATAG ATTGAATCGT CTACACTTAA TTGGACAAAT TCTATGAGAA TAGATATTGT	1140
	TAATTTAAGA AAGTAGGCGA TTTTATTATG ACAAGAGAAA GAAGATCATT TAGTTCAGAG	1200
20	TTTAAGTTAC AAATGGTTAG ATTATATAAA AATGGTAAGC CTAGGAATGA AATTATACGC	1260
	GAGTATGATT TCACACCTTC GACGTTTGTA AATGGCGGTT ATAAAATGTA GGAAATGGA	1320
	TAAAGCAACA TCAAAACACG GGTACATTCA ATCACCAAGA TAACTTATCG GATGAAGAAA	1380
25	AAGAGCTGAT TAAATTACGC AAAGAAGTTC AACATTTAAa AATGGAGAAC GATATTTTAA	1440
	AGCAAGTAGC GCTGATTATG GGGCAAAAAT AGAAGTCATT CAAAAGAATG CACATCAATA	1500
	TTCAGTATCA GCAATGTGTA AAGTCCTGAT AATACTAAGA AGTACCTATT ATGATTCTAT	1560
30	AAAAAGAAAA GATAATAAAA TCACTAAAGA TGATTCAaAC ATAGAACATG CCGTCATAAA	1620
	TATTTTTAAT TCTAATAGAA AAGTCTTTGG TACAAGACGA ATTAAAAATC ATTTAAATGA	1680
	CAAGGGTCTC ACTGTATCTG GACAAAAGAT AGGTCGATCA TGAAAAATC TAGTTTCTGT	1740
35	TTATACGAAA TCTAAATACA AAAATCATCT AAAAGAACT AATGAAAAC GAATTAAAAA	1800
	TCTTTATTAT TAGCTGCTGG TGTATTATTT GTTAGTCCAA TTTCATTATC TTATAATTCA	1860
	GATGTAGCTC ATGCTGAAGA TAAGTTAGAC CATCTCAAG CAAAGGTAAT ATATTTGAGT	1920
40	AACCAAAATT TATTTGATGA ACTTGAGAAA AAAGGTTATA AACTGGAAGA TATATTTACA	1980
	AAAGAAGAAA TAAAAAATA TAAAGCTGAA GACCAATTGA GAGCGGGTAA AACTCAATAT	2040
	GTAGAAACAG GTAAAGATAC TGCAACATTA TATCTTTCTT CTGCATATAC AAAACAATA	2100
45	GCTGCTTTAG GT	2112

(2) INFORMATION FOR SEQ ID NO: 488:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 454 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 488:

5 GTAGGCACAC ATCTGCCATA TAAACATTCT TTTATACTAG TGTTCTCATA TAGTGTAGAG 60
 TTATAGTCTC CTTCTTGAAT CTCGAATAAT TCAATCAACC TATCAACCTT AGTCTCTTCC 120
 GTTACTTCTT TTTCAATATC AACTATGAAG GGGATATCAA TTGGAATAAA ACTTGACGTC 180
 10 GAACACTTAT TTGTATTTGG ATGAAAACGA ACGAATCCAT CACTAAATCC TGTTGAAAAA 240
 AATATTTTTC CTTGTGATAG ATCCGGATTT TCTCGCGCCC ATTTAATTAA TTCATCTAAT 300
 CTCATTCTT TTTTAACTTT GATTTTCATT GTTATATCTC CTCTTGAACA GTAAATTTAT 360
 15 CGTTAACTGA TACGTATCCA GTCACATTAC ATAAGATGCT ATCAACATCA AAAGTCACAC 420
 AACAGTTGCG TTCAACATCA TTTGAATAGA ATCT 454

(2) INFORMATION FOR SEQ ID NO: 489:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489:

30 TTGTCAGAAT TAGAATGCTT TTGAGTTACT TCATAATACT CATCAGTTTT TTGTGTATCC 60
 TTTTGACTTT TATTTATTTT TTTCCACTTA CCAGTATGAC TTTCTTTTTT TACAGTTATT 120
 TTCGGTTTGT TAAAAAATG ATAACCGATG TTCTTTTTGG TATCTGTGGA CTCTAAAAAG 180
 35 ACTGAATTGT TTTCTGATT ATCAGAATTG GTTGTGTTGT TATCGTCTGT ATATAACGTA 240
 TACCCATTCT CTTTGGGATT TTCAATCGTT GTAACGGAT TCTTTGATGA ATCAGTACTT 300
 TTAATGCCAG TTCCTAAGAA GACAATTTTA TCGTTTAATA TGAAATATGA TTTTTTGGCA 360
 40 GTTAAAGTTT TGCTCTGATT TTCAAAATCC ATTCCGATAC TAGCATGTTG GTCATCAACT 420
 TTTGTCCGCG CAACAAAAGT TTTACTCGAC TTTTATCAT CCGTATCTTT TAATATTTCA 480
 TTGTCTAAAG TTGTGTACC TGATAAACGT TTCATATCGG CTGTCACCCA GAAGTTATCA 540
 45 TGATAGTGTT TGACATCGCT GTTATATAAA TAAGACATTC CAGCACCAGT GTGCCAACCT 600
 TTTAAATTCT CTCGGTTGAT ACTTTCATAG CGTGCTACGT TTTTCGACGT CATACTTAAA 660
 CCAATGCAA AGTCTAAGTC TTTGTTATGA TAGGTGACAC GATCCATGTC ATTATATATT 720
 50 TTAAGTTGTT GTGTTAATCC GTTTTTAGAA ATACTGTTAT CTGTCATTAA AGACTTCATT 780

ACTGAAGATT TGACAATCTT TTTATACTTA GCTTTTGTG AATCATCCAT GGCATCACTT 900
 AATCTCAACA ATGATTTTCAT TACTGTTGCA GATGCTGAGT GACTGGTTTC ATTTTCACGA 960
 5 CTGATAGCTC TACCTCGTGA TAAATCCATC ATTTACCTT TATAAATGAG TGGCATAAAT 1020
 CCGTCGTCAA TCCATGACTT TAAGTTGTA TCATTTTGGG TTTTATCATT AAAAGGTGTT 1080
 TCTTTTATCA TCGGCATCAT TTGAGAAATC CCTCTAAGA GTACAACGCC ATAAGCACCA 1140
 10 GTGTATGGAA CGTCTTGATG ATCAATGTAA GAGCCATCTT TATAAATCC ATTACGTTCT 1200
 TTACCAGTGG CAGAATCTTG AACGTAAGTG AAGACTTTAT TAAATGAATC TATAGACTTT 1260
 TTCATCATAT CTTTATCTTC TTCGATAATA CATTCTAAAA GTTTCACCTT AGAAATGTCT 1320
 15 ACTAnATTTC CGCCTTTAGC AAGTTCAGnT TTTCTACAC AAGATAATAT TT 1372

(2) INFORMATION FOR SEQ ID NO: 490:

(i) SEQUENCE CHARACTERISTICS:
 20 (A) LENGTH: 564 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 490:

ACAACAATAT AGTAAAGCGT CGTTAATTAA ACAATGGGAC CAATTTGTTC GTCTTATATA 60
 30 AATGTACTTT aCCTTTATTT TTACAAAAAT AGCATTITTC TATGTCATTT AACTAAACAT 120
 GTAAGTTCGT ATGAACGAGG TTTGTAAAT AGATGATTCT AGGAAAATGC TTTTTCITTT 180
 TGACTTAGTT TAAAATATTT TGCCACTTTG TACTGATAGT AGTTGCATTG TACTGTTGTG 240
 35 CAGATTCTAT GCTATTAAIT GAAACTGTT GCAATTTTTG AGTATTATTT AGTAATTGGT 300
 CGACCTTTTC AACCATTTGA TTGATGTCAC CTTGAGGTAC TAAATAGCCA TTAAATCCAT 360
 CTTGaATCAG TTCTGaTGGa CCATAATCTA CATCATAACT GATCACTGGT GTACCTACTG 420
 40 AAAGCGACTC TAAAATTGCT AAGCCAAAAC CTTCCATTTT ACTTGTCGAT AACATCAGTT 480
 CTGCTTTAGC AATCTCTTCA TTAATATGCG TCTTAAAACC ATGAAATTTA ACATGTTCCA 540
 GATATnATGA TAATCTTCTA CAAG 564

(2) INFORMATION FOR SEQ ID NO: 491:

(i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 1277 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 491:

	TATCCACCCC	ACGAGAGCCC	CGGAACTTA	TTGTGTTACA	AGATATATAA	GCAGAAACGA	60
5	ACAACAGTTA	ACAAAATAAA	TGAAATTAAA	CGTTTTAAAA	ATGAAACAAA	TGAAATCATC	120
	TATTAGGTTA	TGAAACTGTT	TATAGCTTGA	ATAGAAGCAT	TTATTTTTTA	GGAGGACAAT	180
	TATTATGCGT	CAAAACATTTA	TGGCAAATGA	ATCAAACATT	GAGCGCAAAT	GGTATGTTAT	240
10	CGATGCTGAA	GGCCAAACAT	TAGGTCGTTT	ATCATCAGAA	GTAGCATCTA	TCTTACGCGG	300
	TAAAAATAAA	GTAAC TTACA	CACCACACGT	TGATACTGGT	GATTATGTAA	TCGTTATTAA	360
	TGCATCAAAA	ATCGAATTTA	CTGGTAACAA	AGAACTGAC	AAAGTTTACT	ACCGTCACTC	420
15	AAATCACCCA	GGTGGTATCA	AATCAATCAC	TGCTGGTGAA	TTAAGAAGAA	CTAACCCAGA	480
	ACGTTTAATT	GAAAAC TCA	TTAAAGGTAT	GTTACCAAGC	ACTCGTTTAG	GCGAAAAACA	540
	AGGTAAAAAA	TTATTTGTAT	ATGGTGGCGC	TGAACATCCA	CACGCTGCAC	AACAACCAGA	600
20	AAACTACGAA	TTACGTGGTT	AATTAGAAGG	AGGAAATGAC	TTTGGCACAA	GTTGAATATA	660
	GAGGCACAGG	CCGTCGTAAA	AAC TCwGtAG	CACGTGTACG	TTaGTACCa	GGTGAAGGTA	720
	ACATCACAGT	TAATAACCGT	GACGTACGCG	AATACTTACC	ATTGGAATCA	TTAATTTTAG	780
25	ACTTAAACCA	ACCATTTGAT	GTAAC TGAa	CTAAAGGTAA	CTATGATGTT	TTAGTTAACG	840
	TTCATGGTGG	TGnTTCACTG	GACAAGCTCA	AGCTATCCGT	CACGGAATCG	CTCGTGCAAT	900
30	ATTAGAAGCA	GATCCTGAAT	ACAGAGGTTT	TTTAAAACGC	GCTGGATTAC	TTACTCGTGA	960
	CCCACGTATG	AAAGAACATA	AAAAACCAGG	TCTTAAAGCA	GCTCGTCGTT	CACCTCAATT	1020
	CTCAAAACGT	TAATTGTCGG	ACGATATATA	CAAAACACCT	CGATATTATG	TCGAGGTGTT	1080
35	TTTTTGCGGT	TTTTGCGGCG	AATATGGAAT	GTGTAGAATA	TAAATGAATT	TTTACCTTCC	1140
	CACCATAAAA	GATGAAGAAC	CATGAATGTG	GAGAACAATA	AATAGTTGGA	TATTCTGTTA	1200
	TTTTTTTGGA	AGTGGAAGTG	GATTTGGAAT	ACTTTACTCn	AAACGATTAA	AAGGTTTAAA	1260
40	AAAACAACAA	AAAGAAA					1277

(2) INFORMATION FOR SEQ ID NO: 492:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 673 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 492:

TTATTGACAT TGT TTTTATC CAAAATTCAT TGT TAAGACA TTTTCTTTAT GAAATAATAA 120
 ThATATTGAA GTATATTTTT ATTATTATTA AAAATAAATA AGGGGATACT TATGAGCACA 180
 5 AATCAAACAT TTTTAATATT TGT TATAGCA ATTATTCTAC TTACATCTGT AATAGGAATT 240
 GTTGGACGAT ACATGAGTCG TCAACGTCTA TTA AAATCTA TGGAAACATT ATGGCAAACG 300
 ATTTCTCCAT TAGAAGCTTT TATCAGACCG AACTCACATT TCGACTATGA GTATAAGCTC 360
 10 TACAAGGAAA AATTTGAATC ACATTCATTA GTTGATGATA AAAC TTGGTC CGACTTAAAT 420
 ATGAATGCAA TCTTTCATAA GATGAATTAT AATTTAACAG CTATTGGTGA AATGAAGCTA 480
 TATGCCTGTT TACGTGGAAT GCTTTC AATT ACGAACAAAT CATTACTTAG TTTATTTAAT 540
 15 GATAATGCTG AATTTAGAAA AAACGTAACA TATCATTTAG CTTTGATTGG TAAAAC TGTT 600
 aTCCAACATT TCCAGACCAA ATCACACCGG kAAACGTCCA AATATATTGn TCTATGCCCG 660
 20 GTTTACCACT ATC 673

(2) INFORMATION FOR SEQ ID NO: 493:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1240 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 493:

AAAAAAGTT AATGCTGACG GTGTATTAAC TTTTGATATT CTAGAAAATA AATATACTTA 60
 TGAAGTTATT AACGCTATAG GGAAAGATG GATTGTTAGT CATGTCGAAG GTGAAAACGA 120
 35 CAAGAAAGAA TATGTAATAA CTGTCATTGA TAGGAAATCA GAAGGCGACA GACAACTGGT 180
 TGAATGTACT GCTAGAGAGA TTCCCATAGA CAAGTTAATG ATTGATAGAA TTTATGTTAA 240
 TGTAACAGGA TCTTTTACAG TAGAAAGATA TTTTAACATT GTGTTTCAAG GTACTGGAAT 300
 40 GCTTTTTGAA GTCGAGGGCA AAGTTAAATC TTCAAAGTTT GAAATGGTG GTGAAGGCGA 360
 TACAAGGTTA GAAATGTTTA AAAAGGGATT AGAACATTTT GGTTTAGAAT ATAAAATAAC 420
 GTATGACAAA AAGAAAGACA GATATAAGTT TGTATTGACG CCTTTTGCAA ATCAAAAAGC 480
 45 GTCTTATTTT ATTTCTGACG AATCAACGCC AACGCTATAA AACTCGAGGA AGATGCAAGT 540
 GATTTGCGCA CCTTCATTAG AGGATATGGT AATTATTCAG GAGAAAGAAAC ATTCGAACAC 600
 GCTGGGCTCG TAATGGAAGC TAGAAGTGCA TTAGCTGAAA TATACGGCGA CATCCACGCA 660
 50 GAACCATTTA AAGATGGTAA AGTGACTGAC CAAGAACTA TGGATAAAGA ATTACAATCG 720

TATCCAGAAG CAGACCCACA ACCCGGAGAC ATAGTACAAA TAAAATCTAC CAAACTAGGT 840
 TTGAATGATT TAGTCCGTAT AGTACAAGTT AAAACGATT GGGGTATAAA CAATGTAATT 900
 5 GTTAAGCAAG ATGTAACGCT TGGTGAGTTT AATCGAGAAC AACGATATAT GAAAAAAGTT 960
 AATACTGCAG CTAACATGT TTCTGGATTA AATGATGTTA ACCTTTCTAA TCCTAGTAAA 1020
 GCGGCAGAAA ACTTGAAGTC TAAAGTAGCG TCAATAGCTA AATCAACACT CGATTTGATG 1080
 10 AGTAGAACTG ATTTGATTGA AGATAACAA CAGAAGGTAA GCTCTAAAAC TGTGACTACA 1140
 TCTGACGGCA CTATCGTTCA TGATTTTATA GATaAATCma ACATTAAaGA TGTAAAaCG, 1200
 aTTGGAACGa TTGGCGATtC TGTAGCTAGA GGATCACATG 1240
 15

(2) INFORMATION FOR SEQ ID NO: 494:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1311 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 494:

25 ACGGTGGATT TAGCCTAGAA GATTTAACGC ATCAAGGTAA ATTaTCAGCA TTTAGCTTTA 60
 ATGATCAAaC AGGTCAAGCA ACATTGATTA CTAATGAAGA TGAAAACTTC GTAAAAGATG 120
 30 AGCAACGTGC TGGCGTAGAT GCAAATTATT ACGCTAAACA AACATATGAT TATTACAAAG 180
 ACACATTTGG TCGTGAATCA TATGaCAACC AAGGTAGTCC AATTGTTTCA TTAACGCATG 240
 TTAATAACTA CGGTGGTCAA GATAACAGAA ATAATGCCGC ATGGATCGGT GACAAAATGA 300
 35 TCTATGGTGA TGGTGATGGT CGCACATTCA CAAGTTTATC GGGTGCAAAT GACGTAGTAG 360
 CACACGAATT aACACACGGT GTGACACAAG AGACAGCGAA CTTAGAATAT AAGGACCACT 420
 CAGGCGCTCT AAATGAAAGC TTTTCAGATG TTTTGGATA CTTTGTAGAT GACGAGGATT 480
 40 TCTTAATGGG TGAAGATGTC TACACACCTG GAAAAGAGGG AGACGCTTTA CGCAGCATGT 540
 CAAACCCAGA ACAATTTGGT CAACCAGCTC ATATGAAAGA CTATGTATTC ACTGAAAAAG 600
 ATAATGGTGG CGTACATACG AATTCTGGAA TTCCAATAA AGCAGCTTAT AACGTGATTC 660
 45 AAGCAATAGG GAAATCTAAA TCAGAACAAA TTTACTACCG AGCATTACG GAATACTTAA 720
 CAAGTAATTC AAACCTTCAA GATTGTAAAG ATGCATTATA CCAAGCGGCT AAAGATTTAT 780
 ATGACgAGCA AACAGCTGAA CAGGTGTATG AAGCATGGAA TGAAGTAGGC GTGGAGTAAA 840
 50 AATATATAAA CaAGAAGAAG TAATGTAAa CACTTATAAA TAATTAAATT TTAAATACAG 900

ATTAGATGAG AGGAGTGTGA GGGTTGTCTG CCGAAAGACT ACTCGGCAGT CTAAAATCAT 1020
TACAAGTAGT AGATATGTGA TAATTAAATG CTGACTTAGA ATACAAAATT CATTTTAAAA 1080
5 GTTGTACAA AAAATTTACA TGTATTTTGA TTATCTTTTG CAAAACAAAG TGTTAAATTA 1140
TAAATGAaAC ATGCATGAAT TTATTTTGA ATACAAGAAA CGTAACTACC AAAGGAGTTT 1200
ACAAATATGAA GAAAAGTAAA CGATTAGAAA TTGTTTCTAC AATAGTTAAA AAGCATAAGA 1260
10 TTTATAAAAA AGAACAAATc ATTTTCATATA TTGAAGAATA TTTTGGTGTA A 1311

(2) INFORMATION FOR SEQ ID NO: 495:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1761 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 495:

TGCACTTTCT AAAAATACTT GCTTTACTTG TTCCAATTCC TTGTCAGACA TTGATTCGGC 60
25 TAATTTATTC ATTATTGCCT CAAATACGTT AATTATGTCA TCCATTTCAG TACTATAAAA 120
ATCAAACCAT TTTGCCGTAT CTTTTTCTCT ATTAAGCTTA TGATCACTTT GAGAGCGCTT 180
AGCTAATTCT GCATAAATAT ATGGACAAGG TGCCATTGCA GCAATTGTAT AAATAGCATT 240
30 TTCACGACTA TGCGCTTGA AATACATATG TTTTATGTAA TGGTCGCCAC TTGGAGGCCA 300
AACTTTGTGTT TTAATGATTT CTTCGTATGA TTCACCAACA ATTTGCGCTA AAATATCATG 360
CGCAAGTACT TCACCTTCAA CCATAAATTC TATTTGCTCT ACTAAAAATT TTACGTCATT 420
35 CATGCTATTC ATTTTGGAA TTAACAAAGC ATATAAGTTT GTAAATCTT TTAATACGC 480
AGCATCAGCT TTTAAGTAAT GGCCTAATGC GTCAGcTCCT ATATCTCCGG ATAACATCTT 540
CTGAATAAAG TCATCCTCAT AAATATCATT AATGATTGGC TTTGCAGCTT GGTACAATTT 600
40 TTGTGAAAAT TCCATTGTAA AAAATCCTCC CTAAATAAAA AACTACTTC CAACATGAAA 660
GTAGTTTGAT GGCAATGTG CTATACTAGC CCCATCACTT CAATAACTAC TTTCCTACGT 720
TGGTACTAAC CAAATCAGGT CATAAGGGTC TGAACAATTC ATCTCAGCCA TATCATTAGG 780
45 CTCCCCTAGT AGTTCCTTAG TATTCAATTG CAAATTAATC TTAGCAAACG GTTTCAACAT 840
TTCAATTAT TGTGCTCAG TTGTATTATT ATCTTTAAAT AATAATTCTA TAATGACATA 900
TATTTGCGAA ATAAAAAAC CGGAACATAT CGAGAATTCC CCGATATATT CCAATCTAAA 960
50 AGTTACTTAT ATAACATTA ATTAGCTATG CATAAATGGC TTATGCAGTA ACCCAATGTC 1020

TTGCTGGTGA TACACCTTTA TATTTAGCAG GTGCTACTGA ATCCCAAGTT GATTGTAAGA 1140
 ATTGATACTT ACCAGCTGCA CCTGgATGTT GGrTTTACAG CATGAATATT GCCACCTGaT 1200
 5 TCACGTTGrG CAATTTGTTT TAGATGAGCa TTCACATTTA CTGATGAACC TTCTGATGAT 1260
 TTTGATyCAG TTGGTGTTC AGTAACTTGT GAATTGTTTG ATGTTGATGC TTGTGGTTGT 1320
 TGAGTTTGAG CATTTTGTGG TGCTTCAACT TCTTGTGATT GTACTTGATT AGCTTGAACA 1380
 10 GCTGATGGTG CAACATTATT AGTTGCAGGT GCTTGTGCAC TCATGTCTGC TCCATTAGTA 1440
 CCTGTTGCAT GGTAATTCCA AGCAAAGTGT GTACCATCTG ATTCAAAGTG ATAAGTAAAC 1500
 CCTTCATAGT CAAATGTATA ATTATAAGCC CCAGCTTCAA TTGGTTTTTG ATTTAATGTT 1560
 15 TGATCATTTG ATTGCGCCAT TTGCCTGAAA GATGCTTTAT TTAAGTCCGC TTCACnTGCA 1620
 TGGGCTTCGT GGACCTGCAT TTCTGGCTA CGATTCTTAA ACCTACTGGC nAAnGATGAT 1680
 GCGAGTAATG TTTTCTTCAT AATCTTAAAA TCCTCTACA AGTGAATTTG TGTCTCTAAA 1740
 20 AGTTTTACAG TGGACGACTG T 1761

(2) INFORMATION FOR SEQ ID NO: 496:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 794 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 496:

TCATTTATGA AAAATGTCCG AnAGAnCCaa GaAAmaCAAT TGAgCGTGAA GAAAAAGCAA 60
 35 GACTTAAAGA AGAACAAAAG GCACGTCAAA ATGAACAGCC ACAAATAAAA GATGTGAGTG 120
 ATTTTACGGA AGTGCCTCAA GAAAGAGATA TTCCAATTTA TGGGCATACT GAAAATGAAA 180
 GTAAAAGCCA GAGTCAACCA AGTCGAAAAA AACGAGTGTT TGATGCAGAG AATAGTTCGA 240
 40 ATAACATCGT AAATCATCAT CAAGCAGATC AGCAAGAACA ATTAACAGAA CAAACTCATA 300
 ACAGTGTGTA AAGTGAAAAC ACTATTGAAG AAGCTGGTGA AGTTACGAAT GTATCGTATG 360
 TTGTTCCACC GTTAACTTTA CTTAATCAAC CTGCAAAACA AAAAGCAACA TCTAAAGCTG 420
 45 AAGTGCAACG TAAAGGACAA GTACTAGAGA ATACATTAAA AGATTTTGGG GTAAATGCAA 480
 AAGTGACACA AATTAAATTT GGTCTGCAG TAACTCAATA TGAAATTCAA CCAGCTCAAG 540
 GGGTTAAAGT GAGTAAATTT GTAACTTGC ATAATGATAT TGCATTAGCT TTAGCAGCAA 600
 50 AAGATGTTAG AATCGAAGCG CCAATACCTG GTCGTTCTGC AGTAGGTATT GAAGTGCCAA 660

55

ATAAACTAGA AGTTGGATTA GGaAGAGATA TATCAGGTGA TCCAATTACT GTTCCACTAA 780
 ATGAAATGCC ACAC 794

5 (2) INFORMATION FOR SEQ ID NO: 497:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1161 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 497:

15 AGCCAGTTTT GcATTTCGTC AAATCGCAAT GAATATATTG ATTGCTTGTC AAAATTTGGA 60
 AGAAATTTAT TCTTCTTTTT CTTTTAGCA GTTATCATTC CTAATTGTCC TTTCTTTTIA 120
 TCTTAGTGAT AAAGAATCCA TCTGAATTAA AGTCTTGCGG CATGATTTGT AACGTTTTGA 180
 20 CCAACTCTCC AGTTATCGGA TGTGAAACG GTTCAAATTC GAAGTTTTTA TTATTTTTCa 240
 AAAACGTATA AATCACGTTT TCATTTTCTA GTTGCTCAAT TGTACATGTT GAATAGATGA 300
 TTTCTCCACC TATTTTTTACA TTGTTTTTTA CATTITCCAA TATTTCAAGC TGTAAATCAA 360
 25 CTAGTGACTC AATATGTTGT TTGCTTTGAG TATACTTAAT CTCCGGCTTA TGTCTCATTa 420
 CACCTAATCC GCTACATGGT GCATCAACAA GTATCTTATC GTATGTTTTA TCATAAGGTT 480
 TTGTCGCATC ATGTTGAAAA GCTTTAATAT TTGTTAATCG TAATTTTTTT ATATTAAAAAT 540
 TAATTAAGTC TATTTTGTGA TCATGTATAT CTGAAGCGTC AACTTGCCCT TCTGGCATTa 600
 AAACCTCAGC AATGTGACAA GCTTTACCGC CAGGTGCACT ACATGCATCT AATACGTGAT 660
 35 CATGTCGGTC TACATTCATA ATGTGTGCAa CAAACATTGA GCTTTTATCT TGAATTGAAA 720
 CGAATCCATC TTAAATGAA CGAGAATGAA TAATTGGTTG TCCTCCTATA TGGAGACAAT 780
 AAGGTAAGTC ATGATCTTTT TCAACGTCAT AACCTTCGTC TTGCAACTTT TCAATAATAT 840
 40 CATCTAATGA TGCTCGCGTC AGGTTGGCAC GCACAGTTGT TGATGTCGTT TCTAAAAATG 900
 ACTGTAAAAAT TTTTTCAGTT TCTTCGAGAC CATAATGTGT TGCCCAATGA TCTATAATCC 960
 ACTTCGGCAT ACTATACTCG ATTGCCATTC TTTTTTTAGG ATCTGCAATT TCATTAAAAAT 1020
 45 CAGGTAAGTC ACTACGCATC ATTGTACGTA AAATACCATT TACGACATTA CCATTATGAT 1080
 AGCCACCGCG TTCTTTTGCT ATTTCAACTG CTTCAATTAAT AATGGCATGA TTTGGAACCTT 1140
 TATCTAATAATA nACATATTGA T 1161

50 (2) INFORMATION FOR SEQ ID NO: 498:

55

(A) LENGTH: 1504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEO ID NO: 498:

10	AGCTCACGTC ATCTTCGGCG GCGCTAAATT AAAATAATCA ATTTCTGAGT TAAACTTTTA	60
	TTTACAACAT ACTATTACTA TACATTACAA ATTTTTAAAA TATACATTAC ACTCATTACT	120
	CAATGGmAAG CGTATGATTT CmCAGCCCCC CTAGCTTGTA GAAATCATAC TTTCTTTTTT	180
15	TCAATATATA TACAACATT AAATCCCAT AGATTGCAGA GCACATAAGT AAATTTTTTT	240
	AGAGCTTGAG GTTTGTTTTAG CTTAAGCAAC CCATGAGCTC AAACACTTCC TGTTCCACATA	300
	ACACTACAAA TCGCATTATG TTGCTTAATC TTATGTTTTAT ATAAATTACA CACAATAAAT	360
20	AGAAAGAATG TGAACATCAT GAATAAATTA TTGCTACTCG TTACATTTAT CATTTCGTGTG	420
	GGTTCAGGTA TTGTTATGTT AATGCAAGGC TACGAAAAAT TAACGGGCGG ATTTACGCTG	480
	AAAGGTTTAG TACCAGTCAT CGCTAACAACT ACTGATTCAC CAGAGTGGTA TAAGTGGTTT	540
25	TTCGCAAATA TAGTTGCACA TACGACGTCA TTATTTGATA TTGTTGTCCC ACTCGGAGAG	600
	ATTGCAATTG GATTAGGTTT AATTTTTGGA GTTTTTGCAT ATGCTGCTAG TTTCTTTGGA	660
30	GCCTTTGTTA TGATAAATTA TATCTTAGCA GATATGATAT TTACGTATCC TCTTCAATTA	720
	ACTTTCTTTA TCCTTTTACT AATGAGTCAC TCATTGTTAA AACAGATTTT ACTTAAAGAA	780
	ATCATTAAAT ACTTTAGAGG TCGTAAGAAC AGAGGTGAAA AAATAGATGA CCCACTTACT	840
35	GATCGTGGAT GATGAACAAG ACATTGTAGA CATTTGTCAA ACCTATTTTG AATATGAAGG	900
	TTACAAAGTA ACAACGACAA CTAGCGGTAA AGAAGCAATT TCTTTACTAT CAAATGATAT	960
	TGATATCATG GTACTTGATA TCATGATGCC AGAAGTTAAT GGTTACGACA TTGTCAAAGA	1020
40	AATGAAAAGG CAAAATTAG ATATCCCCTT TATCTATTTA ACTGCCAAAA CACAAGAACA	1080
	TGATACCATT TACGCCTTAA CTTTAGGTGC AGATGACTAT GTCAAAAAAC CATTTAGTCC	1140
45	AAGGGAAGTC GTTTTACGTA TTAATAATTT ACTTACAAGA ATGAAGAAAT ACCATCATCA	1200
	ACCAGTTGAA CAACTGTCGT TTGATGAATT AACACTTATT AACTTAAGTA AAGTtGTGac	1260
	tGTAAaTGGT CACGAaGTCC CTATGCGTAT TAAGGAATTT GAGTTATTGT GGTATTTAGC	1320
50	TTCTAGAGAA AATGAAGTTA TTTCTAAATC AGAATTACTT GAAAAAGTTT GGGGATATGA	1380
	CTATTACGAA GATGCTAATA CCGTGAATGT CCATATACAC CGTATTAGAG AnAAAATTAGA	1440
	AAAAGAGAGC TTTACAACAT ATACCATCAC AACTGTATGG GGATTAGGAT ATAAATnTGA	1500

(2) INFORMATION FOR SEQ ID NO: 499:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1623 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 499:

	ATTGAAAGCG ATAATTCGTA nTAATTGAGT TTGTTGAAAA ATTTAGGGTA ATGTAAAGAT	60
	ATAAAAGATA CATAGATGG AGAGATATAA AGATGTTGAA TGAGATACAA ATATTAAATA	120
15	aTGGATACCC GATGCCTTCA GTTGGGTTAG GTGTTTATAA AATCTCTGAC GAAGATATGA	180
	CTAAAGTTGT AAATGCTGCA ATTGACGCAG GCTATAGAGC GTTTGATACA GCATACTTTT	240
20	ATGATAATGA GGCTTCACTA GGACGAGCAT TAAAGGATAA TGGCGTCGAT AGAGAAGATT	300
	TGTTTATAAC AACGAAGTTA TGGAATGACT ATCAAGGTTA TGAGAAAACA TTCGAATATT	360
	TCAACAAATC GATTGAAAAT TTACAACTG ATTATCTTGA TTTATTTCTA ATACATTGGC	420
25	CTTGTGAAGC AGATGGTCTA TTTTAGAAA CATATAAAGC TATGGAAGAA CTTTACGAGC	480
	AAGGTAAGGT AAAAGCAATA GGTGTATGTA ATTTTAATGT TCATCATCTA GAAAAATTAA	540
	TGGCTCAATC AAGTATCAAA CCAATGGTGA ATCAAATTGA GGTACATCCA TATTTTAACC	600
30	AACAAGAATT ACAAGAATTT TGTGATCGTC ACGATATTAA AGTGACTGCA TGGATGCCTT	660
	TGATGAGAAA TAGAGGACTA CTAGACGACC CTGTCATTGT TAAAATTGCT GAAAAATATC	720
	ATAAAACACC AGCACAAGTT GTATTACGTT GGCATTTAGC ACACAATAGA ATTATTATTC	780
35	CAAAATCTCA GACACCTAAA CGCATTCAAG AAAATATAGA TATTTTAGAT TTTAATTTAG	840
	AATTAACAGA AGTAGCTGAA ATTGATGCTT TAAATAGAAA TGCAAGACAA GGTAAAAATC	900
40	CAGATGATGT GAAAATTGGG GATTTAAAAT AACTGGATGT TAAATTTTAC GTTTATGAAT	960
	GCCTTTTAAT GTGTACATTA AAATAAATGA GTTGGTTTTT ACTATTTGAT AAAACAATAC	1020
	TCAGGTACAT TCAAAATCTT TTAAATAAAA AGGATGGACA TAGATGAAAA TTAGAGTCGT	1080
45	CATTCCTTGT TTTAATGAAG GGGAAATCAT TACACAAACA CATCAACAAT TAACTGAAAT	1140
	ACTTTCACAA GATAGTAGTG TGAAAGGCTA TGATTATAAT ATGCTTTTCA TAGATGATGG	1200
	TAGTACGGAT ACCACTATAG ATGAAATGCA ACATCTTGCC ACAATAGATA GGCATGTCAG	1260
50	CTTTATTTCT TTTAGTAGAA ATTTTGAAAA AGAAGCAGCT ATGATTGCAG GTTACCAGCA	1320
	TAGTACTGAA TTTGATGCAG TCATCATGAT AGATTGTGAT TTGCAACATC CACCTGAATA	1380

TAGAAGTGGT GAAAATTTTA GTCGCAAAAC ATTAAGCCAT TTGTATTATA AGTTAGTTAA 1500
 TTGCTTTGTA GAAGAAGTAC AATTTGATGA TGGTGTGGT GATTTTAGAC TTTTAAGCCA 1560
 5 AAGAGCTGTT AAATCCATTG CATCACTTGA AGAATATAAT CGnTTTTCaA AAnGGnTATT 1620
 TGA 1623

(2) INFORMATION FOR SEQ ID NO: 500:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 605 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 500:

20 AAAGTnGGTG AAnCTATATA CTTAATCTAT ATTTATATAT TAACCATTAG GGTAAAAAAT 60
 TACTCTAGCA TTTATGAATA GATGGGAGTT TATTTTATTA TTATATAGGA GAGATGTTGA 120
 ATGACACATC GCGCACTATT AGTTGTTGAC TATTCATATG ACTTTATCGC AGACGACGGC 180
 25 TTACTAACAT GCGGTAAACC TGGACAAAAT ATTGAAGATT TTATTGTTTC TCGTATCAAT 240
 GACITTAATT ATTATCAAGA CCATATATTC TTTTTrTTGG ATTTACATTA TTTACATGAC 300
 ATTCATCATC CTGAAAGTAA ATTATTCCCA CCACACAATA TCGTAGATAC AAGTGGTAGA 360
 30 GAATTATACG GTAAAGTAGG TAAATTATAC GAAACAATTA AAGCGCAACC TAATGTACAT 420
 TTCATTGATA AAACGCGCTA TGATTCGTTT TTTGGTACCC CGCTTGATAG TTTATTGAGa 480
 GAAAGAAGTA TTAATCAAGT CGAAATCGTT GGTGTATGTA CCGATATTTG CGTGTACAT 540
 35 ACAGCAATTT CTGCATACAA CTTAGGtTAT AAAATTTcAG TACCTGCTGA GGGAGTGGCT 600
 CATT 605

(2) INFORMATION FOR SEQ ID NO: 501:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1739 base pairs
 (B) TYPE: nucleic acid
 45 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 501:

50 TAGGTTnAAA GCATAGnTTT nTCAAAAAGA CAAATCATTC ATATATTGGA GGATATTTTG 60
 GTGTAAGATA TAGTGCAACC ACAATTGCTA AAGACTTGAA GGAACtAAAT ATATATCGTG 120

	TGAGAGAAAA GTTTAGACAC TATTGTGAAC ATGAAGTTCT AAGTTCAATC ATCAATGGTT	240
	CATACATTAT CGTCAAAACC TCACCTGGTT TCGCCCAAGG CATAAACTAT TTTATCGATC	300
5	AGCTAAATAT AGAAGAGATA TTAGGTACGG TGAGTGGAAG TGACACTACA TTAATCTTAA	360
	CTGCCTCAAA TGATATGGCA GAATACGTAT ATGCAAAATT ATTTAAATAG ACATGTATCA	420
	AATGAATAAT AAAAATTTGT TTCGTATCAC GTGTACTCAA GTTAGTTACC AAATATTAAAC	480
10	TTGTGTACGC GTTTTTTTTAT GGAAAGAAAG AATTCATAGT CATTCAATTG ACTGTATAAA	540
	AAACTTTATA CAACATGTTT TTATGGGTAT TTTTGAATAA AAAATGTATA TTTTGACCCA	600
15	AAATACCTTT ATTTATGTAT AAAAATCCAT TATTATGTAT TGTATAACAA AAAGATATGA	660
	AATTTTCGAC TTTCTTTATG TGAATATAAT CACATGTAAG CGTTTGAAGA TTGTCTATAC	720
	TCTAAATGAA TTCAAAGATA AAAGGAGGAA ATAGACATGA CAGATGGTCC AATTAAAGTA	780
20	AATAGCGAAA TTGGAGCTTT AAAAAGTGTG TTAAGTAAAG GTCCTGGAAA AGAATTAGAA	840
	AATTTAGTAC CTGATTATTT AGATGGATTA CTATTTGATG ATATTCCATA TTTAGAAGTA	900
	GCTCAAAAAG AGCATGACCA TTTTGCGCAG GTGCTAAGAG AAGAGGGTGT TGAAGTACTT	960
25	TACCTTGAGA AGTTAGCAGC TGAAAGTATT GAAAATCCTC AAGTAAGAAG TGAATTTATT	1020
	GATGATGTAT TAGCAGAGTC TAAAAAACA ATATTAGGTC ATGAAGAAGA AATTAAGGCA	1080
	TTATTTGCGA CACTTTCTAA TCAAGAACTT GTAGATAAAA TAATGTCAGG GGTACGTAAG	1140
30	GAAGAAATTA ATCCGAAATG TACACATCTA GTAGAGTATA TGGATGATAA GTATCCATTC	1200
	TATTTAGATC CAATGCCAAA CCTTTATTTT ACTAGAGATC CACAAGCCTC AATAGGACAC	1260
35	GGTATAACAA TCAATCGGAT GTTCTGGAGA GCACGACGAC GAGAATCAAT ATTTATTCAA	1320
	TATATTGTAA AGCATCATCC TAGATTTAAA GATGCGAATA TTCCAATCTG GTTAGATCGA	1380
	GATTGCCCAT TCAATATTGA AGGCGGCGAT GAACTTGTTT TATCTAAAGA TGTCTTGGCT	1440
40	ATAGGCGTTT CAGAACGTAC ATCTGCACAA GCTATTGAAA AGTTAGCGCG ACGTATTTTT	1500
	GAAAATCCGC AGGCGACGTT TAAAAAAGTA GTAGCAATTG AAATTCCAAC TAGTCGAACT	1560
	TTTATGCACT TAGATACAGT ATTTACAATG ATAGATTATG ACAAATTTAC AATGCATTCA	1620
45	GCCATTTTAA AGGCAGAAGG CAATATGAAT ATATTTATTA TTGAATATGA TGACGTAAAT	1680
	AAAGATATTG CCATCAAACA ATCTAGTCAT TTAAGAGATA CTTTAGAAGA CGTACTAGG	1739

(2) INFORMATION FOR SEQ ID NO: 502:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1745 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 502:

5	CTGTACATAC AGCAATATCG TTAACAAATG AAAACAGTAT TTTAGGATTG TAAACCATGA	60
	TAAACCTAAA ATACTGTTAT TTTTATTACT TAAATTTCTT CTTCAATGCC TTTTCAACAT	120
	AAGGTGGAAC GAATTCAGAA ATATCTGCTC GATAAGCTGC AACTTCTTTA ACAATACTTG	180
10	AACTTATAAA TGAATAATTA GTACTAGACA TCATATATAA CGTTTCAATT TCATTGTTCA	240
	ACTTTTTATT CATTGAAGTT AAGCGTAATT CATATTCAAA ATCACTGACT GCTCTTAAAC	300
	CACGTATGAT TGTTTTAGCT CCTACTTGTT CACAATAATC GACTAGTAAA CCACTAAATT	360
15	GATGAACCTT GACATTAGGT AAATGTTTAA CAGATTGTTT AATTAAATCC ATACGCTCTT	420
	CTAAACTAAA CGTACCTTCT TTTTACTAT TTTTAAGAAC ACAGACATGA ATTTTCATCAA	480
20	ATCTATCTGT ACTTCTCTCA ATAATGTCTA AATGACCATA AGTAATGGGG TCAAAACTAC	540
	CCGGAATGAC CGCTATTGTA TGTTCATGC TATTCTCCCT TTCTAATAA CAATGTGTCT	600
	GTCAACCCAT AATGGTAACG TTTAATCATA TTAAACGGTT GATAATCtAT TTCTTCATGA	660
25	TTGCTAAATT CACAAACGAT GATACCATT TCTTTCAATA AATTAAACTC TGAAATTAGT	720
	TTTAAAGCTT TATCAATGAG ACCTTTATTA TAAGGTGGAT CTAAaGAAAT GACATCAAAT	780
	TGAATATCAC GTTTTGACAA TGCTTTTAAA GCTCTATCTG CATTATTTTT ATAAACTTCA	840
30	GATTGTGCCT CTAAATCCAA ATTGCAAGA TTGgTTTTAA TAACTTTTAC AGCTTTAAAA	900
	TTTTGaTCAA CAAAGATTAC CTTATCCATA CCTCGAGAGA GTGcTTCTAT TCCAAGCGCC	960
	CCGCTTCCTG CAAATAAATC TAAACCTATA CCTGACACAT CATATAAACT ATTAAAGATA	1020
35	CCTTCTTTAA CTTTATCCAT AGTTGGTCTC GTATTACGGC CTTCCATACT TTCTAAAGCT	1080
	TTACTTTTAT GTTTACCTGC AATGACGCGC ATGTTGTTCA CACTTCCAAT TCATTTAGTT	1140
40	ATTTAATATA ATTTATTGAG AAAAAGGAGA ATGATAAACC AATGAAACAA ACATTTATTA	1200
	CACTTGGTGA AGGTCTAACA GATTTGTTG AATTCATGAC GATGATTGAA TATAACCATC	1260
	AACGTATTGA TAAAATTATC TATTTTCATT CACCACAAGC TGAAAATAAA AAGTCATCTG	1320
45	TAGCAATCAT TATGAACCCT ACAACTGGCA ATCATTTCCA AGCATTTTAT ATCATGATAA	1380
	ACGCTATTAA ATATCCATAT CCAGATTCAA ATAAAAAGTT TCAAATGATA AATGATTGTG	1440
	CTGAAAAATT CGACATACCA ATTTTAGGTA TCGATGTACA GCCCCCTCAA GCATTTCATG	1500
50	ATTTATCGTT ATATTATAAT TATTTAATTA GTGTGTTAAG GCTCCAAAAA TGGATACCAG	1560
	aACTTCAATA ATATTAATTA TATATTCGT GTTTCTCTT TTCGTAAGTT TTCTTTAAGT	1620

55

TTTGATCAAC ATCTTGCTCA TTCACATACA TAATTACAAA TTTACGATCT CTATTTGAAT 1740
GAACG 1745

(2) INFORMATION FOR SEQ ID NO: 503:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1035 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 503:

TCGTCTTTAA TCTTGCTGAC TTTAGAAGGC TTACGAGTGC GGTTACCATT TTTTGCATCT 60
TTAACTGATT GAACTAAAGC TtGACGTGTA GATTTATCAG CTAAACTAAT TGCACCACCA 120
ATTACGGCAC CAATTAAAAT ACCAGGAACA AATTTATTTT CCATAAAAAA CTACCCCTCT 180
TTCAAATTTG CATCTTTTAC GATGTAGTCT ATTAAATTAT CACAAGATGA TAATACCATG 240
TCGTATACAC CTTCAAAATT ATTCTGTAG TATGGATCTG GTACATCACT CTCTTCCATA 300
TTACTAAATT CTAACAGTTT GAACAATTGT CCCTTAAGAT TAGGATTGAT AGATTTAATA 360
TTATCAACGT TACTTTGATC CATAGCCACA ATGTAATCAA AATCATCTGT CGCTTCGAAT 420
AATTCATAA TCATGCCATC AAATGGAATA TTGTGTTTGT TGAGAATTTT TTGTGTACCT 480
TCATGAGGTG GCTCTCCTAA ATTCCAGCTA CCACTACCTC TTGAATGTAC TTTAATATCA 540
TGAATATTTT TGTCTTTAAG TCTTTGTGCG ATGATTGCTT CTGCCATTGG AGAACGACAT 600
ATATTGCCAA GACAGACAAA TGCTACATCT ACCATTTTGA TTCCTCCAAA CTATGTAGTT 660
ATATCCCAT TTTATAGCGA CTTTAAACAA TAAGAAAGCA GATTATATAA AATTCTATTA 720
AAGTTTATTA AATTGTGATA CTTTGATAAC ATAATATTA TTAGAGGTGA ACATTGTGGC 780
TATGACAAAT GAAGAGAAAG TnTTAGCTAT TAGAGAGAAG TTAAATATtG TTAATCAAGG 840
ATTATTAGAT CCTGAAAAAT ATAAAAATGC AAATGaAGAA GAATTAACAG ATATATATGA 900
TTTTGkTCaA yCAAGAGaAA GATTGTCGCC AAGTGaAGTG mCAGCTATTG CTGaCGCTTT 960
AGGACAATTG CGACACGAAT AGGAGTGGGA ATTTTGACTA ATTACAAAGA AAAGTTACAA 1020
CAATACGCTG AACTA 1035

(2) INFORMATION FOR SEQ ID NO: 504:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1284 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 504:

5	AmCCTGACAA GATTCTTATA TTCATAGATA TGAGAGCTAA AAATGAAATC AACAAACANA	60
	CAATAAAGTA AACGATAATA GCCCATATAC CATTTTGTAA CCATATTACA AATTGTGTTG	120
	TATTATAGCC ATTTCCAGCT AATAATTGCT GGATAAATGC ATTATTGTTT AATGTATTTT	180
10	CAAGATTAGC AAtCGATGTG TTATTACTGA ATGAAACAAG TGCTATAAAC ATCGTAATGA	240
	CAGTAAGCAC TAATAACATC ACCCAACATA ACCAACCTAG AACTTTTTTCA GTTAATCTAC	300
	TTACTGGACG TTTAATTTGA GTAAATTGTT CTCCAGTCAT TCGTTACAAC TCCTTATAGT	360
15	ACTTATCCCG TTATTATAAC TAAATATACA GTAAATAACT ACTATTTATG ATTTTATTTT	420
	AATGACATTT TGAAATTCAA AAAGTTTTCA TTGTATTAC TTAAACTTC AGGTCCTAAA	480
20	TCTTTATAAA CTTCAAGGCG TTCTTGCTCT TTCTTAGTCG GATAAAAACG ATGGTCGTCT	540
	TTAATCTCTT TAGGCAACAA TTGTCGAGCA GCCTTGTTTG GCGTTGCATA GCCTACGAAT	600
	TCTGTATTTT GCTTGTATT TTTAGCATCT AATAAAAAAT TCATAAATT ATATGCACCC	660
25	TCTTTATTTT GTGCCGTTTT TGAATTACC ATATTGTCGA ACCATAAATT CGATCCTTCT	720
	TTAGGAATAA CATAATTATA TTTATCCCCT TCTTGCACTA GAGGTGCTGC AACACCACTC	780
	CAAACAACCG CTATGTTACC TTCATTTTGT TGAAGCATCA TGGTAATTTT ATCACCTACG	840
30	ACACCTCTTA CTTGTGGTGC TAGTTTGGTT AAATCTCGCT CTGCTTCTTT TAAATGGTGC	900
	GAATTACGGT CATTAGATT ATACCCAAGT TTATTCAAAC TCATGCCTAT AATCTCTCTA	960
	GCACCGTCAA CTAGTAAAT TTGGTTTTTA AATTTAGGAT TATACAATGA CTTCCAATA	1020
35	TCAAATGATT CATTTGGATA CTTTCTTTA TTATATAAAA TACCTACAGT TCCAAAGAAA	1080
	TAAGGTAAAG AATATTTATT GCCTCTATCA AATGACATAT TCATATAATC TGAATCTAAA	1140
40	TTTTTAATAT TAGGTACCTT ATTATGATCT ATTGGTAACA ATAAATGATC TCTTTTCAAT	1200
	TTTTGAACTG nATATTCAC AGGAnAAGCA ACATCATAAT GTGTACCGCC AGTGCGAATT	1260
	TTGGnGTCCA nCGCTTCATT TGAA	1284

(2) INFORMATION FOR SEQ ID NO: 505:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5763 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

	ATAATTAACA TCTTTTGCTA TATACCACCA GTTTGATACA TAAAATATCG CAGCAATAAT	60
	ATCATGTTTA ACCCTAATGA TATTATCTGA TTTTAATAAT AAGGTTGCTG TCCCTACAAC	120
5	CATTAATAAA ACTATGACTG CTGGTAATAA ACGTTTTAAA CGACGTATCC AAAAGCTTTT	180
	CAATTTAATG ATACCTGTGT CATCATACTC TTTGAGTAAT AAGCTTGTA TTAATAAACC	240
10	AGAGATCACA AAAAATGTAT CCACACCTAA AAAGCCACCT GtCAACCATT GCTTATTTAA	300
	GTGGTAAATA ATAATTCCTA GAACAGCGaT TGCCCTCAAA CCATCGAGCC CTGGAATATA	360
	TCTCATTTTC TTATACTTTG TAAAACCCCT TGTTTTGTTT ATTTTTCAT TCTTCCCTTT	420
15	TAAAACGTGTT CTCTTAGAT GCTTAATTAA ATTTAGTTAT GCTGTTTAA AGAATATTGA	480
	AATGCATATG TATATTATTG AATTACGACA TCATCAAAAT CATATTGACT AAAATACGTG	540
	TAAATTAAAA AAATTACCAA TGATGATTCT TACTTCGAAA TCCAATTTGT AATGCAACTC	600
20	GGCAACTTAA AACTATGAA GTATTATGTA TTGTAATATA ACTGTAATAT AAATTCAATT	660
	TATTATAAAA ATTTTCAAGA AAATATTCAA CTAGAAAATG AATTGTGCAC TCTTGGAAGT	720
	GCAAGTCACT GTCTTAATTC ATATTTTTTG AAACAAGTTA GATATAAATT TTCAAAATAA	780
25	AATCAGAAAC TAGAACATAA ATAAGGCTCC CTTCAAAATT TTCATTTTTC AATGTCTACT	840
	TTGAAGGGAG CTTATTCACA ATGAATTATA CTCTACAATG TTATATTGAC TGCGGGCCCA	900
30	AACACAGAGA ATTTGAAAA GAAATTCTAC AGGCAATGCA AGTTTATGTT AGCTCACACC	960
	AAGTGCAATC TTAGCGTAAC GTGACATCAT ATCTTTTGTC CAAGGTGGAC TCCATACGAT	1020
	ATTCACTTCA GTATCCTGAA TTTCAGGAAT CTCTGCTAAT ACTGTTTTAA CTTGaTCAAT	1080
35	AATTTGAGGT CCCATTGGAC ATCCCATGTA TGTTAAAGTC ATATCAACTG TACATACGCC	1140
	TTCATCATCA ACATTCACTT TGTATACTAA ACCCAAATTA ACGATATCAA TTCCTAATTC	1200
	AGGGTCAATT ACCATTTCTA ATGCACCTAA GATACTATCT TTCAATGCCT CTTCCATCCA	1260
40	TATCACCTCT TTAATGTCAT ATTATTCATA ATATATCAAA TATCCGACAA AACGCCAATA	1320
	AAATGCTATG ATGTATCTAT ATGAACTAAG CAACTTATGA GGAGAGAGAT ATGCAACCAC	1380
	ATTTAATATG TCTAGACTTA GACGGAACAT TATTAAACGA TAACAAAGAA ATTTCAATCAT	1440
45	ATACTAAACA AGTATTAAAT GAATTACAAC AACGTGGaCA CCAAATTATG ATTGCGACTG	1500
	GCAGACCTTA TCGTGCAAGT CAAATGTATT ATCATGAATT AAATTTAACG ACACCAATTG	1560
50	TTAATTTTAA TGGCGCTTAC GTACATCACC CTAAAGATAA AAACCTTCAA ACTTGCCATG	1620
	AAATTTTAGA TTTAGGCATC GCACAAAACA TTATTCAAGG ATTACAACAA TATCAAGTAT	1680
	CGAATATTAT AGCAGAAGTG AAAGATTATG TTTTCATTAA CAATCATGAT CCAAGATTAT	1740

55

	AAGAATCCCC TACCTCAATT TTAATTGAAG CCGAAGAAAG TAAAATACCT GAAATCAAAA	1860
	ATATGCTTAC TCATTTTTAT GCCGATCATA TTGAGCATCG ACGCTGGGGC GCACCATTCC	1920
5	CTGTCATTGA AATTGTAAAA CTTGGTATTA ATAAAGCAAG AGGCATTGAG CAAGTTAGAC	1980
	AATTTTTTAA TATTGACCGA AATAATATTA TTGCATTCCG TGATGAAGAT AATGATATTG	2040
10	AAATGATTGA GTACGCGCGT CACGGTGTG CTATGGAAAA TGGTTTGCAA GAACTTAAAG	2100
	ATGTAGCGAA CAATATTACA TTCAACAATA ATGAAGATGG CATTGGTCGA TATTTGAATG	2160
	ATTTCTTTAA TTTAAATATT AGATATTACT GTTAATTTAT AACTAATCAT TTTATAATAT	2220
15	TTTAAACAA TAGGAGGTAA GTTACGATGC CCAAATAGT CGTAGTCGGA GCAGTCGCTG	2280
	GCGGTGCAAC ATGTGCCAGC CAAATTCGAC GTTTAGATAA AGAAAGTGAC ATTATTATTT	2340
	TTGAAAAAGA TCGTGATATG AGCTTTGCTA ATTGTGCATT GCCTTATGTC ATTGGCGAaG	2400
20	TTGTTGAAGA TAGAAGATAT GCTTTAGCGT ATACACCTGA AAAATTTTAT GATAGAAAGC	2460
	AAATTACAGT AAAAaCTTAT CATGAAGTTA TTGCAATCAA TGATGAAAGA CAACTGTAT	2520
	CTGTATTAAA TAGAAAGACA AACGAACAAT TTGAAGAATC TTACGATAAA CTCATTTTAA	2580
25	GCCCTGGTGC AAGTGCAAAT AGCCTTGGCT TTGAAAGTGA TATTACATTT AACTTAGAA	2640
	ATTTAGAAGA CACTGATGCT ATCGATCAAT TCATCAAAGC AAATCAAGTT GATAAAGTAT	2700
30	TGGTTGTAGG TGCAGGTTAT GTTTCATTAG AAGTTCTTGA AAATCTTTAT GAACGTGGTT	2760
	TACACCCTAC TTTAATTCAT CGATCTGATA AGATAAATAA ATTAATGGAT GCCGACATGA	2820
	ATCAACCTAT ACTTGATGAA TTAGATAAGC GGGAGATTCC ATACCGTTTA AATGAGGAAA	2880
35	TTAATGCTAT CAATGGAAAT GAAATTACAT TTAAATCAGG AAAAGTTGAA CATTACGATA	2940
	TGATTATTGA AGGTGTCGGT ACTCACCCCA ATTCAAaATT TATCGAAAGT TCAAATATCA	3000
	AACTTGATCG AAAAGGTTTC ATACCGGTAA ACGATAAATT TGAAACAAAT GTTCCAAACA	3060
40	TTTATGCAAT AGGCGATATT GCAACATCAC ATTATCGACA TGTCGATCTA CCGGCTAGTG	3120
	TTCCCTTAGC TTGGGGCGCT CACCGTGCAG CAAGTATTGT TGCCGAACAA ATTGCTGGAA	3180
	ATGACACTAT TGAATTCAAA GGCTTCTTAG GCAACAATAT TGTGAAGTTC TTTGATTATA	3240
45	CATTTGCGAG TGTCGGCGTT AAACCAaACG AACTAAAGCA ATTTGACTAT AAAATGGTAG	3300
	AAGTCACTCA AGGTGCACAC GCGAATTATT ACCCAGGAAA TTCCCTTTA CACTTAAGAG	3360
50	TATATTATGA CACTTCAAAC CGTCAGATTT TAAGAGCAGC TGCACTAGGA AAAGAAGGTG	3420
	CAGATAAACG TATTGATGTA CTATCGATGG CAATGATGAA CCAGCTAACT GTAGATGAGT	3480
55	TAaCTGAGTT TGAAGTGGCT TATGCACCAC CATATAGCCA CCCTAAAGAT TTAATCAATA	3540

	GTTAGAATTA TGTTGGACTG GTACTACTAT CCAGTCCATT TTTTATGTTT AACATTTTTTA	3660
	GAATCAAAAA AGACATAAGG TCTTGGACTA ATAATTGTCC ATGCCTTATG TCATATACTA	3720
5	TATGTCTTAT CAATTAGCCA ATACCGAATA ATTTTGATAT AGGsCCTAAC GGTAGAATGA	3780
	CACCTAATAC CATTGTGATG ATAATTAATG CAATTGTTAT CCAAAACATT GTGTGACTTT	3840
10	GTTTCATGTCT CTTTCTTTTA GCAATCGACA CTTCCATCAA TCCAACTACT GCAACACCAC	3900
	ACAGCATTTT CAATGTAAGC AACATATGAT TTGCCCCGCC ATTCATAAAT GACTGAATTA	3960
	ATATCCAAAA TCCTGAAATT AACGTCAACA GCATAAATAA GCGTAAAATC ATGTGCAACG	4020
15	GTTTGAAAAA TGGTGATCTG CCTTGATTTT TTGAAATGTT TAAGTATGTA GCGATAAATA	4080
	AAATAATCGC TAATACCCAA CTTAATATAT GTAAATGTAA CATACTGATT CCCCCACTT	4140
	TAATTATTTA TATTATTAAA TTAAAGCTTC TTGGGATTAA TACCCACTTG CTGTGAATTT	4200
20	AATCATGATT TGATTATACA CGAATATATA TTCTACCACA CTTCTATATT TGAGAGGAAG	4260
	AACATGACAT TTTATTCCTT ATTAGAATAT TGTGAATCTG CTGTAAAATA ATCAACTACT	4320
	TCTTAATATC AATATTTTAC TTTTCATCTCA AAATGGTAAC ATTATAAATA ATTTATCTTT	4380
25	AACACCTTTT TAGAAAAGCA AGAAAAAAT AACAATCTA TATAAAGACT GGTAGCTTTT	4440
	TTAAATGATA ATTATTTAGC GATATAAGTT GTCAGCGTTC CAATATTATC AATAGTCACT	4500
	TTAACTTCAT CACCTGGTTG TAAAAATTTA GGTGGTTGCA TACCTGCACC AACGCCTGCT	4560
30	GGTGTACCAG TTGCAATAAT ATCTCCCGGA TGTAGTGCAA CATATTTTGA AATTTCTTCT	4620
	ATTAATTCAT CAATTTTAAG AATCATTTTCG CCAGTGTTAC CATCTTGTCT AATTTCAATG	4680
35	TTAACTTTTG TAACAATATT TACATTTTCA GGTAATGGTA GTTCGTCTTT AGTAACGATA	4740
	TAAGGACCCA TTGGGCAACC GCCAGTTAAA CTTTTTGATA AAAATGCTTG ATCTTGTTCA	4800
	CTTTGTGCTT TGCATCAGT GATATCGTTA ATAATTGTAT AGCCGTAAAC ATAATCTAAA	4860
40	GCTAATGCTT TTGGAATCTT TTCACCAGAC TTACCAATAA CAATACCTAA TTCACCTTCA	4920
	TAATCTAATT GATCAGTAAT ATCTTTATGA TTTGGAATTG TTGCATTATC TCCTGTTAAA	4980
	GATGACGCTG CTTTTGTAAA TACATATAAT TTTTCCACTT CATGATTTAA TTCGTTGCGA	5040
45	TGATCTTTGT AATTTCTACC AAAAGCAATC ACATTATTCG GAGGTGTTAC TGGTGGTAAA	5100
	AATTCAATGT CATTAAATGA AATTTTATAG TCTTCAGCTT TGCCGCTATC TTCTGCTGCT	5160
	ACAACTGCTT TACGTACTTG TTCTTGAAAA TCTAAAGTAT GATTTTGTTG TAAACCAGCT	5220
50	AACAATGTTT TAGGATGGAA ATCTCCTTCT GCAAAGTCAG CAAATACTTG TGTTAAATCC	5280
	CATACAGCAT CTTGCGGTTT TACTTTAACG CCATATGAAG TTTTGTCAAT ATACTTGAAT	5340
55		

TTCGTTATCA AATAACAAAT AAATAAGTAA GACAATTTTG AAAATGAGTT GTGTTTCATTC 5460
 TGCTACAAGG ACTTTGCACT TAATCGAAAT TATTTTTTAT TCTTTTGAAA ATCAAAATAC 5520
 5 TATAGTTGCA ATGTACCAAA TTTGAAGAAG TATAAATAAC CTTTAACTTC TTTATTAAGA 5580
 ATCGTTTGAA GCGTATTTTG ATAATATTTT ATCTGTATCT TATATTTATT TTTAATTGT 5640
 10 GTACCAATTT CTTTCATCTGT CATCCCACGG CGACGATTAA ATGCATCGGT TTTATAGTCT 5700
 ACAAAATAAT GCACACCATC TTTAACAAAG ATTAAGTCAA TCATACCTTG AATAATTGAG 5760
 ACG 5763

(2) INFORMATION FOR SEQ ID NO: 506:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 506:

25 CCATATGATT TTGTGCAATA ACTCTTTTTT TTTCTTCTTT TCGTAAAAAG TTGTACATCG 60
 CTTTGTGTT TAAGAGACTA TTGTTTTCTT TAGGTTTTTG AACTTCACTC AGTGATTTTC 120
 30 TAGTTGTAA CACTAAATT CCAACTGTTG TATCTTTGTA TCTAGCCATA ACTTTATTCA 180
 GATGTTTGTC ATTTGTAATA ACTACGACAT AATTAAACAC TTCATAATAA TCATTAATTT 240
 GATTATCTAA TCTATCCAGC TTATCTAATT CTGTTTTAAT CTCATAGACA ACGCCTTTGC 300
 35 CGTTTAACAA TATAAATCA GCaATACTTT TCCCTATGGG CATCTCAGAA AGTGCACTAG 360
 TTGTATTAAAT AGAATGTCGT CCTAGAAGGA GTkTATTAAAG TATnGTGTTT TTGTAAAAAT 420
 AT 422

(2) INFORMATION FOR SEQ ID NO: 507:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1188 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 507:

50 GCTTATGTAT TCTCAAAATA TTTATGTGAT ACGCAAAGGA GACATGGTTA TTCGACCAGC 60
 ATTTGATGAT GACGATCAAA GAAACGGTAG TGAAATAATT CGGTTTGACA AAACGCGTAT 120

TTATCTTGGC AAGAAAGCAG AGACAAACCG CATTACTGGC ATTTCTAGTA AACCACCTAT 240
 TTTACTAACA CCATTATTTT CAACTTATTT TTTCCCAACA CATTCTGACA GACAAAATGA 300
 5 AAATATTTGG TTAAATATGC ATTATATCGA AAGTATTAAA GAATTAAAAA ATCGTAAATG 360
 TAAAGTGACA TTTATTAATA ATGAATCAAT CATTCTTCAT GTTTCATACC ACAGTTTATG 420
 GCATCAATAT AACAAATCCA TTTTITACTA TTACATGGTA GATAACAAT CTCGCATGAT 480
 10 ATCAAAAAAT CCCGACCAAC CAATAGATT AATAAAGCC ACATTGAATG TGTGGAAGC 540
 ATTGACACGC TATTCTTTAT TTGAAGATAA ATAAATTGTT TATTTTAAA ATATGCGGAA 600
 TGTTTTATAA ATATAGTGTA AATGTTCTGC ATATTTTTTT AAGGTATCTA TTGCAAATTA 660
 15 ACTTAATCTT GTTATAATAA TATTGTGCT TGATATTCAA ACACATACAA ATTAATCCAC 720
 AGTAGCTCAG TGGTAGAGCT ATCGGCTGTT AACCGATCGG TCGTAGGTT CAGTCCTACC 780
 20 TGTGGAGCCA TTGGAAACGT ACTCAAGTTG GCTGAAGAGG CGCCCTGCT AAGGTGTAGG 840
 TCGCGAAAGG CGCGAGGGTT CGAATCCCTC CGTTTCCGTT ACTTGCTAAA ATGGTATATA 900
 CCATTTTAAc TTTTGTGTT ACTTATATAT AATGAATGAG AATTTCACTG TTCTTTTATA 960
 25 TCAATTTTAA AATTCTAAAA ACCTTCTCTA GATAATCTTC TCTAAGAAAG GTTTTTATAC 1020
 TTGTTGAAC TATAATTAAT TTATTACATA GCAATATTTA CCTGTTTTTA ACTATAAAAT 1080
 TATCACTACA TGAAATACGA TAATTGCGAT CTCTTAACCT CTCTGCaATT AATGtACTCA 1140
 30 TTGgTTTCAT CGTATGATTC ATGTATAATA GCATTTkTTA AATAATTC 1188

(2) INFORMATION FOR SEQ ID NO: 508:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 840 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 508:

CCCAACTTCG GTTATAAGAT CCCTCAAAGA TGATGAGGTT AATAGGTTTC AGGTGGAAGC 60
 45 ATGGTGACAT GTGGAGCTGA CGAATACTAA TCGATCGAAG ACTTAATCAA AATAAATGTT 120
 TTGCGAACAA AAtCACTTTT ACTTACTATC TAGTTTTGAA TGTATAAATT ACATTCATAT 180
 GTCTGGTGAC TATAGCAAGG AGGTCACACC TGTTCCCATG CCGAACACAG AAGTTAAGCT 240
 50 CCTTAGCGTC GATGGTAGTC GAACTTACGT TCCGCTAGAG TAGAACGTTG CCAGGCATAA 300
 TATTAATCCA CAGTAGCTCA GTGGTAGAGC TATCGGCTGT TAACCGATCG GTCGTAGGTT 360

5 GCGCGTAACA CGGGTTCGAG TCCCGTAGga GTCATACAAG CAGAAGTGAA ATATCGCTTC 480
 TGTTTTTTTA TTACATATTT ATTGTTGAGG AAGGTTGTCC GAGCTGGCCG AAGaGCACGC 540
 CTGGaAAGTG TGTAGGCGTC ACAAGCGTCT CAAGGGTTTCG AATCCCTTAC CTTCCGTAAA 600
 GGCgCTTAAA TTGGTTTTAC CCATTTTAAG CGCTATTTTT ATTTTGACT CAATCCCTTG 660
 10 ATATATCTGC ATTTGAGCTA TTATCCTCAT TTTTACACTT CTTATTTATT TATATCCATT 720
 TAAAATTTTT TAGCCACAAT GTGACTAATT TTTGaTGAAT AATCCTAATT TTAGtCACAA 780
 GATTTTGAAG TTTAGTCACA AAACAAATCA TTCAGATTTT TTTCyATAAA TTTAGTTTCA 840

15 (2) INFORMATION FOR SEQ ID NO: 509:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1373 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 509:

25 TTCTATGAAA CTATCAGCTG TTTCTTTTTT CTTGGTAAAA ATCGTTTTTAA CGCCATGATA 60
 CATGTATTCT CTTTCAaCAC CCGATTGGTC ATAAGCATCG AAAACACAAA TAATTTTCATC 120
 TGAAATAACT GCATTATAAT TTGCAATTGC ATCTATTAAT TGCATTCTAG CTTCTTCTAA 180
 30 ATTCTCTTTT GCAATGGCGC TTAGCGTTGG TGATTGTCCT ATCATATTGT ATCCATCAAT 240
 GATTAAGTAA CGTCTTTTCA TTATATTTCT CCAACATCAT GTCTTTTTTCG AAATACTTCG 300
 TACATCATTa AACTTGCTGC AACCgaAGCA TTCAAAGTGT TTACATGTCC AACCATTGGA 360
 35 ATCTTAATAT AAAAATCGCA TTTATCACTT ACTAGGCGAC TCATACCCTG TCCTTCGCTA 420
 CCAATTACAA TAGCCAATGA CATGTCCGCT TCTAGATTTT TATAATCTGT TGCATTATTA 480
 40 GCTTCAGTGC CAGCTACCCA AAAGCCATTA TCTTTTAGTT CATCGATAGT TTTAGCTAAA 540
 TTTGTCACTC GAATAACTGG TACATGTTCA ATTGCACCTG TTGAGGCTTT TGCAACTGTT 600
 TGCgTTAGTG TAACTGAACG ACGTTTAGGA ATAATAACAC CATCAACTCC CGTTGCATCG 660
 45 GCTGTTCTTA AAATTGATCC CAAGTTATGT GGGTCTTCTA AGCCGTCTAA TATAAGTACT 720
 GTCAATAAAC CTTCTTTTTT TTTTGTCTGT TTTAAAAATT GATCGAAGTC AGCATATTCA 780
 TATGGTGCAA TAAGCGCTGC AACACCCTGA TGTGGTGCAT TTGCTAAAAA ATCTAATTTA 840
 50 GATTTTGGTA CAGTTTGAAC AATGATTTTT TGATCTTTTG CATTTTTTAA AATTTCAATTA 900
 ATTTGTTGCT TTTTAATACC TTCTTGAATC AATATCTTAT TTAYCGGATG CCCAGTAATA 960

55

TTCGTTTACT ATTGTTATTA TTTTATTTAA TAATGCCTCT AATCGTTCTT CTCTTTTTTC 1080
 TAAATAaAGA AAACCTATCA CTGCTTCTAT CkCTGAACIT TTACGATATG TTTGAACATC 1140
 5 AGTGTTTTTA GCTTTAGTAT GACTTTTCGC GTTACGCCCT CGCTTCAAAA TATCCATTTC 1200
 TTCGTCTGTA AACCATTCTT GCTCCATTAA ATATTCTAAC GTTTGCGCCT GACTTTTGCG 1260
 AGATACATAT TTTTAGACA TTTGATGTAG TTTATTAGGC TTTCTTTTAA GCTTTAAAC 1320
 10 GATATAGGTA CGTACATATT GATCTAaGAC TCGCTCnCCC ATATATGCTA AGG 1373

(2) INFORMATION FOR SEQ ID NO: 510:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 717 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 510:

TCAAGTGCCA ACTACACAAC TGATTTACAT TCTTTAGGTC AATATGTACA AGAAGGCCGT 60
 25 CGCTTCTTAT TCGAAACAGT TGAAAAGTA AATCATCCTA AATATGATAT TACTATTGAA 120
 AAAGATAGTG ATGATCTAGA CGGATTAAAT TATTTAGCTG GTAAAACAAT CGACGAaGTT 180
 AACACAAAAG CATTGGAAGG TACATTATTA GCGCATACTG ATGGTGGTGT TCCTAACATG 240
 30 GTAGTGAACA TTCCACAATT AGATGAAGAA ACTTTCGGTT ACGTCGTATA CTTCTTCGAA 300
 CTTGCTTGTG CAATGAGTGG ATACCAATTA GCGGTAAATC CATTTAACCA ACCTGGTGTA 360
 GAAGCATATA AACAAAACAT GTTCGCATTA tTAGGTAAAC CTGGTTTTGA AGaCTTGAAA 420
 35 AAAGtATTAG rAGAACGTTT ATAAAATACA TTACTTCAAA GATTAGTGAA GTTTGAAAAG 480
 ATAGAACTAG ACGTTAACTA TTAAAGCAT ATTTTCGAGG TTGTCATTAC AAATGTAAAA 540
 40 ATGTAATGAC AACCTCGTTT TTATTTATAT GCAAGAACTA GGTTACTAGC TAaTGTGACA 600
 AGATGTTAAG AGAAAATTAA aGATAAAATA ATATCTGCCT TACAATAATA TTGTTATACT 660
 ACTAGnGACT GATTTATTAG CATGATTACA TGTTAATGTT TCTTTACTTA GTAATTA 717

45

(2) INFORMATION FOR SEQ ID NO: 511:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2700 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

	AATCTAATTT TTCCTCCTAA TCTCGGATTG GATAGTGGTC AATTTTTCTG GCCCCGCCATT	60
5	TTAGCGTTTG TTCTAACTGG GATTGGTTTA CCATTATTAG GTGTGATTGT AGGTGCACTT	120
	GATAAAGAAG GATATATTGG CGCATTAAAT AAAATTTTAC CTAAATTTTC AATATTGTTC	180
	TTAATCATCA TTTATTGAC TATAGGACCA CTTTTTGCAA TACCTAGAAC TGCATCTACA	240
10	TCTTTTGAAA TGACAATTAC ACCAATTATA CATAGCAATA GTAGTATCGC TTTATTTATA	300
	TTTACGATTA TCTACTTCAT AGTCGTTTTG TATATTGTGTT TAAATCCATC TAAGTTAATC	360
	GATCGTATTG GTTCATTATT AACACCATTA TTATTGATTA CTATTTTAGC GATGATTATT	420
15	AAAGGATACT TAGACTTTAG CGGTAATAGT GCTGGAAAGG GCAATGAAGC ACTATATCAT	480
	TCTAATTTTT CAAGTTTTGC TGAAGGCTTT ACACAAGGCT ATTTAACAAT GGATGCCATT	540
	GCAGCAATTG CTTTTTCAAT GATTGTTGTT AATGCAGTAA AACTAACAGG CATTACTAAA	600
20	ACAAATCAAA TATTCAAACA AACTTTGACT GCTGGTTTAA TTGCAGCCGT AGCTTTAATT	660
	TTCATATATA TTTCATTAGG TTATATTGGT AATCATATGC CAGTAAGTGA CATGACGTTA	720
	GATCAATTGA AATCCAAAGA TCGAAACATT GGGACATATT TATTAACGAC AATGGCTTCA	780
25	ACAGGATTTG GTTCATTTCG AAAATATTTA TTGGGCATCA TTGTGGCGCT GGCATGTCTA	840
	ACTACAGCAT GCGGGCTTAT TGTTGCAGTT TCTGAATATT TCCATAGAAT CGTACCTAAA	900
30	GTATCATACA AAGCATTTGT ATTAGTTTTT ATTTTAATGA GTTTTATTAT TGCTAACCAA	960
	GGTTTAAATG CTGTTATCTC AATGTCAATT CCGGTATTAA GCATTGTATA CCCAGTAGCA	1020
	ATAACTGTTG TATTATTAAT TTTAATTGCC AAATTCATAC CGACAAAACG CATTTCACAA	1080
35	CAAATCCAG TTATTATCGT ATTTATATTG TCGATTTTCA GTGTTATTAG TAAGTTAGGT	1140
	TGGCTGAAAA TTAACCTTAT AGAATCATTG CCTCTAAGAG CGTATTCTTT AGAGTGGKTC	1200
	CCAGTAGCAA TTATTGCAAC GATATTAGGC TATCTAGTCG GCATATTTGT AAAACAAGAT	1260
40	CCAATTAAAT ATCAACAGGA ATAACGAATA ATATAAAGA GGTGGGACA TAAATCCCTA	1320
	AAAAACAGC AGTAAGATAA TTTTCAATTA GAAAATATCT TACTGCTGTT CTCTATTTAT	1380
	ACAATACTTC GTATTGAATG GCTTCGCTTT CCTAGGGTGC CGTCTCAGCC TCGGTCTTCG	1440
45	ACTGGCACTG CTCCTCAGG AGTCTCGCCA TTAATACTAC GTATTAACGT GTAATTTTAC	1500
	TTTGAAATAC TTTAAAAAAA TAAGACACTT TGCCCAACTT GCACATAAAT GTAAAATTCA	1560
50	ATAAAATAAA TTTTCTGTGT TGGATCCCTT CGTATAATTT AATAAATACT ACTAAACTAA	1620
	ATTAACGAGG TGCCTTATGT ATAAAAATTA TAACATGACC CAACTTACAC TACCAATAGA	1680
	AACTTCTGTT AGAATTCCTC AAAATGATAA TACGCGATAT GTTAATGAAA TTGTTGAAAC	1740

55

AAGATATGCA TACCGTAATG ATAGATATAG TTTTAAACGT GACTTCAAGC TATATGAATG 1860
 TGATGACTGT TCATCATGTT CTTTGAGACA TCAATGCATG AAGCCAAATT CGAAATCCAA 1920
 5 TAAGAAAATT ATGAAGAATT ATAATTGGGA ATACTTTAAA GyCCAAATTA ATCAAAAGCT 1980
 TTCTGAACCA GAAACGAAAA AAATCTATAG TCAAAGAAAA ATTGATGTAG AGCCTGTTTT 2040
 TGGATTTATG AAGGCTATTT TGGGTTTCAC TCGAATGTCA GTTCGAGGAA TAAATAAAGT 2100
 10 TAAACGAGAG CTAGGTTTTG TATTAATGGC ACTTAATATA AGGAAAATAG CAGCTCAACG 2160
 AGCTGTACAT TATAAAATAC ATATCAAAAA AGCTGATTTT TATCAAATAA TTAATAGAAA 2220
 15 TCAGCTTTTT TACATTGCCT AAGAATTTAA TGTCCCAAGC CCTTTTTATC GAATAACTTA 2280
 TTGTAAACCT TGTCTTTCTT GGTATTGTT TTCGTTATTT TTTTCGTGTT TTTGTTTCCA 2340
 CTCTTTTTGA GTCATTACAT CGTCAACTTG CATGTAACT TCAACTACAT CTAAACCAGT 2400
 20 AATATATTTA ACTTGTTCTT TAACTAATTC AGTTACTTTA CGGAAGATTT TTGGTGCTGA 2460
 TTCACCATAT TCTAAGATT CTTTTAAGTC TACAGCAGCT TGTTTTTTAC CAACTTCAAC 2520
 AGATACACCT TGAGTAACAT TGTGCCACT TGAGAATGCA TTAGTGAATG TATCAGTTAA 2580
 25 GCCACCTTTC ATATCTAAGA TACCTTTAAC TTCACGTGCA GCGATACCAG CAATTTTTTTC 2640
 AACAACTTCA TCAGAGAATG TTAATTTGnt TTTGAATTGA GGCTCCTGAT TInGTTChnTG 2700

(2) INFORMATION FOR SEQ ID NO: 512:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1135 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 512:

40 ATCAAATACA ATAATTAAAA TAaTGGCTAA TACaCCTAaa ATaATAATcG AGTTaGAGTA 60
 GCCTAAGCGA yCACCcTAAC aGTArACATT CTAGGCATAT ATTTaTCTTT AGACATTGAA 120
 GCCGCTAACA TCGGAAAAGC TGTAATCCA GTGTTGGCTG CTAATACCA AATCATCAG 180
 45 GTCGTTGCCT GTACAAAATA GAAGGCAGCA TTATCACCA ATATTTGCAT AGCTAATTGT 240
 GATAAACTG TCGTTTCCGT TTGTGGcAAA ATTCCATAAA CATATGCTAA ACCAACGATA 300
 CCAACTAATA AAAACGCTAA AATTGAACCC ATAGCAATTA ACGTTTTTAC AGCATTTTTTA 360
 50 GCACTTGGCT CTCTAAATT GGTACCAGCA TTTGAAATAG CTTCAACACC CGTTAACGAT 420
 GAAGCCCCTG ATGAAAATGC TTTTAATaSc AAGAATAATG TTACTCCAGG AACCGCAGTT 480

AATATTAATA TCACTAACCC TATAATGAAA AGATATACTG GATAGGATAA TACGGTGGCA 600
 GATTCAGTTA AACCACGTAA ATTTAATATT AAAATAAAAA GTACAAGTAW ACATGCAATC 660
 5 AGT~~r~~CkTTAT GCCCATATAA ACTTGGGAAT GCAGCAACAA ATGCATCAGC ACCAGATGAT 720
 ATmCTAACAG CGACAGTCAG TATGTAATCG ACTAATAATG AGCCTCCTGC AAGCAATCCC 780
 10 CATTMTTCTC CTAAATTGGT CTTGGACACC ATATACGCGC CGCCACCTTT AGGATATGCA 840
 TAAATAATTT GCCTATAAGA CATAATTAAA GCAGCTAATA AGATCAGAAC AGCACCTGCA 900
 ATCGGTAAAG TATACCAAGT TGCAACTGCA CCCACTACTG ATAGTGTAAT CAGTATTTGT 960
 15 TCGGGACCGT AAGCCACTGA AGATAATGCA TCCGACGAAA GAATCGCTAA CCCTTTAAAC 1020
 TTCGATATCT TTTCGTCTTT TAGTTCTCTG TTTTTTTTAG GTTGCCCTAT AATAAGtCTT 1080
 TTAAATTGAT kGAcATAAtC TCCTATTCCT TTTTATAGTT TtnGATGGAA ATCAC 1135

20 (2) INFORMATION FOR SEQ ID NO: 513:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 918 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 513:

30 TTATGGTAAA ACTTTAACAC AAGCATTACA TAAATCGCAA TCATTTATGT GGCAAAAACG 60
 TTGAATTTAA GCCATTTCTA TTCATTTTAC GTAATTATTA GCCGTATATG TGTAATAATA 120
 35 CACATTTTAT TCAGATTTTT TATCGCGCTC CATTAAATCT TTTACGCATT CTTTACTGA 180
 GATATTTTCA AATAATACTC TATATAATGC ATTTGTAATT GGCATATCCA CATTMTTTTC 240
 TTTAGCTAAA TGATAAACTG ATTTAGTTGT ATAAATACCT TCAACAACCA TATTCATTTT 300
 40 AGATAATGCT TGATCCATTG ATTCACCTTG TCCAAGTTTA TATCCTAATG TGAAATTCCG 360
 AGAATGTGTT GATGTGCAAG TAACGATTAA GTCACCGATA CCACCTAAAC CTAGAAATGT 420
 CATAGGATCG GCACCTAACT TTTCACCTAA TCTACTAATT TCCGCTAAGC CACGAGTCAT 480
 45 TAATGCAGCT TTTGCATTAT CACCGTAGCC AATTCCAGCT ACGATACCAC TTGCTACTGC 540
 GATGATATTT TTCAATGCAC CACCAAGTTC AACACCAATC AAGTCATCAT TCGTGACAC 600
 ACGCAAATAA TCATTCATAA ATAAATCTTG CGTTAATTTA CTTACACTTT TATCTTTTGA 660
 50 TGAAGCAGCA ACTGTAGTTG GTTGCTTGAC TACAACCTCT TCCGCATGAC TTGGCCCTGA 720
 CAACACGCCA ATACCTGCAT TATATTCAGG TGAAATAGAA TCTTCAATCA TTTCTGACAC 780

55

CAGCTTATCA TTAATTTGAG AAGCAACTTC TCGCATTGCT TTAGTAGGTA AAGCCATTAA 900
GTAAATATCT GCAAATTG 918

5

(2) INFORMATION FOR SEQ ID NO: 514:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 587 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 514:

20

CAATACTCGG TATTATAGAT AGTCCTACTA CATTTATATG GGTTCGTATC ATTACATTAA 60
TTGCACATCA ACTTGAAGGT AATATCATTACACCAAATGT AATGGGTAAA TCTTTAAGTA 120
TCCATCCTTT AACAAATTATC GTTGTTATTT TAGCAGCAGG AGATTTAGGT GGCTTTACAT 180
TAGTTCTGAT TGCAGTGCCA TTATATGCTG TACTTAAAAC GGTGTTAGT AATATTTTCA 240
AATACCGCCA ACGCATTATT GATAAAGCAA ACAGTAATGT TAAGGACTAA TTCTGTGGAT 300
GTCTTTTAAG AATATAAGAT ACTATCGCAT CAAAAGTTGA AACTACAGCT TTTGAGGCGA 360
TTTTTTTGTG CATAAAAAAT CAGTCMAATG AAATATCAAA TAATTTTCCA TCAGTCCGAT 420
TATTATAAAA GCAAAAAAGC TTTGCTCACA TATATAATAA CGTGAACAAA GCTGTTGAAT 480
GATATTATTT AATTGCGTGG AATCCGCTAT CTACATGAAT ATTTTCACCT GTAACGCCAC 540
TTGATAAATC ACTTAATAAG TAAGCCGCAG TTTTACCTAC TTCTACT 587

35

(2) INFORMATION FOR SEQ ID NO: 515:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 812 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 515:

45

50

GGTTCATTCC AAAAAAGTAC GCGATAATTA GCGAAGAAGA ATTTGAAAAT TTAAATGTTG 60
TTAAACCAAA TAAAAATAAT GTTTTCTGGT CAGTTATAGG AAGTTCGGTT TTGTTTGGAG 120
TTACTTTAAG GAAATACATA CATGTTTTTG ATGTTCAATT AGATAAGCTA GTTGTAATGA 180
TATTGTGTGC TCTCGCTTTA ATTTGTGTGA TAGTTTTTTA TTTTAACTTA AATAGAAAGC 240
TTAAGTTAAA AGTGTGTTGAT ACAAATATTG AAAAAATAA GAGAGTTATA TAAAwACCAA 300

55

TTTCATTAAT TGCCCTTATG ACAATCGAAC CTCAAAATAT AATAATATTT ATTTATTGGA 420
 TTATGATGAC AATGCTTTTC TTTTGTGTTAA ATATGACTTC GATAGGTAAT GAAAAAGTTC 480
 5 GCGTTATAAT GAAAAATAAT TGATTACATT TAAATATTTC TAAATGTTGT CGACACAATC 540
 CTTTAAAGAC GCTAGTAGAA TTTAAATGAC TTCTAATGTA TATGAAAGTG TATCAATATA 600
 AAACCAATTG AAAAGAAGTG GAGACATTGC TTTGTGAAAC TGAAAATATT AATAAGAATC 660
 10 CCAATATAG AATTATCAAA TACAAAGATG AATATTTGAT GATTGATTTA GTAAGTACAT 720
 GGTTAGCACT CTTTTTCCCA ATGATTAAAT GGCTGATTCC AAAAAAGTAC GTCAAAATCA 780
 15 GCGAAAAAGA TTTTGAAACT TTAACATTG TG 812

(2) INFORMATION FOR SEQ ID NO: 516:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 526 base pairs
 20 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 516:

GTTCTATTAT TTCTCCAAT GGTAATGGAT TTTCGTAAAA TGATTGATAT AAGTTGATAA 60
 TCTCAATAAC TTTTAATTGA TCTGGAAACA TCGTTTTTTG GAACATTATA CTGCGATTTT 120
 30 CACTTTGTAA TAACTGTCT TTATCAAATA TCTCACCAGA ATTAGCATTACATTACCAA 180
 TTAATATATC AATTAACGTT GACTTTCCAG CACCATTTTT TCCAATTAAA GCGATACATT 240
 35 TACCTTGTTT AATATCGAAC GAAATATTTT TTAGAACACA TCCTTTATTA AATGACTTGT 300
 TGATATTAGA TATTTGAATC AATTTAATCA CCTCTATTTT TTTCTTAATT TAATATTAGT 360
 AAATTTATTA GATTTAAAAT AGAACACTTT GTCATAGATT TGAAATGACA AATGTCATTAC 420
 40 TTAGATKTAC ATAATATATT TATCGTTATT TTAATTTTGG GCAAAATAAA AAGAGCCTCT 480
 ATAATCGrGc TCCTTACAAA TAAATTATAA AAttGGCGAA cTAAAT 526

(2) INFORMATION FOR SEQ ID NO: 517:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4544 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 517:

	GGAACACCTA AAGAAAGAAA TGATGCATTA AACACAGAGG CTGATATCTA TGTAACCAAT	120
	AAAGAAAATA CTAAATGGTT ATGCGATCAA TATAAAAAAG AATGGCCATT TGATATGGTT	180
5	GTAATTGATG AACTGTCTAC ATTTAAAAGT CCTAAGAGTC AAAGGTTTAA ATCTATTAAA	240
	AAGAAATTAC CACTCATTAA TAGATTTATA GGATTAACAG GAACACCTAG CCCAAATAGT	300
10	TTACAGGATT TATGGGCTCA AGTTTATTTG ATAGACAGAG GCGAAAGACT TGAGTCTTCA	360
	TTCAGTCGTT ATCGAGAAAG GTACTTTAAA CCAACACATC AAGTTAGCGA ACATGTTTTT	420
	AACTGGGAGC TAAGAGACGG ATCTGAAGAA AAGATATATG AACGAATAGA AGATATATGT	480
15	TTAAGCATGA AAGCGAAAGA TTATCTAGAT ATGCCTGACA GAGTTGATAC TAAACAAACA	540
	GTAGTCTTAT CTGAAAAAGA AAGAAAAGTA TATGCAGAAT TAGAAAAAAA CTATATTTTA	600
	GAATCGGAAG AAGAAGGAAC AGTTGTAGCT CAGAATGGGG CATCATTAAAG TCAAAAATA	660
20	CTTCAACTAT CTAACGGTGC AGTTTATACA GATGATGAAG ATGTAAGACT TATACATGAT	720
	AAGAAGTTAG ATAAGTTAGA GGAAATTATA GAGGAGTCTC AAGGCCAACC AATATTATTG	780
	TTTTATAACT TCaAACATGA TAAAGAAAGA ATACTTCaAA GGTTTAAGGa AGCAACCACA	840
25	TTaGAGGATT CAACTATAA AGAACGTTGG AATAGTGGAG ACATTAAGCT GCTTATAGCA	900
	CATCCAGCAA GTGCAGGGCA TGGATTAAAC TTACAACAAG GTGGGCACAT TATTGTTTGG	960
30	TTTGGACTTA CATGGTCATT GGAATTATAC CAACAAGCAA ATGCAAGATT ATATAGACAA	1020
	GGACAAAATC ATACGACTAT TATTCATCAC ATTATGACCG ATAACACAAT AGATCAAAGA	1080
	GTATATAAAG CTTTACAAA TAAAGAACTA ACGCAAGAAG AATTGATGAA AGCTATTAAA	1140
35	GCAAGAATAG CTAAGCATAA GTAATGGAGG TATAAGATGG GAAAGGCATC ATACGATATT	1200
	AAGCCAGGTA CATTTAAATA TATTGAGTCA GAGATATATA ACCTACAAGA GAACAAGAAA	1260
	GAGATAAATA GATTGAGAAT GGAGATACTT AACCCAACGA AAGAGCTAGA CACTAACATT	1320
40	GTGTATGGAC CGTTGCAAAA AGGTGAACCA GTTAGAACAA CTGAACTAAT GGCAACAAGG	1380
	TTATTGACTA ATAAGATGTT ACGAAACCTA GAAGAAATGG TCGAAGCAGT TGAAAGTGAA	1440
	TACTTAAAGT TACCTGAAGA TCATAAGAAA GTAATTAGGC TAAAGTATTG GAATAGAGAT	1500
45	AAGAAGCTAA AGATAGAGCA AATAGGAGAT GCATGTCACA TGCATCGTAA TACAGTTACT	1560
	ACTATACGAA AGAACTTTGT TAAAGCGGTA cGwtATCATG CAGGTATCAA ATAACATTGT	1620
50	GCAAAGATTG TGCAAAAGGC CTACAAATCT GTAGTAATAT GATAGTATCG GAAAGATGTA	1680
	TAAAGTTATC TAAAGTTAT ACGACACAAG TACACGAGGC ACATCGCTAT GCGtGTGTCT	1740
55	TTTGTATGTC AATCAAAGAG GTGTAAGAGA TGACCAAGCA TAATAACATT TATAAGCATG	1800

	AGATAGCATT AGATAGGGAT AATCATCTTT GTCAAATGTG TTTACGTGAA GACATAGTAA	1920
	CAGATGCAAA CATAGTGCAT CATATTATTT ATGTTGATGA AGATTTTAAT AAAGCTTTAG	1980
5	ACTTAGATAA TTTGATGTCA GTTTGTTATA GCTGTCATAA CAAAATTCAT GCAAATGATA	2040
	ATGACAAAAG TAATCTTAAG AAAATTAGAG TATTAAAAAT TTAAATAAAA AAATAATTTA	2100
	TTTTTATAGC CCCCTACCCA TCGGCTTAAA ATGTTTTTTC GACGGGTACC GGCGGGGGCC	2160
10	CTTCGCTTGC AACGCGGATA AACTTTTATG AAAGGGGGTC TTTATATGAA ATTAACAAAA	2220
	AAACAGCTGA AAGAATATAT AGAGGATTAT AAAAAATCTG ATGACATATT AATTAATTTG	2280
	TATATAGAAA CGTATGAATT TTATTGTCGG TTAAGAGATG AACTTAAAAA TAGTGATTTG	2340
15	ATGATAGAGC ATACAAACAA GGCTGGTGCG AGCAATATTG TTAAGAATCC ATTAAGCATA	2400
	GAACTGACAA AAACAGTTCA AACACTAAAT AACTTACTCA AGTCTATGGG TTTAACTGCA	2460
20	GCACAAAGAA AAAGATAGT TCAAGAAGAA GGTGGATTCTG GTGACTATTA AAGTTTTAAA	2520
	TGAACCTTCA CCAAACTAT TAACAACATG GTATGCAGAG CAAGTCACTC AAGGGAAAAT	2580
	AAAAACAAGC AAATATGTTA AAAAGAATG TGAGAGACAC CTTAGATATC TAGAAAATGG	2640
25	AGGTAAATGG GTATTTGATG AAGAATTAGC GCACCGTCCT ATTCGATTCA TAGAAAAGTT	2700
	TTGTAAACCT TCCAAAGGAT CTAAACGTCA ACTTGATTAT CAACCATGGC AACATTTTAT	2760
	TATTGGCAGT TTGTTTGGTT GGGTTCATAA AGAAACAAAA CTGCGCAGGT TTAAAGAAGC	2820
30	TTTGATATTT ATGGGGCGAA AAAATGGTAA AACAACTACT ATATCTGGTG TTGCTAACTA	2880
	TGCTGTTTCT CAAGATGGAG AAAACGGCGC TGAAATCCAT CTTTGTAGCA ACGTAATGAA	2940
	ACAAGCTAGG ATTCTATTTG ATGAATCTAA GGCGATGATT AAAGCTAGCC CAAAGCTTAG	3000
35	AGAAAAATTT AGACCTTTGA GAGATGAAAT TCATTACGAT GCAACTATAT CTAAAAATTAT	3060
	GCCACAGGCT TCAGACAGTG ATAAGTTGGA TGGTTTAAAT ACACATATGG GCATTTTTGA	3120
40	TGAAATTCAT GAATTTAAAG ATTATAAATT GATTTCAGTT ATAAAAAACT CAAGAGCGGC	3180
	AAGGTTACAA CCCCTTCTTA TCTACATTAC GACAGCAGGG TACCAACTAG ATGGACCACT	3240
	TGTTAATATG GTAGAAGCGG GAAGAGACAC CTTAGATCGA ATCATCGAAG ATGAAAGAAC	3300
45	TTTTTACTAT TTAGCTTCTC TCGATGATGA CGATGATATA AATGATTCGT CGAATTGGAT	3360
	TAAAGCAAAT CCTAACCTAG GTGTTTCTAT CGATTTAGCT GAAATGAAAG AAGAGTGGGA	3420
	AAAGGCTAAG AGAACACCAG ATGAACGTGG AGATTTTATA ACCAAAAGGT TTAACATCTT	3480
50	TGCTAATAAT GATGAGATGA GTTTTATTGA TTATCCAACA CTTCAAAAAA ATAATGACAT	3540
	TATTTCTTAA GATGAGTTGG AAGGTAGACC ATGTACTATA GGTATGATT TATCAGAAAC	3600

55

AACACATTCT TGGATTCCCTA AGCATAAAGT TGAATATTCT AACGAAAAGA TACCCTATAT 3720
 AGAATGGGAA GAAGACGGAT TACTAACAAT ACAAGATAAT CCTTATATAG ACTACCAAGA 3780
 5 TGTTTTAAAT TGGATAATAA AGATGAATGA GCATTATGTT GTCGAAAAA TCACTTATGA 3840
 TAGGGCGAAT GCTTTTAAAT TAAATCAAGA GTTAAAGAAT TATGGCTTTG AAACAGAAGA 3900
 AACwAGACAA GGGGCTTTGA CCTTGAGCCC TGCaTTGAAG GATCTAAAAG AAATGTTTTT 3960
 10 AGATGGGAAA ATAATATTTA ATAATAATCC TTTAATGAAA TGGTATATCA ATAATGTTCA 4020
 GCTGAAACTA GACAGAAATG GGAAGTGGCT GCCATCTAAA CAAAGCAGAT ATCGTAAAAT 4080
 15 AGATGGTTTT GCAGCATTTT TAAACACATA TACAGATATT ATGAATAAAG TTGTTTCTGA 4140
 CAAGGGTGAA GGAAACATAG mATTTATTAG TATTATAGAT ATAATGCGTT AAGGAGGTGA 4200
 ATGTTATCGC AAAAGAGAAT ATTGTCACAC GCATAAAGAA AAAATTGATA GACAATTGGA 4260
 20 TTGaTCAGTC AGCTTCTAAG CTTTATGACT TTAGCCCATG GAAAAATAAA TCTTTTTGGG 4320
 GTGTAATCAA TAATACGCTT GAAACTAATG AAACGATATT TTCAGCTATT ACnAAGTTAT 4380
 CTAATTGAT GGCTAGTTTG CCCTTGAAAA TGTATGAAGA TTATAAAGTA GTTAATACAG 4440
 25 AAGTATCTGA TThACTTACA GTGTCACCGA ATAATTCTCT GAGCAGTTTT GATTTTATTA 4500
 ATCAAATTGA AACAATCAGA AATGAAAAAG GTAATGCATA TGTG 4544

(2) INFORMATION FOR SEQ ID NO: 518:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 881 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 518:

40 CATGGCAGAA TATTGAAGCT GTAAAAAAG GACATGTAAT TTCATATAAA GCAGAAGATT 60
 ATTGGTTCAC AGATCCTATT ACATTAGAAC ATTTGAGAAG TAAATTAAAA AAAGAAATTT 120
 TAAATAAAAA ATAATAGAAA TAAGTTGTAA AAATTTTCTT ATGCATTGGT ACTAATGTTT 180
 45 TTAAGGAGTG ATTAAATGAA GCAACTGGTT GGAATCCCG AATCAATGTT AATTCCTTTG 240
 ATAGCTCGAG CAAAAGAGTA CGAAAACGAA AAACCAATAA TAAAAGACGC ACTATCAAAA 300
 AAAATATTTG ATGGTTTAGA TGATATGTAC AAAAATGTTA CATGTGATGA CATGTCTCAA 360
 50 ATTGGAATTA GTATACGTAC TGTGATAATA GATTGTGTTA CTAAGAGGCT TATCAAGGAT 420
 AATAAAGATT TAATCGTGGT CAATATAGGT TGTGGCTTAG ATACAAGGTT TCAAAGATTT 480

ACATTTTTTA AAGAAAGTAA TAGTTATAAG ATGATATCTA AATCTATGCT AGATTACAGT 600
 TGGATTGATG ATGTCAAAAA TTATAAATTT TTTAATAGTA AGTCAGATAT ATTGTTTATT 660
 5 ATTGAAGGTG TATTGaTGTA TTTTGATGAG AGTGTAATGA CTCaATTATT GGACACTATT 720
 ATCAAAAAGA TGGGAGATCA TAATTTGACA TTTGCGATTg AATTTTGCTC aAAAAcAATT 780
 GCGAATAATA CmaAGAGACA TCAATCGGTA TCCAAGTTAT CCTCACCACC TGTTTTTAAA 840
 10 TATGGGTACC ATGATTTanA AAAATTGGAT GAnATTTACC C 881

(2) INFORMATION FOR SEQ ID NO: 519:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3122 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 519:

TGAATaAAAA TATATTAATA GATAAACACA AATGTGTCCa AATACCCCTA GAGGTATTTG 60
 25 ACnAGTTCCA TCCAACGTGT TAAATACCC CTACAGGTAT TTTTAGGGAG GTTATTATGA 120
 AACAAATACGG AGAAaGTTT ATCGATGAAT TTAGTAAAGC AGAATTGGAA AACTAGCCA 180
 AGCAAGGGCA ATTAATTGAC GTTAGAACAG AAGAGGAGTA TGCATTAGGA CATATCAATG 240
 30 GTTCCATACT TCATCTGTG GATGAGATTG AGTCATTCAA TAAAGAAAAA AATAAACCT 300
 ATTATGTAAT CTGTAGAAGT GGTAACAGAA GTGCTAATGC TAGTAAATAT TTAGCTAAAC 360
 AAGGTTATAA CGTTATAAAT CTTGATGGTG GTTATAAAGC TTATGAAGAA GAAAACGATA 420
 35 GTTATGATAC ACAAGAAGAA TATAAAAGTA TAGAAATTAA AGCAGATCGT AAACAATTTA 480
 ACTATCGTGG TCTTCAATGT CCAGGGCCAA TTGTAAAAAT TAGTCAAGAA ATGAAGAATA 540
 40 TTGAAGTAGG TGACCAAAAT GAAGTCAAAG TCACAGACCC TGGATTCCCT AGTGACATTA 600
 AAAGTTGGGT GAAACAAACA AGGCATACTT TAGTTAAGCT TGATGAAAAT AACAATGGAA 660
 TTAATGCGAT TATTCAAAAA GAAAAAGCAA AAGATTTAGA TATAAATTAT TCTGCTAAAG 720
 45 GTACTACAAT TGTATTATTT AGTGGAGAAT TAGACAAAGC TGTAGCAGCG TTGATTATTG 780
 CAAATGGTGC TAGAGCTGCT GGAAAAGATG TAACTATCTT CTTTACTTTT TGGGGGCTTA 840
 ATGCATTAAA AAAAGTGCAA ACAGTTAATG TTAAAAAGCA AGGTATTGCA AAAATGTTTG 900
 50 ATTTAATGTT GCCCAAAAAG AATATACGAA TGCCTCTTTC CAAAATGAAT ATGTTTGTTG 960
 TAGGAAATAT GATGATGCGC TACGTAATGA AAAAGAAAAA TGTTGATTCA TTACCAACAC 1020

55

	TCATGGGTAT TCAGAAAGAA GAACTTAGAG ATGAAGTTGA GTACGGTGGT GTAGGCACTT	1140
	ATATTGGTGC TACTGAAAAT GCGAATCATA ATTTATTTAT CTAATTAAAT CTATTAATAA	1200
5	AAGGAGTTGT TATCATGTTT TTTAAACAGT TTTACGATAA TCATTTATCT CAAGCATCAT	1260
	ATTTAGTGGG TTGTCAACGT ACAGGAGAGG CAATAATAAT AGACCCTGTT CGTGATTTAT	1320
10	CGAAATATAT AGAAGTTGCA GATTCTGAAG GTTTAACAAT TACACAAGCT ACAGAAACAC	1380
	ATATTCATGC TGATTTTGCT TCAGGAATTC GTGATGTGGC TAAACGCTTA AATGCAAATA	1440
	TATATGTGTC TGGCGAAGGT GAAGATGCAT TAGGGTATAA AAATATGCCA TCAAAAACAC	1500
15	AATTTGTTAA ACATGGAGAT ATCATTCAAG TAGGCAATGT TAAATTAGAA GTTCTGCATA	1560
	CTCCAGGACA CACGCCTGAA AGTATTAGCT TTTTACTCAC TGATTTAGGT GGTGGTTCAA	1620
	GTGTTCCGAT GGGATTATTT AGTGGTGA CT TATTTTTTGT TGGTGATATA GGTAGACCTG	1680
20	ATTTAaTAGA AAAATCTGTT CaaATAAAgG GTTCTACAGA AATTAGCGCG AAACAAATGT	1740
	ATGAGTCCGT TCAAAATATT AAAAATTTAC CAGACTATGT TCAAATCTGG CCGGGTCATG	1800
	GTGCTGGAAG CCCTTGTGGT AAAGCATTAG GTGCCATACC TATATCTACA ATAGGTTATG	1860
25	AGAAAATTAA TAACTGGGCA TTTAATGAAA TTGATGAGAC TAAATTTATT GAATCATTAA	1920
	CATCAAATCA ACCAGCACCA CCGCATCATT TTGCACAAAT GAAACAAGTT AATCAGTTTG	1980
30	GTATGAATTT ATATCAATCA TATGATGTTT ATCCTAGTTT AGATAATAAG AGAGTAGCAT	2040
	TTGATCTTCG TAGCAAAGAG GCCTTTCACG GTGGCCACAC AAAAGGAACA ATCAATATAC	2100
	CATACAACAA AAACTTTATT AATCAAATTG GTTGGTACTT AGATTTTGAA AAAGATATAG	2160
35	ATGTAATTGG AGATAAATCT ACTGTTGAGA AAGCGAAACA CACTTTACAA TTAATTGGGT	2220
	TTGATAAGGT AGCAGGCTAT CGTTTGCCAA AATCAGGCAT TTCAACCCAG TCCGTTTATA	2280
	GCGCTGATAT GACAGGTAAA GAAGAACATG TATTAGACGT ACGTAATGAT GAAGAGTGGG	2340
40	ATAATGGACA CTTAGATCAA GCAGTTAATA TTCCGCATGG TAAATTATTA AATGAAAATA	2400
	TTCTTTTAA TAAAGAGGAT AAAATATATG TACATTGTCA GTCAGGTGTT AGAAGTTCAA	2460
	TTGCAGTGGG TATATTGGAA AGCAAAGGTT TTGAAAATGT GGTGAATATT AGAGAAGGCT	2520
45	ATCAAGATTT TCCAGAATCA TTAAAATAAT TTAAGGATGT GGAAAAATG AATAAGCATT	2580
	ATCAAATTGT TATTATTGGT GGCGGTACAG CAGGTGTTAC CGTAGCATCA AGACTATTAA	2640
	GAAAAAATCA AAACTTAAAA GAGAAAATAG CAATTATAGA TCCAGCAGAC CATCATTACT	2700
50	ATCAACCATT ATGGACGTTG GTTGGTGCAG GGGTATCTAG TTTGAAAAGT TCTCGTAAAG	2760
	ATATGGAAAG TGTTATACCT GAAGGTGCTA ACTGGATAAA ACAGGCTGTT TCAAGTTTTT	2820
55		

5 TAGTAGCTCC AGGATTACAG ATTAATTGGT CTTCAATTAA AGGACTAAAA GAAAATATAG 2940
 GTAAAAATGG TGTGTGCTCT AACTATTAC CTGACTATGT TAACgAAACT TGGAAACCAA 3000
 10 TTTCTAATTT TAAACAAGGA AATGCCATTT TTACGCATCC AAACACTCCT ATAAAGTGTG 3060
 GAGGTGCgCC TATGAAAATT ATGTATTTAG CTGAAGATTA TTTTAGGAAA CATAAAATCC 3120
 GT 3122

(2) INFORMATION FOR SEQ ID NO: 520:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3982 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 520:

ATAGAGATAT AATTAATAT ATTGAGGTCA AACGATGATA ATTAAAATTT TAACAATTCT 60
 ATTACTACTT TGTATATTGA GCTATTTGGT TACAAATAGA AAGAAGCCTT TTCTGTTCTT 120
 25 AAAGACACTC TTTATGGGTG TGGTATTTAT CTTTATAGGA TATATTTTAC TGGCAATATC 180
 TGCCGTAATT ATTTATGGTA TTATTCAATT TATCACAATT GATTTTGGTA GTTTTTCTT 240
 AATGGGTATT ATATTGATCT TGATTTCAAG TATATTCCAA TTATTTATAG TTAGATTACT 300
 30 TTTTAGAAAA AAGAATGTCG ATTTGACAGA GGTGTGCGTT TTAsASCATT TAATTCAATG 360
 GTTCTTAGTT TACTTTGCGA TCTATCAAGC AGTAAATGAA AAAATGGACA TTAATGATAT 420
 TAATATCGAC AATTTCCAAT CTGTCTTTTT TGACGTGTCT AATTTGAATT TAGTAATTCT 480
 35 ACCAACGTTA ATCATTAGCT GGGTCACAAT ATTTAACTAT AGAATGAGAA GTTACAAATA 540
 AAATCTATGA GATTATACCT TCAGACACCA ACATTCAAAT GGTGTCTTTT TTGTTGTGTG 600
 40 GTTTTATTTT TGAAATTGCG AAAAGTAGAG GCATGAATTT TTTGACTAGT GTATAAGTGC 660
 TGATGAGTCA CAAGATAGAT AGCTATATTT TGTCTATATT ATAAAGTGTT TATAGTTAAT 720
 TAATAATTAG TTAATTTCAA AAGTTGTATA AATAGGATAA CTTAATAAAT GTAAGATAAT 780
 45 AATTTGGAGG ATAATTAACA TGAAAAATAA ATTGATAGCA AAATCTTTAT TAACATTAGC 840
 GGCAATAGGT ATTACTACAA CTACAATTGC GTCAACAGCA GATGCGAGCG AAGGATACGG 900
 TCCAAGAGAA AAGAAACCAG TGAGTATTAA TCACAATATC GTAGAGTACA ATGATGGTAC 960
 50 TTTTAAATAT CAATCTAGAC CAAAATTTAA CTCAACACCT AAATATATTA AATTCAAACA 1020
 TGACTATAAT ATTTTAGAAT TTAACGATGG TACATTCGAA TATGGTGCAC GTCCACAATT 1080

	TCAAAATCTT	GTGAGAGAAT	TTGAAAAAAC	ACATACTGTC	AGTGCACACA	GAAAAGCACA	1200
	AAAGGCAGTC	AACTTAGTTT	CGTTTGAATA	CAAAGTGAAG	AAAATGGTCT	TACAAGAGCG	1260
5	AATTGATAAT	GTATTAAAAC	AAGGATTAGT	TAAATAAAAC	TTCAATCGTT	GCTGTTATCT	1320
	GGAAATAATT	AATTAAATGT	TATGTTAATT	TTTGTTAATG	AAAAAAGTAA	TCTATTTAAT	1380
	GACAGGTTAA	TGTAATTGTC	CTGAAATTGA	CTATATACTC	AGTAAGTATC	AATTTTAAGG	1440
10	AGAGCTTATA	ATGAAATTTA	AAAAATATAT	ATTAACAGGa	ACATTAGCAT	TACTTTTATC	1500
	ATCAACTGGG	ATAGCAACTA	TAGAAGGGAA	TAAAGCAGAT	GCAAGTAGTC	TGGACAAATA	1560
	TTTAACTGAa	aGTCAGTTTC	ATGATAAACG	CATAGCAGAA	GAATTAAGAA	CTTTACTTAA	1620
15	CAAATCGAAT	GTATATGCAT	TAGCTGCAGG	AAGCTTAAAT	CCATATTATA	AACGTACGAT	1680
	TATGATGAAT	GAATATAGAG	CTAAAGCGGC	ACTTAAGAAA	AATGATTTTCG	TATCAATGGC	1740
20	TGATGCTAAA	GTTGCATTAG	AAAAATATA	CAAAGAAATT	GATGAAATTA	TAAATAGATA	1800
	ATAAATAAAA	CAGGTTGAGA	CAAAAAATGG	TCTTAACCTG	TTTTCAATTT	GCATATGTGA	1860
	TAAATTCTAT	ATCAAAATGC	TTATGTATAA	TGAATGACAT	TTAAAGTAG	GGGAGACAAA	1920
25	TATAAATACA	ATAGTTCCTA	GGATTACTCT	CAAATAACT	ATATCAATTA	TTTACTTTGC	1980
	TCTCCTATTT	TTTAAATAT	GTACATGTTT	AAACAATCAA	AAGTGTACAA	TATTAAATTA	2040
	TCATTTCCAG	TTCTAGTGCT	ATATTGGTAG	TAGTTGACTA	AATGAAAATA	AGCTTATAAC	2100
30	AAGTTTTTTC	AATACTCGTG	GGGCCACAAC	AGAGAGAAAT	AGGATCACCA	ATTCCAACAG	2160
	ACAAATGCAAG	TTGGCGGGGC	CCCAACATAG	AGAAATTGGA	TCACCAATTT	CAACAGACAA	2220
	TGCAAGTTGG	GGTGGGGCCC	AACACAGAAG	CTGGCGAAAA	GTCAGCATAC	AAAAATGTGC	2280
35	AAGTTGGCGG	GGCCACAACA	GAGAGAAATA	GGATCACCAA	TTCCAACAAA	CAATGCAAGT	2340
	TGGCGGGGGC	CCCAACATAG	AAGCTGGCGA	AAAGTCAGCT	TACAAAAATG	TGCAAGTTGG	2400
	CGGGGCCCCA	ACATAGAGAA	ATTGGAACCC	CAATTTCTAC	AGACAATGCA	AGTTGGGGTA	2460
40	GGACATCGAT	AAAGAAATAC	TTTTTCTTTA	GCAATTAGTA	TTTCTTATGC	ATGAGCTTTA	2520
	CTCATGTATT	CATTTTTTAA	GTACaCATTa	GCTACAGCTA	ATGATAAAGA	ACCACTACAT	2580
45	AATAAATCAT	TAGTGGTTCT	TTATCATTTT	TATCTCACTC	TTTACTGGA	AGAAAAAGTT	2640
	TACGTTTGTA	GAACATGCCA	CAATACCAAA	AATAATTAAG	AAAAATAAGA	CGATAAGCAT	2700
	GATGACACTT	TTCAAACAAC	CTCTATCAGT	TTCTTTTCGAT	TTTCTTTGTT	GAACCTTTTT	2760
50	ATAATCTTCA	AGTAGTTTTG	CGGCTTTTTT	ATTTATATGT	TTATTCATGA	TGTTGACTCC	2820
	TTATAATATA	TGTTTAATTC	ATTAAAATAG	TTGAAAACAT	GACTTGAAAT	AAAGATATAA	2880
55							

AATTTGATGA TGATATTTGC TTTTATTTT CCAAATGGAA TTTACTTAAA CTGATGCATT 3000
 AAAATATTAA TGAAGCACTA GAATACATAA ATGAATAGTa ATGGTGcACA GTATAGAATA 3060
 5 ATTAAGGCTA TATTAAGTAT AAATATCGTT AACTGTAAGC TATCTTTAGT TTTAATATAA 3120
 ACTATTAGGA TAATCGACGT AAGAAGAATC ATATATATTA ATGATGAAGA AGTCCATACA 3180
 AAATCCGCAT CATTGTGTGT TAATAATGGG ACTATAATTA ATCCGAAATT AATCATGCAT 3240
 10 GCTATATATA CTATAATGTT ATACACAATG TTAATTTTTG TTCACCACCT TATACTTCTA 3300
 TTTTAAAAAC TTCTTTATAA TGATATATTG TTTAATGTTG AAATAATTAG ATTATCTAAT 3360
 TTTCAATTTGC TTTACATGTA AAAGGCTATA TATAGTATGC TCTTTATGAT TCTAAATGCT 3420
 15 TTTTAATATT TAATGCTCAT CAACATTTGG ATTTTGAATA TTCAATTCAA AAACCTTTATT 3480
 AGCTACGTCA ATTGTAAAAT CAGAACCATA GTTGACATGA GCTACTTTTA ATTTTCCATC 3540
 TAAATAATAG ATTGCGATTG CAACATCGTA AAATTCGTCA ATGACAAATA AACTCTTTTC 3600
 20 GTTTGTTACA ACCTCATGCT CTCCTGAGTA TACAACGTTA ATTTCCCAAT CATTAAAAAC 3660
 CATTGTAA CCTCCTTGAA CATTTAAATT GATTCAACTT AAGTTTAACT TATTCATACA 3720
 25 ACTTCGTACA ATATCTAGAT GAACATTAAT TGTATTTCTA GAAATCTTTT TCAATTATAT 3780
 GTACTAATTA TACTTTTAAA TTTCTTATTT CAGTATAGTT TTAAAACGAT TTTAAAATAA 3840
 TTCTGCAAAT ATATTAACAC ATAATGTGTT CAAAAAGTTT TGAACAATTT CAAAACCTTT 3900
 30 ATATAAAGGG nTTGACAACA TGGATTCAAA TnTCTTATTT TAAAAATTAC CTCATATAGT 3960
 GTCATGTTAG CCAATTTTTTA AG 3982

(2) INFORMATION FOR SEQ ID NO: 521:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1353 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 521:

45 AGCTTGgATG ATTAAATaTG GtCctTTCCC AaCctTAGAT AATGAAAGAT tTGATAATCC 60
 TGAATATAAA GAAGCTATGA AAAAATATcA ACAGAGATTT ATGGCTGmAG ATGAGGCTTT 120
 GAAGAAATTT TTTAGTGAAG AGAAAAAAT aAAAAATGGA AATACTGATA ATTTAGATTA 180
 50 TCTAGGATTA TCTCATGAAA GATATGAAAG TGTATTTAAT ACTTTGAAAA AACAAAGTGA 240
 GGAGTTCTTA AAAGAAATTG AAGATATAAA AAAAGATAAC CCTGAATTGA AAGACTTTAA 300

GTTAGGTAAA ACATTTTATC AAAACTATAG AGATGATGTT GAAAGTTTAT ATAGTAAGTT 420
 AGATTTAATT ATGGGATATA AAGATGAAGA AAGAGCAAAT AAAAAAGCAG TTAACAAAAG 480
 5 GATGTTAGAA AATAAAAAAG AAGACTTAGA AACCATAATT GATGAATTTT TTAGTGATAT 540
 AGATAAAACA AGACCTAATA ATATTCCTGT TTTAGAAGAT GAAAAACAAG AAGAGAAAAA 600
 TCATAAAAAT ATGGCTCAAT TAAAATCTGA CACTGAAGCA GCAAAAAGTG ATGAATCAAA 660
 10 AAGAAGCAAG AGAAGTAAA GAAGTTTAAA TACTCAAAT CACAAACCTG CATCTCAAGA 720
 AGTTTCTGAA CAACAAAAAG CTGAATATGA TAAAAGAGCA GAAGAAAGAA AAGCGAGATT 780
 TTTGGATAAT CAAAAAATTA AGAAAACACC TGTAGTGTC A TTAGAATATG ATTTTGAGCA 840
 15 TAAACAACGT ATTGACAACG AAAACGACAA GAAACTTGTG GTTCTGTCAC CAACAAAGAA 900
 ACCAACATCA CCGACTACAT ATACTGAAAC AACGACACAG GTACCAATGC CTACAGTTGA 960
 20 GCGTCAAAC T CAGCAACAAA TTATTTATAA TGCACCAAAA CAATTGGCTG GATTAAATGG 1020
 TGAAAGTCAT GATTTACAA CAACGCATCA ATCACCACAA ACTTCAAATC ACACGCATAA 1080
 TAATGTTGTT GAATTTGAAG AAACGTCTGC TTTACCTGGT AGAAAATCAG GATCACTGGT 1140
 25 TGGTATAAGT CAAATTGATT CTTCTCATCT AACTGAACGT GAGAAGCGTG TAATTAAGCG 1200
 TGAACACGTT AGAGAAGCTC AAAAGTTAGT TGATAATTAT AAAGATACAC ATAGTTATAA 1260
 AGACCGAATA AATGCACAAC AAAAAGTAAA TACTTTAAGT GAAGGTCATC AAAACGTTT 1320
 30 TAATAAACAA ATCAATAAAG TATATAATGG GCA 1353

(2) INFORMATION FOR SEQ ID NO: 522:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1987 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 522:

GTCTGAGTCA GGTGCTGTTT GGTTAGATGC TGAAAAACA AGTCCTTATG AATTTTATCA 60
 45 ATTCTGGATT AATCAATCAG ACGAAGATGT AATTAAATTC TTAAATACT TTAATTTCTT 120
 AGGAAAAGAA GAAATTGATC GCTTAGAACA ATCTAAAAAT GAAGCACCGC ATTTACGTGA 180
 AGCTCAAAAA ACATTAGCTG AAGAAGTAAC TAAATTTATT CATGGTGAAG ATGCATTAAA 240
 50 TGATGCAATC CGTATTTTAC AAGCATTATT TAGTGGTGAT TTAAATCAT TATCAGCGAA 300
 AGAATTAAAA GATGGATTTA AAGATGTGCC TCAAGTGACA TTATCAAATG ACACAACAAA 360

TGTTAACAAT GGTGCGATTT ATATTAATGG TGAGAGACAA CAAGATGTTA ATTATGCTTT 480
 AGCACCAGAA GATAAAATTG ATGGCGAATT TACGATTATT CGTCGCGGTA AGAAAAAATA 540
 5 CTTTCATGGTT AACTATCAAT AAATATAATT GCATAGCTAA ATAAATTAGA GCCTACTCAT 600
 ATTCATTCTT AAGAATGTAA TGAGTAGGCT CTTAATGTAC TTTTCTGTCT GTAAATTATC 660
 TAAAGAAATC TCTATCGCCT TGGCCAGGTG ATTGACGTTT TGATCGGCTT TGACGTTTTG 720
 10 GTTGTTCCTT TTGTTGTTTT AATTTCACCT TAACTTCTTT TGTTTTACCA TCACGGATAA 780
 CGGTAACAGT GACTGATTCA CCAGGTTTTT TATTTTCATA TAAATAGCTT CTTAAATCAA 840
 CATCATCTTT AATTTTCTTG CCATCAATTT CTGTAATAAT ATCACCTTTT TTAAGATCAA 900
 15 TATCACTATC AGCTTTGGCG ACATAAATAC CGTCTTCTCT ATCAGTATGA AGTTGCTCGC 960
 GCTCTTCTTC AGGAATATCT TTCAAATTAA TTAAACCAAT ACCAATCGAA GGGCGGTCAA 1020
 20 TTTTACCATG TTTTACAAGT TGTTCAATTG TTACTTTAAC TTCATTACTT GGAATAGCAA 1080
 ACCCGATACC TTCAACTTGT GTCGCAGCAA TTTTCATTGA GTTAATACCA ACTAAATTAC 1140
 CATTAAATATC TACTAATGCG CCACCTGAGT TACCTGGGTT AATAGCAGCA TCTGTTTGAA 1200
 25 GAACGCTAAC TTTTGTATTG CCACCAGTTG TCTCAGCGTC AATCGTACGT TCGCTTGCTG 1260
 AAATGATACC AGATGTTACA GAGTTAGCAA ATTGTAATCC TAATGGGTTA CCCATTGCGA 1320
 ATACGCTATC GCCAGTTTGT ACTTTTGAAG AGTTGGCAAA TTGAATCGCT TTAATACCTT 1380
 30 TTGTATTTTC AATTTTAAGT ACAGCAATAT CAGTTACTGC ATCTTTACCA ACTAATTTCTG 1440
 CTTTAACTTG TTTTTATTA TGTAATTGGA CTCTAATTTT ATTTGCGCCA TCAATAACAT 1500
 GATTGTTTGT AACGATATAA GCTGAATTGT TGTTTATTTG ATAGATAACA CCTGAACCTA 1560
 35 CTCCAGCTTC AGATGGTTTA GATGATTTGC CyTTTAAATAA GTCGTCTACA CTTGATGCTT 1620
 TTtGCATGk AATAACTCCA ACAATTGTAG GGGAGACAGA TTTTATCATT TCATGAACGG 1680
 TACCGAATTT CTGCTTTGA CCGTCTAATT GATTGCCACC TTTATTATTT GTTGTCTGAA 1740
 40 CAGTTGAACC ATCTTTATTT AAAATTGTAC TATTTAATAC TTTGCCTATA CCAAGTACTA 1800
 GAAGTGCACC AATAATTCCA GCAATCAATG CAACGATGAC TGtTTTAAAC CATGGAAT 1860
 TAGGTCTTCT GTATCTAGGT GTTTGGCTAT GGTtTGTTGT AGAATGATCT GTATGATTAA 1920
 45 AATCTGACAT ACTTAACCTC CATTATATGA TTTATATATG CTTTAAATTAT GTCTTTTtTT 1980
 TATAATT 1987

50 (2) INFORMATION FOR SEQ ID NO: 523:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1351 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 523:

	TTAGAAGTCA AATCATTANT GCGGTCNTAT CGAGTATACT ATTAACCTCA ACTATTTTAG	60
10	CAATTGCATA TATTTTAATG TGGTTTAACG GCCATATGAC aCTAACTTTG ACCTTAACGA	120
	CAATAATTAC AAGCTGTTTA ACCTTATTAA TATGTAGTAT TTTTATTAAT CCACCTTATAC	180
	AAAAAATTAA GCAGTTTAAT ATAAAACTA AGCAATTTGC TAACGGAAAT TACGCAAGCA	240
15	ATGATAAAAC GTTTAATTCA CAAAAGAAA TTTATGAATT AAATCAATCT TTTAATAAAA	300
	TGGCTTCTGA AATTACGCAA CAAATGAATC AAATTAAATC CGAACAACAA GAAAAACAG	360
	AACTGATTCA AACTTAGCC CATGATTTAA AAACACCTTT AGCAAGCATT ATTTCATATT	420
20	CTGAAGGACT ACGTGATGGT ATAATCACTA AGGATCATGA GATTAAAGAG TCATACGACA	480
	TATTAATTAA ACAAGCAAAC AGATTATCAA CATTATTGTA TGATATGACT CATATTATCA	540
	CTTTAAATAC AGGTAAAACA TATCCCCCAG AATTAATACA ACTAGACCAA TTACTTGTAT	600
25	CAATATTGCA ACCATATGAG CAACGTATCA AACATGAAAA CCGCACATTA GAAGTGAATT	660
	TCTGTAACGA AATTGATGCA TTTTATCAAT ATCGAACGCC ACTTGAGCGT ATTTTAACAA	720
	ACTTACTTGA TAATGCGCTA AAATTTTCAA ATGTTGGTAG TCGCATTGAT ATTAATATTA	780
30	GTGAAAACGA AGATCAAGAT ACTATCGACA TTGCTATTAG CGATGAAGGT ATTGGCATT	840
	TACCAGAACT ACAAGAACGT ATATTGGAAC GTACATTGAG AGTAGAAAAC TCTCGTAATA	900
35	CAAAAACGGG TGGTCTGGA TTAGGCTTAT ATATAGCTAA TGAAGTCGCG CAACAAAATA	960
	ACGCAAAAAT CAGTGTAGC AGTGATATAG ATGTAGGAAC TACGATGACT GTAACATTAC	1020
	ACAAATTAGA CATTACGTCA TAATCCGATT TATTTATAAA ATAAAATGCA AAGACTAAAA	1080
40	AGAAGCTCCC ATTAATGAGG GCTTCTTTTT TTGTTTATTT AGAATAAACT TTATGGGTAT	1140
	CCTTCTCATC ATTTTCAAGA CTTGAAAGAT TTGTAGCTTG AATAATATAT TTAGGACGTG	1200
	CCTTAACCTC ATAATATATC CTGCCAATAT ATTCACCTAC AACACCAATT GAAATTAACT	1260
45	GTATGCCGCC TAATAATAAA ATAGCTGCAA TCGTTGAAAA ATATCCCGGA ATATTAACAC	1320
	CAGATATCAT AATATTGATG GAATAAATAG A	1351

(2) INFORMATION FOR SEQ ID NO: 524:

50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 524:

5 TTGTTGTCAG TTTAACGCAA CGTGTACCG ACTTTTCAAG TACACAACAA TATATATGTC 60
 TATTCCATTT TTAGCCCcTG CCATTTTCAT CATTGGTGGT ATTATGTTGT TTATTTCAAC 120
 ATTTAATAGC TTAGATGAAA CTGCTGAAAA TAATAACAAA ATAAAGAAAC TAATGATTAA 180
 10 AGGACTTATC ATTATTAACA TTTCATTTAT CGTTATGATG GTTTTAACAC CATATTGGTA 240
 CTTGTATTTA ATCGTCTATC TTATTTTCTT GTTGTTTTTA TTGTGGCAAA AGGTTTATAA 300
 ATTTTAATAC CAAACTATT AAACACTTCT GATATTCTTA GTTCAAAATA TCAGAAGTGT 360
 15 TTTATAGTGT TATCTAGTTC AGATAAATAT TTCCTTACTT AAAAAACGC CCTCCTCTTA 420
 TTTTGACCCC nAT 433

(2) INFORMATION FOR SEQ ID NO: 525:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1845 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 525:

30 CCCCCACTAT GATATGCTTG GCCTATTGCG TCAATCCCCT TATATTTTCGG CAATAGAGAT 60
 GGATGTATAT TCAATATTTT ACCTTCAAAT GAAGCTAATA AGTCTGGACC AATTAGACGC 120
 ATGTAGCCAG CTAGAATAAT CCACTCTACC TTATCTTCAT TTAATAATGT TACTAAATGT 180
 35 TGTTCATACG CTGCTTTTGA ATCAAATTGT TTTGGTTCAT TAATATAAAC AGGAATATCG 240
 TGCTTTTTTG CTCTATCTAT ACAAACGCA TTTTGATGAT CCGTATATAG CGCCGTAAC 300
 TCAATATTTT CAAGTTTTCC TGATTCAACA TGCTCAACTA TATTTKCAA GTTACTTCCT 360
 40 GAACCTGATG CAAAATCGC AATTTTAACC ATTGTTATAC CCCCACAAT TCAATTGCAG 420
 TTGACTCATT TTTCACAATA TGACCAATTG GATAGGCTTC CACATTTTGT TCTGCTAAAA 480
 TCTTCAAAGC GCGTGATACA TCTTTTTTCAT CAACGATAAC CGTATAGCCA ATACCCATAT 540
 45 TAAAAATGTT ATACATTTCA TTTGTGTCTA TATTGCCTTG TTGTTGTAAC CAATCAAATA 600
 TTTTGGTGT TGGAAATGAT GTAGTATCAA TTCTAGCAGC ATATCCGGCT GGCAATGCAC 660
 GTGGAATATT TTCATAAAAA CCTCCACCAG TAATATGATT CATTGCCTTA ATAGAAACTT 720
 50 CTTTTTCAA AGCAAGTACA GGTGACAT ATAATTTAGT TGGTTCTAAA AAGACATCGA 780

55

5 GCACTAAACT GSTATCCATT GAATGAATGC CACTTGACGC AAGCCCTATA ACAACTTGTC 900
 CCTCTTTCAC TTCTGAACCA TCTACATAGT CATCCTTTTC AACTGCTCCA ACAGCAAATC 960
 CAGCTACATC ATATTGCGCT TCGTGATACA TTTCACCCAT TTCAGCAGTC TCTCCACCGA 1020
 TAAGTGCAGT ATTGCTTTCA ACACATGCAT CACTAATACC TTAAACAATT TGTTCAATAA 1080
 CTTCAGGAAC AACTTTGTTT GTAGCAATGT AATCTAAAAA ATATAATGGT TCTGCACCTG 1140
 10 TCGTTAAAAAT ATCATTAAACA CACATTGCGA CTGCATCGAT ACCTATCGAA TCATGTTTAC 1200
 CATAGTCGAT AGCTAGTTTT AATTTGCTAC CTACTCCGTC TGTTCAGAA ACTAAAACTG 1260
 GCGCTGtCAT ATTTAATTGT GATAAATCAA ATGTAGCACC GAATCCACCT AAACCACCGA 1320
 15 TAACTTCTTT ACGCATCGTA CGtTTAAcaw GaCtAGaCaT TctTTcACA GCTTCATAAC 1380
 CAGCATGAAT ATTTACACCA GATTGTTTAT ATGCTTTAGA CATTTAAATT TCCCTCGCTA 1440
 TCAAAAAAGT GTTTGTTTTT AGAAATATAT TGCTTTTGTC GATGACTTAA ATGCGCTTTA 1500
 20 TAATTTGCTT CATAATCATA TAACCCTGCA GGATAATCTC CAGTGAAACT TTCTACACAT 1560
 AAGCCACTAT ATGGCGCGTC ATAATCTAAA CCAATTGATT CAATTAACCC ATCTACAGAT 1620
 AGATATGCTA ATGAATCAGC GCCAATATAA TCTTTAATTT CTTCAGGTGA TTTGCTTGCA 1680
 GAAATTAATT CTGCCGTAGT TGAAACATCG ATTCCGTAAA AACTTGGAAA CATAAATTC 1740
 GGTGATGCTA TACGCACATG CACTTtATTT GCACCAGAAT CkTTTAACAT TTTCACAATG 1800
 30 CGTCGAATTG TCGTACCGCG nACAATGGAA TCATCAACAA GACTG 1845

(2) INFORMATION FOR SEQ ID NO: 526:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1141 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 526:

TCATAAATnG TAAATGTTAA TCGTCATATA ATATTAATTT ACAACACCAT TTTGTTTATT 60
 TGAAGCTTGT GCGCCTTGTT GTGTGCCACC TTGATTTTGA TTTGAGTTTT GATCTGTAGC 120
 45 AGGTTGTTGT TGATTGCTGG AmtCACTGTT ATTAGTTGAm TCACTGTTCT CGTTAGATGT 180
 CTTATCTTTA TCTGTCGTAT CATTTTGTTT TTTTCTCAAT AACTATTAT CTAAAGGCGT 240
 TAATGGTATT AATGAACCAT AATGATTAAT GACACGTTGA TCTAAGAAAT CATTTTATC 300
 50 ATTAATAGGT GATAATTCTA AGTCTTTACG AAGTAAGTTT GCATATTTT GAATGCTTTC 360

AGTTTTAATT TCAACATCAT TTGTTaAGTA CATTTTTGCT AAAGCTTTAA TTTcAGAGTT 480
 AGTTAAATTA TGCTTTGCAT TTTTACCTAC AATTTGAATC ACGTTATCAA GTTTATCAAT 540
 5 AGAATCAACT TCCTGTGCTT TTTGGAATAA AATCTTAATT AATTCATT T GACGTTGTCC 600
 ACGTTTTAAG TCTGAATCAT GATGTCTAGT TCTAGCAACT GCTAAAGCCT CATCACCATT 660
 10 TAATTTTGG TACCCTTTTT TAATTTTAAT CTTACCAGTA TCATCTGTGT TAGGTTcATT 720
 TAAGTCGTAT GGCACATCAT AGTATATGCC ACCAAGCTCA TTTACAGCCT CGACAAATGC 780
 TTTcATATTG ACTCTCACAT AATAATCAAC AGGTACATTc ATGGTAGCTT CTACCGAATC 840
 15 CATTGCGGCA ATTGGACCAC CATATGCATG TGCATGGGTA ATCTTATCGT AATAGCCAAC 900
 TTTAGGAATG TAGCTGATAG TATCACGTGG AATACTAAGC ATTCTAATTT GATGTTTTGA 960
 TTGATTAAAA GTAGTTAAAA TCATAGCGTC TGATCTAGAG TGTTCAGCAT CCTGTCCTTT 1020
 20 TTTTCTTCTT CCATCGTTAT CATCGATACC TAAGAAAAGA ATAGAGATAG GTTGTTCCTTc 1080
 GGGATGACTT TATTATCTCT TAAGTTGGAT TGACGTTAGC ATTTTGTCTG TCTTGAGAAG 1140
 A 1141

(2) INFORMATION FOR SEQ ID NO: 527:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1565 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 527:

35 GCACCATTAG CATTGGGcnc ACTGATTGGA GTAnCAGTTG TTGAAAATTC GGCGCCAACA 60
 AGTAAACmGG CmCaGGCaGC mATamCCCCa TATTATAcTT ATAATGGTTA TATTGGTAAT 120
 40 AATGCTAATT TTATTCTGGA TAAGAATTTT ATTAACGCGA TTAAGTATGA TAATGTGAAA 180
 TTCAATGGTA TTAAATTAGC TAAAACGAAT ACGATAAAAA AAGTAGAGAA ATATGATCAA 240
 ACTTTTAAAG GTGTTTCTGC AAAAGGAAAC GAAGCAAGTC AATTGCAATT TGTAGTTAAA 300
 45 AATAATATTT CATTAAGA TATCCAAAAA GCTTATGGCA AGGACTTGAA AAAAGAAAAT 360
 GGTAAACAA AGGAAGCTGA TAGCGGTATT TTTACTATC AAAATGCTAA AAAGACATTA 420
 50 GGCATCTGGT TTGTCGTTGA TCATAATAGA GTTGTGGAAG TAACAGTTGG ACATACACCA 480
 TACAAAACAA GTAAATAAAA TAATGGCATA TTAAGGCTAG AGTGTGAGGA GTGATACCGC 540
 ATTCTAGTCT TTTTATTAA ATAATAACGA TTATTGCGTC TTACATAGTT GTTTGAAATT 600

5 GATTAAGTAT ATAGAGCACT ATTTTGTATT TGTTAATATT TTCACAAAAA TAAAGCCTTG 720
 ATAAATTTTA AATATAATT AAGCTCAATT TTTAAATTTT TATTTAGCTA CAGATAACAT 780
 TTTTAAAAAA GAAAAGAATC AATAAATAAA ATCAACGAAC AAAAAGTATA GAAATAAATA 840
 GAAATAATCA TTTACTTTTC TGAAAAATTA AATTAATATT TTATTTATAA GCTGTTTTTA 900
 AGATTTTCAGG AGGAATGAAA TGTGaGGAAA TTTTCAAGAT ATGCATTTAC AAGTATGGCA 960
 10 GCATTAACCT TGTGAGCAC TTTATCACCA GCAGCATTAG CGATTGATTC AAAAAATAAA 1020
 CCAGCTAATT CTGATATTAA ATTTGAGGTG ACTCAAAGA GTGATGCGGT CAAAGCATT 1080
 AAAGAATTGC CTAAATCCGA AAATGTAAAA AATATTTATC AAGATTACGC TGTTACTGAT 1140
 GTAAAAACTG ATAAAAAAGG ATTTACGCAT TATACATTGC AACCGAGTGT TGATGGTGTT 1200
 CATGCACCTG aCAAAGAAGT GAAAGTACAC GCAGACAAAT CAGGAAAAGT CGTTTAAATC 1260
 20 AATGGGGATA CTGATGCGAA GAAAGTAAAG CCAACGAATA AAGTGACATT AAGTAAAGAT 1320
 GACCGAGCCG ACAAAGCATT TAAAGCAGTT AAGATTGATA AGAATAAAGC GAAAAATCTT 1380
 aAAGATAAAG TCATTAAAGA rAACAAAGTT GAAATCGATG GTGACAGTAA TAAATACGTT 1440
 25 TATAATGTTG AGTTAATTAC AGTGACACCA GAAATTTTAC ATTGGAAAGT TAAAATTGAT 1500
 GCTCAAACCTG GCGAAATTTT AGAAAAAATG AACTTAGTTA AAGAAGCTGC AGAAACTGGT 1560
 AAAGG 1565

(2) INFORMATION FOR SEQ ID NO: 528:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2870 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 528:

40 GGATAGGTAA ATCCATTCAA AATTGGTTGT ACTCATTTTA AATCATATTT TTAAATATAG 60
 AACACATAA TTAAAGGAGT GATGATATTA ATGAATAAAC TTCGAGACAC TACTTTTCTA 120
 45 TCATATTTAT TTAATATCAT ATTGTGGGGA TCTGCATTTT CAATGATAAA GATTGCGTTA 180
 AATGATTTTA GTGCAGAATC ATTGTGGGCA TTTGTTTAA TTTTAGCAAC AATAATTCTC 240
 TTGCCGTTTC TAATTATAAA GAAATTGCCT ACCCCTGAAC TAAGAGATAT CCCTGTTATT 300
 50 TTTATTTTAG GATTTTGTGG ATTTGTGATA TATCACACAG CTTTAAATTT TGGTGAAaCT 360
 TTGATTAGTG CAGGTATatC TGGrATTcyA GTCTCTACAA CGCCTATTTT TTCTAGTGCT 420

	GCATTTATAG GAATATCCAT TATTTCAATA AGTAAAGATG ATTACACAAC TATTAATGTA	540
	TTAGGTGTTT TTATTATTTT ACTTGCATCT TTTAGTGAAA GTTTGTATTT TACTTTCCAG	600
5	AAAAAATACA TAGAAAAATA CGGCTTCATC GCTTTACAC TATATACAAT AATGGCAAGC	660
	TCACCATTTA TGCTTATTTT TATTCCTGAA ATCATCAACG ATATACACGG CGCCACTTTT	720
10	ACATCAATAG TATCGGTACT TTATTTAGCT ATATTCCCTA CTATAATTCC ATACGTTTTG	780
	CTTGCTTATA TTGTGAAGTC AGTTGGTGTC TCTGATGCAA CAATGTCTCT TTATTTAACA	840
	CCTATCGTTT CTTTATTATT ATCTTATCTG TTATTAGACG AGCTACCAAC AACCCTTGCT	900
15	ATTATAGGCG GAATTATCAC TCTACTAGGC GTTAGTTTAA GTAACCTCTT TCAAAATACA	960
	TAATTATTCC AAGTCCCGCA CCTCAGAATC CAAAAACATT CGAGTGATAA AATTTTAAAA	1020
	ATCAAAAATA TAAAAATGAT CTAATTTCTG CAAATTTACC AATATAAATA CTAATATTTG	1080
20	CAATTCACAA AGGGGTATAG TCTGAGTGTA TTCTAATACG AAAGGACTTG GTGGATATGT	1140
	ATTACAGTTA TGGAAATTAT GAAGCATTTG CGCGCCCTAA AAAACCTGAA AATGTAGAAA	1200
	ACAAATCCGC TTACTTAATC GGATCTGGTC TAGCTTCACT TGCTGCAGCT TGTTTTTTAA	1260
25	TAAGGGATGG TCAAATGGAA GGTTCAAGA TTCATATTTT AGAAGAGTTA CCTAAAGCAG	1320
	GTGGTAGTCT TGATGGTGAA AATATGCCTT TAAAAGGCTA TGTGTCCGC GGTGGTCGTG	1380
30	AAATGGAGAA CCACTTTGAA TGTGTGTTGG ACTTATTCAG ATCTATCCCT TCATTAGAAA	1440
	TCGATAACGC GTCTGTATTA GATGAGTTCT ATTGGCTAAA CAAAGAAGAC CCTAACTATT	1500
	CTCGTGTCTG TGTTATTGAG AAACAGGGTC AACGTTTAGT CACAGACGGA GACTTCACTT	1560
35	TGACTAAAAC GGCGATTAAA GAAATTTTAG ATTTATGCTT AACGAATGAA GAAGATTTAG	1620
	ATGATGTCAA AATAACAGAT GTATTTTCCG ATGACTTCTT TAATTCAAAC TTTTGGATT	1680
	ACTGGAAAAC GATGTTTGCA TTTGAACCGT GGCAATCTGC AATGGAAATG CGTCGCTATC	1740
40	TAATGCGATT CGTTCATCAT ATTAGTGGTC TCGCAGACTT TTCAGCTTTA AAATTCATA	1800
	AATATAATCA ATATGAATCT TTAGTATTAC CTATGGTTGA ATATTTAAAA TCGCATGGGG	1860
	TTCAATTTGA ATACGATGTA AAAGTCGAAG ATATTAAAAT AGATGTTACG ACAAGTCAAA	1920
45	AAATTGCCCC AGAAATATTA ATTGACCGTA ATGGTAATGC AGAATCTATT AAAGTACTA	1980
	TAAACGATCT TGTCTTTGTG ACAAACGGTA GTATTACAGA AAGCTCTACT TATGGTGATA	2040
50	ATGATACACC AGCGCCACCA ACTGACGAAT TAGGTGGTAG TTGGACACTA TGGAAAAATT	2100
	TAGCGCGACA AAGTCTGAA TTTGGTAATC CTGATAAGTT TTGCCAAAAT ATTCCTAAAA	2160
55	AAAGTTGGTT TGTTCAGCA ACTTCTACAA CAAACAATAA AGAGATTATC GATACAATAG	2220

TCAATGATTC TGCAATGGCAA ATGAGTTTTA CAATCAATCG TCAGCAACAG TTTAAAGACC 2340
 AACCTGAAAA TGAAATATCT ACATGGATTT ATGCCTTATA TTCAGATGTA AACGGCGATT 2400
 5 ATATTAAAAA GCCAATTACA GAATGTAGTG GTAATGAAAT ATGCCAAGAA TGGCTGTATC 2460
 ACTTAGGTGT ATCAACTGAC AAAATTGAAG ACTTAGCAAA ACATGCATCT AATACGATTC 2520
 CTGTTTATAT GCCATATATC ACATCTTATT TCATGACGCG TGCTATCGGC GACAGACCTT 2580
 10 TAGTCGTCCC GCATCAATCT CAGAACTTAG CATTATTGG TAACTTTGCA GAAACAGAGC 2640
 GAGACACTGT ATTTACAACA GAATATTCGG TTCGTACTGC CATGGAAGCT GTTTATCAAT 2700
 TACTAAATAT AGATCGTGGT ATTCCAGAAG TCATCAATAG TCCATTGAT CTTCGCGTCT 2760
 15 TAATGGATGC CATATACGAA CTGAATGACC ACCAAGATTT GCGTGAGATT ACTAAAGATT 2820
 CGAAAAATGCA AAAACTCGCA TTAGCAGGAT TCCTTAAAAA GATAAAAGGT 2870

20 (2) INFORMATION FOR SEQ ID NO: 529:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3105 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 529:

30 CnGTTTAGCT GCAAAAAGTA AATAACGACA CTGTATTGG AATTTTGCAA TTAGAAACAC 60
 TTTTGGGTGA CATTAACTCA ATTTTCAGCG AGATTGAAAG CGAATACAAA ATGTCTAGAG 120
 AAGAAATTTT AATTTTACTA ACTTTATGGC AAAAAGGTTT TATGACGCTT AAAGAAATGG 180
 35 ACAGATTGTG TGAAGTTAAA CCGTATAAGC GTACGAGAAC GTATAATAAT TTAGTTGAAT 240
 TAGAATGGAT TTACAAAGAG CGTCCTGTTG ACGATGAAAG AACAGTTATT ATTCATTTCA 300
 40 ATGAAAAGTT ACAACAAGAG AAAGTAGAGT TGTTGAATTT CATCAGTGAT GCGATTGCAA 360
 GTAGAGCAAC AGCAATGCAA AATAGTTTAA ACGCAATTAT TGCTGTGTAA GTTTAATAGC 420
 AAAAAAGAG GTTTTCATTA AGTTGAAAAC CTCTTTTGTG TGTTGGCATT AATTTTTCAA 480
 45 ATGTTGACTA CTCAATCCTA AATTATAAAT AGTATAGCGC aCAAaTGCTT AAGAAATTTT 540
 TTCTATGGCA CAAATGAATG GAGCATGATT ACGTTGGTTT AAAAATTGAT ATTGCAAAAC 600
 TTGCGCATGC TTTTGATCCA AAGTACTCAA GTAATCAAGC AATGCATGCT TCTCAATTTG 660
 50 TCCTTCGCTA TGACCATGAT ATATAACAAG TACAATAATA CCTTCAATTG ACATTAATGA 720
 TAGCAATGAA TTAATAGCTT GGATTGTCGT GTCAGGCTTT GTCACGATAG ATTTATCACC 780

55

EP 0 786 519 A2

	TATATGATGT	TCAATATTTT	CATGTCCATC	TTTTATTAAA	GAAACATGAT	TGAAATCCTT	900
	AACTTTATCA	CGTGTATTTT	CCAAAGCTAA	ATCTTGAATG	TCGAAACCAT	AAACATGTCC	960
5	TTCTGGTACT	TGTTCCGGCTA	AAAATAAAGT	GTCATTGCCG	TTACCGCAAG	TGCGTCTAC	1020
	AACAATACTT	TCTGGTGTTA	TATGTTGTTT	AATAAGTGTT	TTTGAAAAAG	GGAGTATACG	1080
	TTCTAATTTT	ATTGCTTCAC	CTTAGACTTG	TAACGCAAGC	CTTGATAAGA	ATTTCTACGT	1140
10	GCTAATTCAG	CATCGATGCC	ATTTAATACT	TCCCATTAT	TAACACTCCA	CATTGGACCT	1200
	ACCATGATAT	CTATTGGACC	ATCACCGGTA	ATTCGGTGAA	CGATCATTTT	AGGGGGAATC	1260
	ACTTCTAATT	GGTCCACAAC	TAGGTTTGTG	TACTCTTCTT	GAGTCATAAA	AGTTAATAAA	1320
15	CCTTTATCGT	ATTGTTTTAC	CATCGGTGTA	CCTTTTAACA	AATGAAGTAA	ATGAATTTTA	1380
	ATACCTTGTA	CATCCATTTG	TGCACCTCTT	TGGCAGTAGC	CATCATCATG	TCATAGTCTT	1440
20	CGCCAGGTAA	GCCATTAATG	ATGTGTGTAC	ATACATTGAT	ATTATGCTTA	CGTAATTTTG	1500
	CCACACCATC	ATAATAAGTT	TTCATATCAT	GGGCACGATT	GATTAAATCA	GATGTTGACT	1560
	GATGGATTGT	TTGTAGTCCT	AATTCAACCC	ATAAGTATGT	TCGTTGATTG	AAATCTGCTA	1620
25	AATATTTCGAC	AACATCGTCT	GGTAGACAGT	CAGGACGCGT	ACCAATAGAT	AATCCCACAA	1680
	CACCCGGTTC	TTTAAGTACA	GGTTCGAATT	TTTCTTTTAA	TACTTCAACC	GGTGCATGTG	1740
	TATTTGTAAA	TGCCTGAAAA	TAAGCAATAT	ATTTTCCTTC	GTGCCATTTT	TCATGCATCT	1800
30	TTTCTTTAAT	TTCTTTAAAT	TGTACTGCCA	TTGAATCTGC	ACGATTACCT	GCAAAGTCTC	1860
	CGCTACCTGC	AGCAGAACAA	AATGTACATC	CACCATGTGC	TACAGTGCCA	TCGCGGTTAG	1920
	GACAGTCAAA	CCCGCCATCC	AATGCAACTT	TAAATATTTT	TTGTCCAAAT	TTATTTTTTA	1980
35	AATGGTAATT	CCATGTGTGA	TAACGTTTGT	TTTCAAAGC	GTATTGGAAA	TGATTGCCCA	2040
	TATGTCATTT	TCCTTTCTAT	AAAAAAAGAG	TTCTAAGTAC	AGATTTTAAC	ATATTTTAAT	2100
40	GTTATAGTGT	TTATTATAGT	TTGACAAAAA	AGAGAGAGGA	ACTATGAAAT	ATGAATATAC	2160
	CTAAATCAGT	CTGGTGGCTA	GTAATTGGCA	TGGCGTTAAA	TATTACTGGT	TCCAGTTTTT	2220
	TGTGGCCTTT	AAATACAATT	TATATGAAAC	AAGAACTTGG	AAAAAGTTTA	ACTGTTGCTG	2280
45	GTTTAGTGCT	AATGATAAAT	TCATTTGGCA	TGGTTATTGG	AACTTTATTA	GGTGGTTCAC	2340
	TATTTGATAA	ATTAGGTGGA	TACAAGACGA	TTTAAATTGG	AACTTTCAC	TGTCTTTGTA	2400
	GTACAACGCT	ACTTAATTTT	TTTCACGGGT	GGCCTTGGTA	TGCTGTATGG	CTTGTAATGT	2460
50	TAGGGTTTGG	TGGCGGAATG	ATTATTCCTG	CGATATACGC	TATGGCTGGA	GCAGTGTGGC	2520
	CAAATGGCGG	AAGACAAACG	TTTAATGCGA	TATACTTAGC	GCAAATATT	GGTGTGGCTG	2580

55

ATCTTATTAT GTATGTTGTG TTTGCGCTTG TCGCGGTAAC GCAATTTAAT ATTGAAATTA 2700
 ATGCGAAAGT TAAATATCCA ACTCATTAG ATATTACTGG TAAAAAGAAT AAAGCAAGAT 2760
 5 TTATTTTCATT AGTACTAATT TGTGCAATGT TTGCAATTTG TTGGGTTGCA TATATTCAAT 2820
 GGGAGTCTAC AATCGCTTCA TTTACACAAT CTATTAATAT TTCAATGGCA CAATATAGTG 2880
 TTTTATGGAC AATTAAACGGA ATAATGATTT TAGTAGCACA ACCATTAATT AAACCGATTG 2940
 10 TCTATCTGTT AAAAGGAAAC TTAAAGAAGC AATGTTTGT CGGCATCATC ATTTTATGT 3000
 TGTCTTCTT TGTACAGAGT TTTGCCGAAA ACTTTACAAT ATTTGyTGTC GGTATGATTA 3060
 15 TTTTAACTTT TGGAGaATGT TTGTATGGCC AGCAGTTCCA ACTAT 3105

(2) INFORMATION FOR SEQ ID NO: 530:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5532 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 530:

TAATGATTAA ACCAGATGCA GTACAAAGAA ATCTAATTGG TGAAGTAATT TCAAGAATTG 60
 AAAGAAAAGG ACTAAACTTT GTCGGTGGTA AATTAATGCA AGTACCAATG GAACTTGCTG 120
 30 AAACACATTA TGGTGAACAC CAAGGTAAAC CATTTTATAA TGATTTAATT TCATTTATTA 180
 CATCAGCACC AGTGTTGCA ATGGTAGTTG AAGGTGAAGA TGCAGTTAAT GTATCTAGAC 240
 35 ATATTATTGG CAGCACCAAT CCTTCAGAAG CTTACCAGG ATCAATTAGA GGTGATTAG 300
 GTTTAACTGT TGGTAGAAAT ATCATTACAG GTTCAGATTC ATTAGAGTCT GCTGAACGTG 360
 AAATTAATCT ATGGTTTAAT GAAAATGAAA TTACTAGCTA TGCTTCACCA CGTGATGCAT 420
 40 GGTATATGA ATAAATATA AACTGTAAAC CTTTACGATT TATTTATAAA GGTAGAAAGG 480
 GTTTTGTAT GTGGTTAGTC ATTATGaTTA TACATAACAA GGCCCGTTTT TTATGTTGTA 540
 GTAAATTACT TGAAAAATTT TATAGTTTTT KGGTAACACG TaTtAAAAaG AGAGGAATAT 600
 45 TCTTTATCAA ATGAAACTAA ACAGAGAGAA GGGGTTGTTA AAATGAaGAA TATTATTTCC 660
 ATTATTTkGG GGATTTTAAT GTTCTTAAAA TTAATGGAAT TACTATATGG TGCTATATTT 720
 TTAGATAAAC CACTTAATCC TATAACAAAA ATTATTTTTA TACTGACTCT CATTTATATT 780
 50 TTTTATGTAT TAGTAAAAGA ATTGATTATA TTTTGTaAGT CAAAGTATAA CAAAAGCGCT 840
 TAACATATGT ATATTTTAAT ATCATAATTT TTTTAAACGG ACTGATTAAC TTTATTAATA 900

	GATACGATTA TATTAAAACG GCTAATCATT TTTAATTAAT GATTATATGA TGCAACTGTT	1020
	TAGAAATTCA TGATACTTTT CTACAGACGA ATATATTATA ATTAATTTTA GTTCGTTTAA	1080
5	TATTAAGATA ATTCTGACAT TTAAAATGAG ATGTCATCCA TTTTCTTAAT TGAGCTTGAA	1140
	AACAAACATT TATGAATGCA CAATGAATAT GATAAGATTA ACAACATATT ATAATGTTAT	1200
	CGTGGAAGTA TGAAAGGAGC GAGTGTGTAT GAGATACCTA ACATCAGGAG AATCACATGG	1260
10	ACCTCAATTA ACAGTTATTG TTGAAGGTGT ACCTGCAAAT ATAGAAATTA AGGTTGAGGA	1320
	TATTAATAAA GAAATGTTTA AGCGTCAAGG CGGTTACGGA CGTGGACGTC GTATGCAAAAT	1380
	TGAGAAAGAT ACAGTAGAAA TAGTATCAGG CGTTAGAAAT GGTTATACAT TAGGTAGTCC	1440
15	AATTACTATG GTTGTAACCA ATGATGACTT TACGCATTGG AGAAAAATTA TGGGAGCAGC	1500
	TCCAATAAGT GAAGAAGAAC GTGAAAATAT GAAACGTAAT ATTACAAAAC CAAGACCTGG	1560
20	TCATGCAGAT TTGGTTGGAG GTATGAAATA TAATCATCGT GATTTACGAA ATGTGCTAGA	1620
	GCGATCATCT GCTAGAGAAA CAGCAGCTCG AGTTGCAGTC GGTGCCTTAT GTAAAGTGTT	1680
	ATTACAACAG TTAGATATCG ATATATACAG TCGTGTGTTT GAAATAGGTG GAATTAAAGA	1740
25	TAAAGATTTT TATGATTCAG AAACATTTAA AGCAAATCTT GATCGTAATG ATGTTCTGTG	1800
	AATTGATGAC AGTATCGCAC AAGCAATGCG AGATAAAATT GACGAAAGCTA AAAATGAAGG	1860
	AGATTCAATT GCGGGTGTCTG TTCAAGTTGT AGTTGAAAAT ATGCCTGTTG GTGTAGGTAG	1920
30	TTATGTGCAT TATGATCGTA AGTTAGATGG TAAGATTGCA CAAGGTGTTG TCAGCATAAA	1980
	TGCTTTTAAA GGTGTAAGCT TTGGTGAAGG ATTTAAAGCA GCTGAAAAGC CAGGTAGTGA	2040
	GATTCAAGAT GAAATTCTAT ATAATAGTGA AATTGGTTAT TATCGTGGAT CTAATCACTT	2100
35	AGGTGGTTTA GAAGGCGGTA TGTCAAATGG AATGCCAATT ATCGTTAATG GTGTAATGAA	2160
	ACCAATTCCA ACGTTATATA AACCATTAAA TTCAGTAGAC ATTAATACTA AAGAAGACTT	2220
40	TAAAGCAACA ATTGAACGTT CTGATAGTTG TGCTGTTCTT GCAGCAAGTA TCGTCTGCGA	2280
	ACATGTCGTA GCATTTGAAA TAGCAAAAGC ATTATTGGAA GAATTCCAAT CAAATCATAT	2340
	TGAGCAACTT AAACAACAAA TTATTGAGCG CAGACAATTA AATATTGAGT TTTAACAACA	2400
45	AGAACAATTG AGGTGTAATC ATGAAATTAC AAACAACATA CCCTTCAAAT AATTATCCAA	2460
	TATATGTTGA ACACGGTGCA ATTGACCATA TTAGCACGTA TATTGATCAG TTTGATCAAA	2520
	GTTTTATATT AATTGACGAG CATGTAAATC AATATTTTGC TGATAAATTT GATGATATTT	2580
50	TATCATATGA AAATGTACAT AAAGTTATTA TTCCAGCTGG TGAAAAGACG AAAACATTTG	2640
	AGCAATATCA AGAAACATTA GAGTATATTT TATCCCATCA TGTAACCTCGT AATACAGCAA	2700

55

	ACGAGGCGTG CACTTTATAC AAGTGCCAAC GACTATACTA GCGCATGATT CTAGTGTGG	2820
	CGGTAAAGTG GGTATTAACT CAAAGCAAGG TAAAAACCTT ATCGGTGCAT TTTATCGTCC	2880
5	AACTGCTGTG ATTTATGATT TAGTCTTTTT AAAGACGTTA CCATTGAGC AAATATTAAG	2940
	TGGCTATGCA GAAGTTTATA AGCATGCGTT ATTGAATGGT GAATCAGCGA CGCAAGATAT	3000
	CGAACAGCAC TTAAAGATA GAGAGATATT ACAGTCATTA AATGGTATGG ATAAATATAT	3060
10	TGCTAAAGGT ATTGAAACGA AGCTGGATAT TGTTATTGCA GATGAAAAAG AACAAAGGTGT	3120
	ACGTAAATTT TTAAATTTAG GTCATACATT TGGTCATGCT GTTGAATACT ATCATAAAAT	3180
	ACCTCATGGT CATGCAGTGA TGGTTGGCAT TATCTATCAA TTTATAGTTG CGAATGCTTT	3240
15	GTTTGATTCT AAGCATGATA TTAATCATT AATTCAATAT TTAATACAAC TCGGCTATCC	3300
	TTTAGACATG ATAAGTACT TGGATTTTGA AACGTTATAC CAATATATGC TAAGTGATAA	3360
20	AAAGAATGAT AAGCAAGGTG TACAAATGGT CTTGATTAGA CAATTTGGAG ATATCGTTGT	3420
	ACAACATGTT GATCAACTAA CATTACAACA TGCATGTGAA CAATTAAAAA CATATTTTAA	3480
	GTAGGTGAAT GAAATGGTAA ATGAACAAAT CATTGATATT TCAGGTCCGT TAAAGGGCGA	3540
25	AATAGAAGTG CCGGGCGATA AGTCAATGAC ACACCGTGCA ATCATGTTGG CGTCGCTAGC	3600
	TGAAGGTGTA TCTACTATAT ATAAGCCACT ACTTGGCGAA GATTGTCGTC GTACGATGGA	3660
	CATTTTCCGA CTGTTAGGTG TAGAAATCAA AGAAGATGAT GAAAAATTAG TTGTGACTTC	3720
30	CCCAGGATAT CAATCTTTTA ACACGCCACA TCAAGTATTG TATACAGGTA ATTCTGGTAC	3780
	GACAACACGA TTATTGGCAG GTTTGTTAAG TGGTTTAGGT ATTGAAAGTG TTTTGTCTGG	3840
	CGATGTTTCA ATTGGTAAAA GGCCAATGGA TCGTGTCTTG AGACCATGA AACTTATGGA	3900
35	TGCGAATATT GAAGGTATTG AAGATAATTA TACACCATTA ATTATTAAGC CATCTGTCAT	3960
	AAAAGGTATA AATTATCAAA TGGAAGTTGC AAGTGCACAA GTAAAAAGTG CCATTTTATT	4020
40	TGCAAGTTTG TTTTCTAAGG AACCGACCAT CATTAAAGAA TTAGATGTAA GTCGAAATCA	4080
	TACTGAGACG ATGTTCAAAC ATTTTAATAT TCCAATTGAA GCAGAAGGGT TATCAATTAA	4140
	TACAACCCCT GAAGCAATTC GATACATTAA ACCTGCAGAT TTTCATGTTT CTGGCGATAT	4200
45	TTTCATCTGCA GCGTTCTTTA TTGTTGCAGC ACTTATCACA CCAGGAAGTG ATGTAACAAT	4260
	TCATAATGTT GGAATCAATC CAACACGTTT AGGTATTATT GATATTGTTG AAAAAATGGG	4320
	CGGTAATATC CAACTTTTCA ATCAAACAAC TGGTGCTGAA CCTACTGCTT CTATTCGTAT	4380
50	TCAATACACA CCAATGCTTC AACCAATAAC AATCGAAGGA GAATTAGTTC CAAAAGCAAT	4440
	TGATGAAGTG CCTGTAATAG CATTACTTTG TACACAAGCA GTTGGCACGA GTACAATTAA	4500

55

	AAACTTGTTA GGGTTTGAAT TACAACCAAC TAATGATGGA TTGATTATTC ATCCGTCAGA	4620
	ATTTAAAACA AATGCAACAG TTGATAGTTT AACTGATCAT CGAATAGGAA TGATGCTTGC	4680
5	AGTTGCTTCT CTACTTTCAA GCGAGCCTGT CAAAATCAAA CAATTTGATG CTGTAAATGT	4740
	ATCATTTCCTA GGATTTTTTAC CAAAACATAA GCTTTTAGAA AATGAGGGAT AATATAAAAT	4800
	GGAAGATATC TATAAATTAA TAGACGATAT CAATCTACAA AACTAGAAA ATTTAGACTC	4860
10	TCGTGTTAAT GAAGCAATAA CTA CTGACAA CGATGACGCA TTATTTATTC TAGGAGAGAC	4920
	ACTTTACAAT TTTGGATTAA TGCCaCAAGG TTTGGAAGTA TTCCGCGTGT TATATCACAA	4980
15	ATATCCAGAC GAAAGTGAAT TGCTGATTTA TTTTATTGAA GGTTTAATGT CTGAAAATCA	5040
	AACTGACGAA GCGTTAGAAT ATTTAAGTTA TGTGTAACCA TCACCTGAAA AGTTGATGTT	5100
	AGAAGCAGAT TTATATCAAC AAATTAATAT GATGGAAGTT GCTATTGATA AATTACAAGA	5160
20	AGCACTTGAA CTAGAGCCAA ATGATCCAAT AATCCATTTT GCATTGGCTG AAATGTTATA	5220
	TTATGATGGT CAATATTTAC GTGCTACCTC TGAATACGAA ACCGTTTTAG AACTGGTGA	5280
	ATATCAAGTT AATGGTGTA ACTTATTCTC TCGTATGGCA GATTGTAGTT TACAAAGTGG	5340
25	kAACTATAGT GATkCCGATt CgCTTATACG ATGzAATTAA TGAAGATGAA ATGACTTCAG	5400
	AAGATTATCT CAAAAGAGCC ATTTCTnACG ATAAAAATGA CATCACTCAA GAAGCAATTA	5460
	AAATAATGAC TACATTACTT TCTAAAGATC CTGATTATAT TCAAgGCTAC TTGTATTTAC	5520
30	aATCaTTATA TG	5532

(2) INFORMATION FOR SEQ ID NO: 531:

(i) SEQUENCE CHARACTERISTICS:

35	(A) LENGTH: 942 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 531:

	AATTGGGTTA TACTATAGGT AAATTTAAGG AGGTAAGAAA ATGGATAAAA AAGAATTAGC	60
45	GAAATTTATA GGCAATAAAA TCAGATACTA TAGAACCaaa TTGaACTTAA CTCAAGATCa	120
	ACTTGGAGAA AAACTCmACa CTAAAAArGC TACTATTTCA AATTATGAGA CAGGGTACAG	180
	AACTCCTAAA CAAGATGATT TGTTTGAAAT TGCTCATATT TTAAATATCA GTATCGATGA	240
50	TTTGTTCCTT ACAAGAAATA ATAAAAAaCG CGACATCACT TCCATATACA ACAAACTCAC	300
	ACCTCCCCGC CAAGAAAACG TACTTAACTA CGCAAATGAG CAATTAGATG AACAGAATAA	360

55

AACTGGTGCT GGCATAGGAG AAGAATTATA TGATGACATA TTGCATGAAG AAGTATTTTT 480
 TAAAGAAGAC GAAACGCCAT CAAATGCTGA TTTTGTATT TTAGTTAATG GTGATTCAAT 540
 5 GGAACCTATG TTAAAACAAG GAACATACGC TTTTATTAAG AAAGAAGATT CTATTAAAGA 600
 TGGTACAATT GCACTCGTTG TATTAGATGG AGTAAGTCTT ATCAAGCGTG TAGATATATG 660
 CGAAGACTAT ATTAATTTGG TATCTCTAAA TCCGAAGTAT GATGATATCA AAGTCGCTTC 720
 10 GTTTAGTAAT ATTAAAGTAA TGGGCAAAGT TGTATTGTGA TTAATAGCGC CTATATGGCA 780
 CTTTAATATA AAAGACGTCT ATTTACAGCAG TGTTTAAAG GAGTTTATAA TGAAAATAAC 840
 TAATTGCAAA ATAAAAAAG AAACATAGT ATATGAAGTT TTAAC TAGTG GTAATCAACC 900
 15 ATTCACCTAT GAGTTACCTA AAGATTTATC GTCACATAAT GC 942

(2) INFORMATION FOR SEQ ID NO: 532:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 417 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 532:

TGGnAAATGC AAACCAAAT ATGATCCTCG TGcAAGTTGA AGCGGGACGT TATGAAGAAT 60
 30 GGGTAAAGAA TGGTTATTTT AAACCGTCAG AAGATAAATC AAAAGAAACA TATACAATTG 120
 TTATCCCGCC ACCAAATGTA ACTGGTAAAT TACATTTAGG ACATGCATGG GATACGACTT 180
 TACAAGATAT CATTACACGT ATGAAACGTA TGCAAGGATA CGATACGTTA TACTTACCAG 240
 35 GTATGGATCA TGCTGGTATT GCGACACAGG CAAAGGTAGA AGCTAAATTA AATGAACAAG 300
 GAATAACTAG ATATGATCTT GGTCTGTAAG AGTTTTTAGA ACAGGCATGG GATTGGAAAG 360
 40 AAGAGTATGC GTCATTTATT CGTGCGCAAT GGGCTAAATT AGGTCTAGGT TTAGATT 417

(2) INFORMATION FOR SEQ ID NO: 533:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 533:

GATCCTGAAC CCGCATTTGT TTCCACTAAA ACAGTATGCC CACTTTCTAC TAAAGCGTGC 60

ATTTTCATAC CATCCACCTC CATAATCATC TTAACGCGAA CATTTTGAAA GCGCAATCAA 180
 AAATCCACAA AATTGTAAAG GTTATTACAC TGACTTTTCC GAAAATTGTG GTAAATATA 240
 5 ATTAAGAAAG AACAAGGAGG CACTTACTAT GATTACTTAC AAAAAATATT TAATCGCAGT 300
 TGACGGTTCA CATGAAGCGG AATGGGCATT TAACAGAGCA GTTGGTGTTG CTAAACGTAA 360
 CGATGCGAAG TTAACAATTG TGAATGTAAT TGATTCAAGA ACGTATTCTT CTTATGAAGT 420
 10 TTATGATGCT CAATTTACTG AAAAATCTAA GCATTTTGCA GAAGAATTAT TAAATGGTTA 480
 TAAAGAAGTA GCTACTAACG CTGGTGTTAA AGATGTAGAA ACGCGTCTAG AGTTTGGyTC 540
 15 TGyTAAATCT ATCATTCTTA AAAAGCTTGC ACATGAAATT AATGCAGACT TGATTATGAG 600
 TGGTACATCA GGCTTAAATG CCGTGAAAag ATTTATTGTT GGTTCTGTAT CAGAATCTAT 660
 CGTTCGTCAT GCGCCATGTG ACGTGTTAGT TGTTCGTACT GAAGAGTTAC CAGCAGACTT 720
 20 CCAACCACAA GTT 733

(2) INFORMATION FOR SEQ ID NO: 534:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 6060 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 534:

TGATGATCCT GAAGCGCTAT TGGATAATTA CAACACTGAA GATGTTGATG CACACAATTA 60
 35 CAATAATATA AATCATGTTA TTTTGCCTG CGATGCGGGT ATGGGTTCTA GTGCAATGGG 120
 TGCArGCATG TTACGTAATA AATTTAAAAA GGCGGGCATT AATGATATTA CAGTTACAAA 180
 TACTGCGATT AATCAATTGC CAAAAGATGC TCAATTAGTT ATCACTCAGA AAAAATAAC 240
 40 TGATCGTGCT ATTAAACAAA CACCAAATGC CATCCATATT TCAGTGGATA ATTCCTTAA 300
 TTCACCAAGA TATGAAGAAC TTTTAAATAA TCTAAAAAAA GATGATCAAG CATAATAATT 360
 AAATAAATTA AAAATGGAG GATACCGCCA TGTTATTGAG TACACGTGAA AAAGAAATGA 420
 45 TAGCCctAtT GATTAAAGTAC CACGGtCAAt ATATCACTAT ACACGACATT GCTCAGCAAC 480
 TTGCGGTGTC CTCTCGTACT ATTCACCGTG AATTAAAAGG TGTTGAAGCA TATTTAACTT 540
 CATTTTCATT AACTTTAGAA CGCGCAAACA AAAAAGGGcT ACGCATTGCT GGCACAGATT 600
 50 CTGATTTAAA CGATTGAAG CAATCGATTG CACAACATCA AACCATTGAC TTATCTGTTG 660
 AAGAGCAGAA AGTAATTATT ATATACGCTT TGATACAAGC CAAGGAGCCA GTTAAACAAT 720

55

	TAGAGCTTGA	TTTAAATAAG	TACCAACTAT	CTTTATCTCG	AAAGCGTGGC	GAAGGCATTT	840
	ACTTGGTAGG	TACTGAATCA	AAGAAACGTG	AATTTTAAAG	TCAATTAATG	GTGAATAACT	900
5	TAAATAGTAC	TAGCGTTTAT	TCAGTAATTG	AAAATCATTT	TGTCTTTCAT	TCATTAAATC	960
	AAATCCACAA	AGACTTTGTT	GACTTAGAGC	GCATTTTAA	TGTTGAAAGA	CtATTaATGG	1020
	ACTACCTAAG	TGCCTTACCC	TACCAACTTA	CCGAATCAAG	TTATTTAACT	TTAACTGTCC	1080
10	ATATCGTGCT	CTCCATTTCA	CGTATAAAAA	ATGGAGAGtA	TGTCGCATTA	AACGATGATA	1140
	TTTATGATTC	TGTACAAAAC	ACATTTGAAC	ACAAAGTaGc	AAGCGAACTT	GCTGATAAAC	1200
	TTGGTCAAAT	ATATGACGTC	ACGTTTAATC	AGGCAGAAAT	TGCTTTCATT	ACTATCCATT	1260
15	TACGTGGAGC	TAAACGAAAA	AATCTTAATG	ATACATCATT	AAATAATCGT	TGTGAAGAAA	1320
	ACAAAATTAA	AGCGTTTGTG	AACAAAGTAG	AAATGATTTT	CGGTATGACA	TTTGCAGATT	1380
20	TGGATACTTT	AGTAGATGGA	CTGACGCTAC	ACCTTAATCC	TGCAATCAAT	CGTTTGCAAG	1440
	CTAATATCGA	GACCTATAAT	CCGTTAACAG	ACATGATTAA	GTTCAAATAT	CCAAGACTAT	1500
	TTGAAATGT	AAGATTAGCT	TTAAATGATT	GTTGGCCTGA	TTTGATTTTT	CCAGAGAGTG	1560
25	AAATTGCTTT	TATAGTtTTA	CACTTTGGTG	GCTCGATTAA	AAACCAAGGT	AATCGATTTT	1620
	TAAACATATT	AGTCGTTTGC	AGCAGTGGTA	TGGGAAGTAG	TCGTCTATTA	TCAACTCGTC	1680
	TAGAGCAAGT	TTTTAGTGAG	ATTGAGCGTA	TTACACAAGC	ATCAGTCAGC	GATTTGAAGT	1740
30	CACTAGATTT	AAGTCAATAT	GATGGCATT	TTTCTACTGT	GAATTTAGAC	ATCGACTCCC	1800
	CCTATTTTAA	GGTAAACCCA	TTATTACCAG	ATAGTGATAT	CAGTTATGTC	GCACAGTTTT	1860
	TAAATACAAA	GTCTACGTTT	CAAGAGACGC	ATGATAAATC	ATCAAACATG	ATTGATAAGG	1920
35	ATGATGTTCA	TGTTGAAACG	AAAGATGTTG	ATGGCAACAC	ATCTTTTGAA	AATGAACAAA	1980
	CTTCATACTT	AACTTCAGTT	TTCGAAAAAC	ATTTAAGTGA	CGAAAAATCA	GAACAATTAT	2040
40	TGCATCATAT	GCGTTCGGGT	TTAACTTTGC	TTGATTCACT	GAAAATAGTT	AGTACCGAAG	2100
	TTAAACAGTG	GCAAACATAT	ATCGCAGATT	ATCTATATCA	ATGCGATGTA	ATAAACGATC	2160
	CAACGTCATT	CGCTGAACTA	CTAGAGCAAC	GATTGATTGA	CAATCCAGGA	TGGATATTAA	2220
45	GTCCATATCC	TGTTGCAATA	CCACACCTAA	GAGACAATAT	GATTAAACAC	CCTATGATTC	2280
	TAATCACAGT	TTTAGAAGAA	CCGTTAACAT	TGCCTAGTAT	TCAAAATGAC	AATCAAACAA	2340
	TTAAATATAT	GATTTCCATG	TTTATTTCTG	ACAATGATTT	TATGGCATCA	CTGGTAAGTG	2400
50	ACTTGTCCGA	ATTTTTAAGT	TTGAAATTAG	AATCTATTGA	TACTTTTATG	GAAAATCCAC	2460
	AGGAACTTGA	AACATTATTA	AGAAACAAAT	TTTTAGAACG	AATTAAAAAA	CAATTTATTT	2520

55

EP 0 786 519 A2

	TAACAGCCAA AATGAAGCAA TTGAAAAAGC AGGTAAAGCC TTAGTTGATA GTGGTGCTGT	2640
	AACAGATGCT TATATTCAAG CAATGAAAGA TCGTGAGCAA GTCGTATCAA CATTTATGGG	2700
5	AAATGGCTTA GCAATTCCTC ATGGCACAGA TGAAGCTAAA ACAAATGTGA TTCACTCAGG	2760
	TTTAACATTA TTACAAATCC CTGAAGGCGT TGAAGGCGT GCGGAAGTAG TTAAAGTTGT	2820
	CGTGGAATT GCTGGTAAAG ATGGCGAACA TTTAGACTTG TTATCTAAAA TTGCAATTAC	2880
10	ATTTAGCGAA GAAGAAAATG TGGATCGTAT CGTTCAAGCA AAATCTGCAG AAGAAATTAA	2940
	ACAAGTATTC GAGGAGGCAG ATGCATAATG AAAGCAGTTC ACTTTGGTGC TGGTAACATA	3000
	GGTCGTGGTT TCATTGGTTA TATTCTgCAG ACAACAATGT TAAAGTAACA TTTGCAGACG	3060
15	TCAATGAAGA AATCATTAAAT GCTTTAGCTC ATGATCATCA ATACGATGTT ATTTTAGCTG	3120
	ATGAGTCTAA AACAACGACG CGCGTGAaTA ATGtTGATGC AATTAATTCA ATGCAACCTT	3180
20	CTGAAGCGTT GAAACAAGCA ATTCTAGAAG CTGATATTAT TACAACAGCT GTTGGTGTTA	3240
	ACATACTACC TATTATTGCT AAATCTTTTG CGCCTTTCTT AAAAGAAAAA ACAACCATG	3300
	TTAATATTGT TGCTTGTGAG AATGCTATTA TGGCAACTGA TACATTGAAA AAAGCAGTAC	3360
25	TTGATATTAC TGGCCCTCTT GGTAACnaTA TTCATTTTGC TAACTCAGCA GTTGATAGAA	3420
	TTGTACCATT ACAAAGAAT GAAAATATAT TAGACGTTAT GGTGAGCCA TTTTACGAAT	3480
	GGGTTGTTGA AAAAGATGCA TGGTATGGTC CAGAACTAAA CCATATTAAA TATGTTGATG	3540
30	ATTTAACACC ATATATTGAG CGTAAATTAT TAACTGTGAA TACAGGACAT GCATATTTAG	3600
	CGTATgCTGG tAAATTTGCA GGTAAAGCTA CAGTTTTAGA TGCAGTTGAA GATAGTTCAA	3660
	TTGAAGCTGG CTTACGCCGT GTTTTAGCTG AACTAGTCA ATATATTACT AATGAATTTG	3720
35	ATTTTACTGA AGCGGAACAA GCTGGTTATG TTGAAAAAAT AATAGATCGT TTCAACAATT	3780
	CITATTTATC TGATGAAGTA ACACGTGTCG GACGAGGTAC ATTACGTAAA ATTGGCCCTA	3840
40	AAGATAGAAT TATAAAACCA TTAACATATC TTTATAATAA AGATTTAGAA CGCACTGGTT	3900
	TATTAAATAC AGCTGCATTG TTATTGAAGT ATGATGATAC AGCAGACCAA GAAACTGTTG	3960
	AGAAAAATAA TTACATTAAA GAACACGGTT TAAAAGCGTT TTTAAGTGAA TATGCTAAAG	4020
45	TTGACGATGG CTTAGCCGAT GAAATAATTG AAGCGTACAA TTCACTTTCA TAATTTATTG	4080
	AGCTTTGTTT GAAACAAGAA GTTTCCAACG TTATTCGTTA ACAATCAGTA ATAATGTAGT	4140
	AGTTCCCTTG AATTAACAAT ATTAAATTTT TGAACATAAA AAATACTCCC TTCAACATAG	4200
50	ACACTTAACT TGTGTTATGT ATGAAAGGAG TATTTTTGCG TTAATAATTT GTTTTATTTT	4260
	CGAGCCACAG CCACCTATTC AATGGCTATT GGTCACTACT AAAACAAATT CATATTAACT	4320

55

	TTGAATAAAT TTTATTCTTC AGTTTGTTGG TCTTTCTTAG TGAATCTTCT AATTAAGAAT	4440
	GCCATACCTG CACCTAGAGC TAATTCAGCA TATGGTAAAT CGTCATTATG TGACATACCA	4500
5	GTATCTGGTA AAGTTTTAGC TTGTTGTTTA GCTTTATTAA CTTTTCCTTG TTGAGCTGAT	4560
	TTTGTCTTAG CTGGGTGGTC GTCAGTGTTA GTTACATTAA GCATATCTTG ATTAGCACTA	4620
	TTGCTTCCAT TTGAAACTGT AGCTGGAGAT GCATTGGCAC CGTCGTTTTG CGTAgyTTTA	4680
10	TTGTTTGCAG CTGAACCAAC TGATTTTTGC GTATCATTAG TATCTGCTGT TGCCGTATCA	4740
	TCTTTTTGGC TAACATTAGT TGAAGTCATT TTTTCTTTTG CTTCAGAAGA TGCAGATGTT	4800
	GATGGTTTAT TCGAAACTTC AGTATCAGCT TTGCTTGGCG ATTTATCTGC TTCGTTAGAT	4860
15	GCAACGTTAG TTTCAGACTT AAGTTGTCCT GCATCAGTTT GATTTGTCGT ACTTTCTTCT	4920
	TTATCTTTTG ATGTATTAGA AGGTACATTT GGTCTGTTA TGTCTGCTGA AGGCAATGTT	4980
20	TCAGTTGTtG ATTCAACCAT ACTTTGATTT GTTGAATCAC TACCATCTTT TTCTGCCTTA	5040
	GCTTTATTTT CAGATTTTGG TTGTGCAACC TTGTCATTAG TTGATTGAGA TTCAGCACTA	5100
	TTATTTACTT CAGCATTTTG TTTTGAATCA TTTACAGATG CATTATCTTT GCTATCAGCA	5160
25	GATGATGCTG CTTCTGTGCT CGCAGTTGTT GGAGCCGTTG CTGTTGATCC TGTGGTGCA	5220
	TTCTCGTTTG TTGCTGTAGT TGTACTATTG TTATTTGTTG TGCTTTCTGC TGGCGTTGCA	5280
	TTATCAGTTT CTGTTACAGG TTTATCAGTT GTGCCGTTAT TAGTTGATTG TACTTCTGGT	5340
30	TTACTAGTTA CATCGTTATC CATTGTCCGA CTGTTTGTTG ATGCATCTAC ACTAGAATTG	5400
	TTATTAGCTT GCGGTTTATC APTTGCATCA TCAGTTGCTG ATGTTGCTGT TGTTCACCT	5460
	GTTGCCGCAT CACTATTATT TGGTGTGTGC GGAGAAGCGT CTGCTTTGCC ATTAGCTGTC	5520
35	GTCTCAGATA CGTTAGGTTG TCCAGTATTT TCTGGTGTTG CATTAGCATT TGAATTTGCT	5580
	GTTGCATCAT TATTATCTAT ACCATTATTA GTATCATTAG CATCTGGATC ATTCTGAGGC	5640
40	ACAATCGCTT CAATTGCAGG TATCGTTACA TTTTGTAATT CAGCAACTTC TGCATTTGTT	5700
	TGTGTTTTAT CTAATTTATC AGCAAATCTG TCAAATATC TACCTAAATC CGTACGTGCA	5760
	ATTTCTTTTCG CCGATGCATC TGCATCTGCA TTTTAAATTA TTTCTATTTG CTTGTTAACC	5820
45	ACTTCTCTGA TTGCTTCCAA AGCATTTTTT TTAAC TTCAG GATTAATACG TTGTGCTTTA	5880
	AGTTGTTCAA GCGCACTATT TTTGACAGTA GCGATTTCTG CATTTGTAGT TTGATCAGAA	5940
	ATATCTTCAG TTGCTTTTGA TAAAATGTCT TCTAAAGCAT TCGTAAACGC TTCTTTTCT	6000
50	TCAGTTGTAG CATCAGCGTT GACATTTACA CCTGCTTCAA TCTGGTCTAG TGCAGTTTCT	6060

(2) INFORMATION FOR SEQ ID NO: 535:

55

(A) LENGTH: 977 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 535:

10	AACAAAGCCT TCCAATTATC TCGTCGGTA GAACAAGTAT TAGCAACTTT ATCACCTACG	60
	CTAAACAGTC CTTACGATTT ATACGGCAGC AAAAAATGC TAGATATTAC ATTCGATTCA	120
	TTTGAACATG ATGGTACAAC GTACCCTGTC GACTATGCTA CGTTTGAAAA TGATTATGAA	180
15	GATAATAAAG ATCCTGAGTT TAGACGTAAA AGTTTCAAAT CGTTTAGCGA TGGGATTGCA	240
	AAATATCAGC ATACTACCGC GGCTACATAT AATATGCAAG TACAACAAGA AAAAATTGAA	300
	GCTGATTTAC GTGGATTTGA ATCAGTCATC GATTATTTAT TACATAGTCA AGAAGTAACG	360
20	CGTGATATGT TTGACCGTCA AATCGATATG ATTATGCGTG ACTTGGCACC AGTTATGCAG	420
	AAATATGCTA AACTTTTACA ACGTATTCAC GGATTAGATA ACATGCGTTT TGAAGACTTG	480
	AAGATTTCTG TAGACCCTGA TTATGAACCA GAGATTTCAA TTGAAGACTC AAAAAATTAT	540
25	ATTTTCGGTG CGTTAAGTGT TTTAGGTGAT GACTATACAA ACATGTTACG TGAAGCATAC	600
	GATCAGCGAT GgATTGATTT TGCACAAAT AAAGGTAAAG ATACAGGCGC ATTTTGTGCA	660
30	AGTCCATACT TTACACATTC ATATGTGTTT ATTTCTTGGA CTGGTAAAAT GGCTGAAGCA	720
	TTTGTCTTAG CACATGAATT AGGTCATGCA GGTCATTTTA CATTAGCTCA AAAACATCAA	780
	CCATATCTTG AATCAGAAGC ATCAATGTAC TTTGTTGAAG CCCCTTCTAC AATGAATGAA	840
35	ATGTTGATGG CCAATTATTT ATTTAACACA AGTGATAATC CAAGATTTAA GCGTtGGGTT	900
	ATTGGCTCAA TTTTATCTAG AACATATTAT CATAATATGG tACCCmTTTA TTAGAAGCnG	960
	CTTATCCACG GGGAGTG	977

40

(2) INFORMATION FOR SEQ ID NO: 536:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1440 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 536:

50

	AGACAGTGAT yGaATTTCAT TTACAGTACA CAAATCATCG AAAAAATTGGT AACATTCTTC	60
	TCTATTTTCT AACGTTAgwA TTgCATCAAA CAATTCATCT AACGCTGCAC CTCGTAATTT	120

55

AATCTATTAA AGTGTAGCGa TTTATATTTT ATTAAATCTG AATCGTTACT TTATTTAATT 240
 TTATGCTAAT CCAGCGCGTT CGAAAATAGT GTCAACTTGa TTCAAATGAT GTTTAGGATC 300
 5 GAAACATTCA TCCAATTCTT CTTTGTGTAA AACACTTGTA ATAGACTCAT CTTGTTTCGAT 360
 TAATTCACGG AACGGTGTIT TCGTTTCCCA AGATATCATC GCTTTTGGTT GTACTTTGTC 420
 GTATGCTTCT TCACGAACCA TACCTTTATT AATTAATGCT AATAAGACAC GTTGTGAGAA 480
 10 AATCAGACCA AATGTTTTAT CTATGTTATT ACGCATATTA TCTTCAAATA CAGTTAAACG 540
 GTCCACAATA TTTGTGAACG ATTCAATGCA TAATCTAGTG CtATTGTAAC ATCTGGTAAC 600
 15 ATAATACGCT CAGCAGAAGA ATGAGAAATA TsTCTTTCAT GCCATAATGG CACATTCTCA 660
 TAAGCTGTAG TAATATAACC ACGAATGACT CTTGAAATAC CTGTGATATT TTCAGAACCA 720
 ATTGGATTTC GTTTATGAGG CATTGCAGAT GAACCTTTTT GGCCTTTTGC AAATGCTTCT 780
 20 TCAACTTCTC TCGTTTCGGT TTTTGAAGG TTACGTATTT CAACGGCAAA TTTTCTAGT 840
 GATGTCGCGA TTAATGCTAA TGTCGCAATA TAGTATGCAT GTCGATCGCG TTGCAATGTT 900
 TGCGTTGATA CAGGCGCTGT GCCAATACCT AAATGTTTAC ACACATAACT TTCTATTTCA 960
 25 GGAGGAATGT TAGCAAAAGT ACCTACTGCA CCACTCATTT TCCCTACTTC AATTCTTCT 1020
 CTTACTTGTT TGAAACGTTG TAAGTTACGT TGCATTTCCG TGTACCACAA TGCCATTTTG 1080
 ACACCAAATG TAGTTGGTTC TGCATGCACT CCATGTGTAC GTCCCATCAT CAATGTATAT 1140
 30 TTATAATTTT TTGCTTTTTT AGCTAAAACG TCGATAAATC TTTCTAAATC TTTTCAATA 1200
 ATGTCATTTG CTGTTTTAAT AACGAACTT AAAGCTGTAT CTACAACATC AGTAGAAGTT 1260
 AAACCATAAT GTACcACTTA CGTTCTTCAC CTAGCGTTTC AGAACTTGT CTAGTAAAGG 1320
 35 CTACAACATC ATGGCGCGTT TCTTGTTCAA TTTCTGTGC ACGTTCGACA TTTACCTTTG 1380
 CGTTTTGACG AATTTTTTGT ACGTCAGCTT TCGGTATATG TCCTAATTCA CTCCATGCTT 1440

40 (2) INFORMATION FOR SEQ ID NO: 537:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 784 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 537:

50 GATAAATCTA TnCAGTTTCC GTCCAAATAT CtGCaCCTAA AGCTTTTAAg TGTTCTACAA 60
 TATCTgTATA ACCTCTATAA ATATGTTTAA CATTGTAAAT TGTAGTTACA CCCTCAGCAA 120

55

CATGTAATGT TGATGGTTTT ATCGTTGCTG TGCCTTCGTC AACTTCAATA TTTGCACCCA 240
 TGGCGCTTAA TTCTTCAACA TGTTTAAAC GCTCCGGATA AATCGTATCA GTTACAAATG 300
 5 AAGGACCATT TGCCATAAAT AATAATGGTG TAATAGGCTG TTGCAAATCA GTAGCAAAAC 360
 CTGGATATAC TAGTGTTTTA ATATCAACAA ATTGATATGG CGCATTATTA TTGATGCGAA 420
 10 TTCTTTCGTC TCTTACATCA ACATTCACAC CTAATTCACT AAATTTAGCA GTTAATGTTT 480
 CTACATGTTT CGGAACAATA TTATTTAATA TAACATTTTC TCCACATGCT GCAGCGATAC 540
 ACATATATGT GCCTGCTTCA ATTCTATCAG GTATAACTTG ATACTCAGAA CCATGTAATT 600
 15 CTTTGACGCC ATTGATTTTA ATTGTTGATG TACCCGCTCC CTTAATATTA GCTCCCATAC 660
 TTGTTAAGAA GTTAGCAACA TCAACTACTT CCGGTTCTTT AGCAGCATTT TCAATTACAG 720
 TTTGTCCTGT TGCATAAACT GCAGCTAGCA TAATGtnAAT TGTTCACCT ACGCnaACCA 780
 20 TATC 784

(2) INFORMATION FOR SEQ ID NO: 538:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 3733 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 538:

CAATCTCCAC CAATGACACG TCGTCATTCTG TGACCTCATA CCArACAAAA AACAGTCTCG 60
 CAATCAAGAC TGTTTTCCAC TCAATATATT CATCCATTAG CGTAATAGAT TATTTGACTT 120
 35 CTGTAGCTAC AAAGaTTTTA CGTTTTTCCC AAACGCCTGT cTTTTcATTG TAATCATCAC 180
 AAGTAATTAA TGTTAATTGT TTATCTTTAC CTTTTTGTTT ATCTAGAACT CCTACATCTG 240
 40 TAGGCCTAAC ATCTCTTATA CTTGTCAATT TATACTTACG TGTTTCATTA CCAACTTTAA 300
 AGTACACCAT ACTACCTTTT TTGGCTGCTT TAAGATTTGT AAATTGATAG TTCGGACGGT 360
 CAATGAAAGT GTGTCTGCA ATTGAAATAT TTTGATCATC TAGTGATTCA TTTCTTCTG 420
 45 CAAAGCTTAC ACCTCTATTT AATTGTTTCTG GTGTTGCTGG TCCTGGATAT ACTGGTTCTT 480
 TAATATCAGC ATCTGGAATT TCAATATAGC CTGCCACTTT CGATTTATCT TTCGGAATTT 540
 GAGGTTTAGC TTGCTGCTTT TTATCTTTAC TCGCCTGTTT TTTTACATTT TTATCATATT 600
 50 GTTCAATCTT TTCATCTTTA TCTTTATCGT GAAGATAATT ATCGATATGT GGTTTAGCAA 660
 ACAAATATGC TGCCACTAGG aTAAGTACCA CACCAGCGAT TGTCATTAAT CGATTTGTCC 720

55

	AGTATACCAT TAATTTCAAA ATGACTCATA GCAATTCATT TTATACTATA AAATTTACAT	840
	GTATACTTTT ACGTTAGATT TCATTACACA TATTTGCATT CAAATAACGA AACGCTTTTA	900
5	ATAATTACTA AGGGGGAATT GATATGATTA GATACGCTAA AAAAGAGGAT TTAAACGCTA	960
	TATTAGCGAT ATACAATGAT GCCATTATCA ATACTACAGC TGTTTATACT TATGAACCAC	1020
	AAACCATAGA CGAACGTGTC GCATGGTTTG AAACGAAACA ACGTAAGCAT GAGCCTATCT	1080
10	TTGTATTTGA GGAAAATGGA AGTGTCTTAG GGTTCGCGAC GTTCGGTTCA TTTAGACCTT	1140
	GGCCAGCATA CCTATATACA ATCGAACATT CTATTTATGT CGATGCTTCA GCTAGAGGAA	1200
15	AAGGTATTGC TAGTCAATTA CTACACCATT TAATTGTGGA AGCAAAAGCT AAAGGTTATC	1260
	GTGCGCTAGT TGCAGGCATT GATGCTTCCA ACAAAGCGAG TATTCAGTTG CATCAAAAGT	1320
	TTGCTTTTAA GCATGCCGGC ACACTGACCA ATGTAGGTTT TAAATTTAAT AGATGGTTAG	1380
20	ATTTAGCATT TTACGAATTA GATTTACAAG ACTAGTAATG TTTGAATCAC ATAATATAAA	1440
	CAAGACAACC ATGTTAATTC CCTTAACATA ACAAGCCAAC ATATAAAATT TTAACTTCT	1500
	CAGGGGAGTG GGACAGAAAT GATAAAGAGC CACTAATGAT TTATTATGTA GTGGTTCTTA	1560
25	CACATTAGCC ACAGCTAATG TGTACTTAAA AATAGGAATA CATGAGTAA ACTCATGCAT	1620
	AAGAAATACT AATTTCTATA GAAAAAGTAT TTCTTTATCG TCGTCCCACC CCAACTCGCA	1680
	CATTATTGTA AGCTGACTTT TCGTCAGCTT cTGTGTTGGG GCCCAAAAAG CTTGTTACAA	1740
30	GCGCATTTTC GTTCAGTCAA CTACTGCCAA TATAACTTTG TAGAGCATAT TACATTGATT	1800
	TACATTGTCC CTTTTATTTA TTCTTTTCAA ATACTATCCC CATAGCTTTG ATTTAACGCT	1860
35	TTTTCTCAAT AACAAAACGA ATATAGTAGA ACATGAAAAC GATAATCATG CTGAGCGATA	1920
	AAGATTTAAA TAATAGATTG ACCCACGTTT CCTCAGTCGT ATATCCATAT GTAATCGTTG	1980
	TGTTAATGAT GAATGCTATA AAGATGATTG ATAGTCTTAG CATATCATCA CTCCTTTTAA	2040
40	GTTATTTTAG ATATACGGGG GCGCTTTTGC AATCACTATT TTGATTAGTA TGCATTTTCC	2100
	ATAAATCTTT CAACTTCTTC AGAGATAATT AAGAAGCATC TATCTGGTAC TAATGATCCA	2160
	GACAGATGCT TCTTTTTTAT CAATATTTTA TTGTTATCTC ATTAATTATT TTAAACCATA	2220
45	TCTTCAGCTG TGCCAAAGAT TTTACGTTTA ATTGCTTCGC CAGTTGGTGT GCCTGCTAGT	2280
	CCACCCAATC CAGTTTCACG TAATGATGCA GGAAGGTTAC GACCAACCTT ATCCATTGCT	2340
	TCAATAACTT CATCAACAGG GATTCTACTT TCAATACCTG CTAATGCTAA ATCTGCTGAA	2400
50	ATTAAAGCGT TACCCGAACC AATTGCATTT CTCATAACAC AAGGAATTTT AACAAGTCCG	2460
	GCTACTGGAT CACAACTAA ACCTAATAAA TTACTTATCG CTAATGCCAT AGCGTGCCCC	2520
55		

	GAACCAACTT CAGnTTGGCA GCCACCTGTT GCACCAGCTA CACTTGCAAT GTTTGCTACG	2640
	ACACGCCCAA ACAATGCTGA AGTGAATAAG AAATCAATCA TTTGCTCTTC TGTTAAATCA	2700
5	TGTGTTTTTT CTAATTTAAA AAGTGCACCG GGAATGGTAC CCGAGGAACC AGCTGTTGGC	2760
	GTtGCACAAA TAATACCCAT CGCAGCATTG ACTTCATTG TTGCAATGGC AcCtTTGcTG	2820
	CGTCAATCAT TTCATATCCA GACAAAGCAT GATGTGTTTC ATTATAATCA CGTAGTTTAG	2880
10	CAGCATCATG ACCAGTGTAG CCCGTTACAC TTTCAACCCC ATCACCTGTC GTCCCTTTGA	2940
	TTACTGCGTC TCGCATGACA TCTAAATTTT GTTTCATTG CGCTCGCACT TCATCACGTG	3000
	ATTTACCGCT TAATTCCATT TCTTCTTTAA CCATGATATC CGCAAATGAC ATATTATTTT	3060
15	CTACGGCATA ATCTATAGTC TCTCTAATTG AATCAAACAT GTTTATTCCC CCTCTAATTT	3120
	ATATAGGAAA CGTTTACGTC ACTGTATTTT TCTTTAATTG TATTTAATGT TGATTCTGAG	3180
20	ATTGCTTTAT TTAATGGTAT TACAACCAAG CATTTATCTT CATCTATCTT AATAAATTCA	3240
	TCTTTACAGT CTAATTTTAT ATCGTTGATA TCATTGATGA AATGATTTAC TTGTGCTTTA	3300
	GTCATATTTT CGTCAACAAC TAAAATTGGT AATCCATGAT TTAAATCTAC TTCTAGTCCA	3360
25	TTTATATGAA TACCTTTAAT TTAAATTGTA CCACCACCGA TTGAAATACC GATAATTTCA	3420
	ATGTAGCGAC CATCATTACG AGATGATTTG ATATAAGCAC AGTTTGGATG TTGACCAATA	3480
	CTATCGCCTT CTCTTCGAT GATATCTATT TTAATACCAT CATCAGCTGC AATTTCTAAT	3540
30	GAAGATTTAA TTCGGTTATC AAATGTTGAA TATCCCATG CTCCACCCAC AATAGCGACA	3600
	TCTGTACCAT GTCCTTGGTG TGTTTGAGCA AATGATTCAT AATAATGTAT TTCAATATTT	3660
	TTATATCTCC CAATATTGCG CGTGCTGAAT TCCCCTTTAC TGCACCAGCC GTATGAGAAC	3720
35	TTGAAGGGCC CAT	3733

(2) INFORMATION FOR SEQ ID NO: 539:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 525 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 539:

	TGGCTGTCTT CTCTATGAGT GTAGTAAGTA AGTTAACGGA TTTAACGCCA AGGCAAATAC	60
50	GTTACTATGA AACACATGAA CTCATCAAAC CTGAAAGAAC AGAAGGTCAA AAACGTCTGT	120
	TCTCACTCAA TGATTTGGAA AGATTACTAG AAATTAAATC ATTATTAGAA AAAGGATTTA	180

AAGAGATAAG AAAAAAGATG ATTGTAGATG CCACGCAAAA GCCTATTGGA GARACTTTGC 300
 CAATAAATCG TGGTGATTTA TCCCGATTTA TTAAATAAAA TTTGGAGGAT TTAAAAATGC 360
 5 CAAAACGTAC TTTCACTAAA GACGACATTC GTAAATTTGC AGAaGAGGAA AaTGTAaGaT 420
 ATTTAAGATT ACAATTCACCT GATATTTTAG GAACAATTAA AAATGTTGAA GTGCCTGTAA 480
 GCCAATTAGA AAAAGTACTT GATAACGAAA TGATGTTTGA CGGTA 525
 10

(2) INFORMATION FOR SEQ ID NO: 540:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1408 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 540:

TTGATTTGCT ACAAAGTATC TnCTCATTTC TGTATCCTGA AAAATCTTTA GTGTAATAAT 60
 GTTGTTTCAGT TTTAATATTT TCAGTCATAG TTGACTACCT CCGTATATTT TGATTTAATT 120
 25 AAGTTGTATA TTTTGATGAA CACTTATTGT TACTTGTGG CGCAAGTAGC AGTTTTTTTCA 180
 TTCTTCATAA AAGTATTCCT TATAGAATAT GAATGTTGCG ATACTTGCGA ATCCTGCAAT 240
 TGaCCATGCT GtAGTGAAGT ATAGAAACGG CATAAGTACA ATCGCTAAGA CTGTGAAGCA 300
 30 TAGTACTGCT ACTAGGTAGC TTTTATAAAT GTTACTCATT TTCTTTTTTC AACTCCTCCA 360
 TTATTCTCTG GTCTGATAAG TCGTGATAAG GGAATTTTTT CcAGCTAAT TGGACTGGTA 420
 TTCTGCCTCG TATCGCAATG TATCCTTCAT CTTCAAGCTC TTTATTTCAGT TCTCTTATTA 480
 35 TTTGTCCTGC TTTGGATTTA GAAACAGATA AAATTACCgC AAGTTCTTTA GCTTGCAAAC 540
 TATTTTTTCAT CATATCTTTT CCTCCTTTAA AATAACTGTT GATTCTCTGG GTTATCTGCT 600
 TCGTAATTAT CTGCAATAAT ACTTTTAGCG AAAAAGTCCA AACTGACCTT ATATAGGTTG 660
 TTCATAGATT TCTTTACGTT AACCCCTTCC TCAAGTACAT AAGGCACCCT AAAATCATT 720
 ATAAACAGTC CGTTTTTCGTC TAAAGTAACG GTTGGTAATT CAGGTTTGTT CCGTCTATAA 780
 45 ACTTCTCCTA GTGTAGGTTT TTGCTTTTCA GCTTGTTTAG TGAAGTCGGA AAATGCCTTA 840
 AGTAGTTTTA TTCCTGAATC AGGATCACTG TGTCGCTCAA TCGTTTCTGC TGTAGACTCT 900
 TTACTAAAAT CATTTCTATT GATTACAGGC TTTCTCGTAT TTCGTTCAAT CTTCCAAACC 960
 50 TTCCACGTCA CAACTGCCAT TGTGGTGAGG AGGGTTGTTT TGTATAGTGC GTTCATTTGT 1020
 AATTCCTCCT ATTAAGTTGT TTGTTCAATT GTGTGTGTTA TTCTTCTTCG TCTAAATCAA 1080

55

CGACTTCGTA AGTGTGCTCA ATCTCGCCTG cATATGTCAC AGTAAGAGTA TCTTTGTGTG 1200
 TGTATGTTTG ACTTTTGTTY TCTtTAACTG CATAAAGTGT TAATACTATA TTGTTTAGCT 1260
 5 TTYCTTTTTG TTCTGGTGTC ATTACGCTC CCCCTAmATT AGCyTCATAA CCGAATTCAG 1320
 TCATGATTTC ATGTATTTTC AATCTGCCTT TTTGTGTCCA TCTAGTTTGT AAAACTGTGT 1380
 10 CTTCTCTGCC ATCAGAACGC ACAATTGT 1408

(2) INFORMATION FOR SEQ ID NO: 541:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 432 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 541:

GTTTCGCAAT GATTTTTATC CGATGCTATG AGCATTatCA AATACAAAT GCTCTCTTAA 60
 AAGCAGTTAT TGA CTGAAAA TCTACTTCTA AGAGAGCACT TTATTTAATT ACTTAAGAAA 120
 25 TCTTGAAATT TCAATATACG ATGTTTATGA TAAGTCGCTT ATTTTCATCTT TAGGCTTGTT 180
 ATTAGTAAGT AGTTTAATAC CACTGATTAA CCATAAGCA AATGTAATTA TGTTACCACT 240
 TATTACAGCT CCAATAATCA ACAATATACC ACTCATTTTT TTGTTTTT TAGTGCTTTAAA 300
 30 CATACGATT GCACCTAAAA TAATTGAAAT GATTCCAAAT ATGAATAGGG ATAAGAATAA 360
 TACAGTGAAA ATTGCTGCTG CTGTTTCTGc ATCAACTGGG nCAACCTCAC CATTAACTGT 420
 35 TGTTGGACAC AT 432

(2) INFORMATION FOR SEQ ID NO: 542:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2426 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 542:

ATAATCATGA AGTnGCTAAA nCGCCAAATA ATGATGGTTC TGGACATGTT GTGTTAAATA 60
 AATTCCTTTC AAATGAAGAG AATCAAAGCC ATAGTAATCG ACTCACTGAT AAATTACaG 120
 50 GAAGCGATAA AATTAATCAT GCTATGATTG AAAAATTAGC TAAAAGTAAT GCCTCAACGC 180
 AACATTACAC ATATCATAAA CTGAATACGT TACAATCTTT AGATCAACGT ATTGCAAATA 240

GTATAAAAAG TCAACGAAAT ATTATTTTGG AAGAACTTGC ACGTACTGAT GATAAAAAGT 360
 ATGCTACACA AAGCATTTTA GAAAGTATAT TTAATAAAGA CGAGGCAGTT AAAATTCTAA 420
 5 AAGATATACG TGTTGATGGT AAAACAGATC AACAAATTGC AGATCAAATT ACTCGTCATA 480
 TTGATCAATT ATCTCTGACA ACGAGTGATG ATTTATTAAC GTCATTGATT GATCAATCAC 540
 AAGATAAGTC GCTATTGATT TCTCAAATTT TACAAACGAA ATTAGGAAAA GCTGAAGCAG 600
 10 ATAAATTGGC TAAAGATTGG ACGAATAAAG GATTATCAAA TCGCCAAATC GTTGACCAAT 660
 TGAAGAAACA TTTTGCATCA ACTGGCGACA CGTCTTCAGA TGATATATTA AAAGCAATTT 720
 TGAATAATGC CAAAGATAAA AAACAAGCAA TTGAAACGAT TTTAGCAACA CGTATAGAAA 780
 15 GACAAAAGGC AAAATTACTG GCAGATTTAA TTAATAAAT AGAACAGAT CAAAATAAAA 840
 TTTTAAATTT AGTTAAATCG GCATTGAATG GTAAAGCGGA TGATTTATTG AATTTACAAA 900
 20 AGAGACTCAA TCAAACGAAA AAAGATATAG ATTATATTTT ATCACCAATA GTAAATCGTC 960
 CAAGTTTACT AGATCGATTG AATAAAAATG GGAAAACGAC AGATTTAAAT AAGTTAGCAA 1020
 ATTTAATGAA TCAAGGATCA GATTTATTAG ACAGTATTCC AGATATACCC ACACCAAAGC 1080
 25 CAGAAAaCGt TAACACTTGG TAAAGGTAAT GGATTGTTAA GTGGATTATT AAATGCTGAT 1140
 GGTAATGTAT CTTTGCCTAA AGCGGGGGAA ACGATAAAG AACATTGGTT GCCGATATCT 1200
 GTAATTGTTG GTGCAATGGG TGTACTAATG ATTTGGTTAT CACGACGCAA TAAGTTGAAA 1260
 30 AATAAAGCAT AATTATATTG GGGGAAGAGC ATCTATATAT TTTTTTAACT ATATAAGACG 1320
 TCTTATTTCC CCTTAATTTA TTGTGAAGTA TATGCAAAT GCAATGAATA GATTGTCCAT 1380
 CATTTTAACG TTATAATGAA TTTAACGACT TAGAACTACA CAAGTAAAGG AGAATGAAGA 1440
 35 TGTCTCGAAA AACGGCGCTA TTAGTTTTGG ATATGCAAGA AGGTATAGCG AGTAGTGAC 1500
 CTAGAATAAA AAATATTATT AAAGCGAATC AGAGAGCAAT TGAAGCAGCA AGACAACATC 1560
 GAATACCACT CATTTTCATA CGTTTAGTGT TAGATAAGCA TTTTAATGAT GTCTCCTCGA 1620
 GTAATAAAGT GTTTTCAACA ATTAAAGCTC AAGGATATGC GATTACTGAA GCAGATGCAT 1680
 CTACACGAAT ACTTGAAGAT TTAGCACCAC TAGAAGATGA GCCGATTATT TCTAAGCGAC 1740
 45 GCTTTAGCGC ATTTACAGGT AGTTACTTGG AAGTTTATTT ACGTGCAAAT GATATTAATC 1800
 ATTTAGTATT AACGGGTGTC TCTACAAGTG GAGCTGTATT GAGCACGGCA TTAGAAAGTG 1860
 TAGATAAAGA CTATTATATT ACTGTTTTAG AAGATGCTGT TGGTGATAGA TCAGATGATA 1920
 50 AACATGACTT TATTATTGAA CAAATTTTAT CACGCTCATG TGACATTGAA TCCGTAGAGT 1980
 CATGGAAAAG TAGTTTATAG TTAATATAAC GTCAATTAAA GCTCGGCAGT AATGTTTGAG 2040

55

GAGGAACATT TGAACATAAA ATAATATATT TATATAAAAC GACCgAGGCG TTCGAACTGA 2160
 ATGtCCTCGG GTTTAATTGA ATAGAAATCG GACTTATGAA CGAAATATGT TTAAGTCGAA 2220
 5 CTCCTTGTTT ATACTTATAA ATTTTACGGG TTTAATATAA TACTTATTTA CCTGTAATAT 2280
 ATGCATAATT nCTTCAGTCG GTCAGCCTGT CGTTGCATAG TTCCTATGCA GCAAATGCAT 2340
 ATCCTAATCC TTTAACATTG GCATThCTGC AAATGAACGC ATAGAATCCA TTTACTGTIA 2400
 10 ACTTTTTThCA ACAAATGTCT nACATG 2426

(2) INFORMATION FOR SEQ ID NO: 543:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1874 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 543:

GAGTTGGGGA ATGTGCTCAA AATATGCGGA CTTTATGCAT TyCGGAATTG sCCaATTGCA 60
 25 GCTTTAAGCT ATGGTCAAAA AAAGAGGGTC ACTATAGCAT CTGTTTTAGT CTTAAATCCG 120
 GAAATAATCA TATTGGATGA ACCGACTGCT GGTCAAGATT TCTATCATT TAATGAGATA 180
 ATGTCATTTT TAATTGAACT AAACAGACAG GGGAAAGACGA TTATTATGAT TACGCATGAT 240
 30 ATGCATTTAT TGTCTGAGTA TAGTTCAAGA ACAGTTGTAT TATCAAAAGG TCAAGTCGTT 300
 GCTGATACCA CGCCAgTATT GGTTTTAAAT GATAAAAAAA TCTGTGAGAT TGCATCATTG 360
 AGACAAACAT CGCTATTGTA AATGGCCGAA TATATAGGGA TTAGCGAGCC ACAGAAATTA 420
 35 GTACAATTAT TTATTAACCA TGATAGGAAG GTGAGACGcC AATGAATCAA TATAATACTA 480
 TAGGTTTTCA CCCGGGAAAT AGTCGTATTc ATCAATTAAA TCGGACTGTT AAACTTTTAT 540
 40 TCTTATTAGT TGTTTCTATT TCTGCAATGG TGACTTATGA CACAAGATAT TTAATTTTAA 600
 TTAGTGCTTC ATCTATTTTA TTGGTCAAAT ATGCTCATAT TGAATGGAAA CAAGTTCGCT 660
 TTGTTGTTAA ATTCATTCTG TTTTTCACAA TAtTAAATAT TATTGCCGTG TACATATTTG 720
 45 ACCCTGAATA TGGTGTAAG ATTTATAATC AGCGTACAGA GTTAGTCAAT GGTATTGGTC 780
 GATTACGCT AACATCACAG GAATTATTCT ATCTTTTTAA TCTAATATTA AAATATATTA 840
 GTACAGTTCC TTTAGCGTTA ATATTTTTAT TCACAACGAA TCCGAGTCAT TTTGCTGCAA 900
 50 GTTTAAATCA GCTAGGTGTG AATTATAAAA TCAGTTACGC AGTCTCACTA GCATTAAGGT 960
 ATATTCCAGA TATTCAAGAA ACATATTTtA ATATTTTACA AGCGCAACAA GCAAGAGGAT 1020

55

TACCTTTAAT ATTTTCTAGT ATCGAAAGAA TTGACACTAT TAGTACTGCT ATGGAGTTAA 1140
 GACAATTCGG ACAGTATAAA AGGAGAACCT GGTACGTCAA AAAACAATTA AAAAAAGATG 1200
 5 ATTATGTTGT TTTGTGTTTG ACGTTAATAC TTCTGATGTT AGTAGTTACA TTATTCTTTT 1260
 TAAATAATAG TCGATATTTT AACCCGTGGC ATTAGTATTC ATATAAATAG TCTTTAAATA 1320
 GAAATAGGAG GGAGACATTT AATGATAAAT ACTGAAAGAT TAAATTTAAT GATTCCAAGT 1380
 10 TCCTCGCATT TAATTGAACT TTATAATATT TGTAGTCATC CACAAGCAAA TATATACACT 1440
 CCCAAAGGTT TACATAATTC CAAATTAGAC ACACAACGGT GGATTGAAAA ATGGCGAAAC 1500
 15 CATTGGCAAC AATATCAATT TGGTTACTTT GTATTGGTAA AAAAAATAGA TTGTAGTGTT 1560
 ATTGGTATTT GTGGATATGA ATATCGACAA TTAAAGCAAG AaACAGTATT AAATTTATTT 1620
 TATAAATTAC ATCCAAGTTT TGAAGGACAA GGTACGCAT GTGAGGCTAT TACAGCAATC 1680
 20 ACAAATTTTG TGAATTATAT CGATCAAGAA ACAGTAAAAG TTATCAGGAC AAATAAGTGT 1740
 AACCAACGTT CAATAAATTT AGCAGAAAGG CTTAAATTCA AGCGAGACGA TACTATGGAC 1800
 GACATTATCA ATCAAGGAGA TATTGTGTTT TAAaATAAA ATACTATGAC ATTATCTAAA 1860
 25 AAATAAAATT AAAA 1874

(2) INFORMATION FOR SEQ ID NO: 544:

(i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 5280 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 544:

TCAACATTC TAACACCAAT GTGAAAATGA TCTATGTGAT TTGCAATGGC TTGATTTGTA 60
 40 ATATGTGTGC CTAAATGACC TGTAGCACCT GTTAACATAA TATTCATTCA CTTCATCTCC 120
 TAATCTTTAT ATACATAACA TAATACTTAT TTGATGGTTT TCAAAACATT TGATTTTATA 180
 AAAAATTCTA ATCTGTATTT ATTGTCGACG TGTATAGTAA ATACGTAAAT ATTATTAATG 240
 45 TTGAAAATGC CGTAATGACG CGTTTTAGTT GATGTGTATC ACTAATATCA TTGAAAATTT 300
 TAATCaGTA CTACGACAAT ATGATGTCTG TTTTGTGTCT GAAAGTTTTA CAGTTTTTAA 360
 50 AATAAAAATG GTATAAAGTG TGATTTGTAT AAAAAAGAGT CTCGACGGAT AAGAATTGAT 420
 TAATAACAGT TAGCATTTTA TTAATTACCT TAACAATGAT TCAAGTTTAG TTAAATGAGG 480
 TTTAATTTGA AAGGGGATAG CGCCTCAATA TAATGTAGGT AGATTGTTCA TATTACGTAA 540

55

EP 0 786 519 A2

	AAGCAAGTTT	AGCATTGGGA	ATGTTAGCAA	CAGGTGTAAT	TACaTCGAAT	GTACAATCAG	660
	TACAAGCGAA	aGCAGrAGTT	AAaCAACAAa	GTGAaTCAGA	GTAAAAACAC	TATTATAATA	720
5	AmCCAATTTT	AGAGCGTAAA	AATGTGACTG	GATTTAAATA	TACTGATGAG	GGTAAACACT	780
	ATTTAGAAGT	CACAGTAGGG	CAACAGCATT	CTCGAATCAC	TTTACTTGGA	TCTGATAAAG	840
	ATAAATTTAA	AGACGGAGAA	AACTCAAATA	TAGATGTGTT	TATCCTTAGA	GAAGGTGACA	900
10	GTAGACAAGC	AACAAATTAC	TCAATTGGTG	GCGTTACAAA	ATCAAATAGT	GTGCAGTATA	960
	TTGATTATAT	CAATACGCCA	ATTTTAGAAA	TCAAGAAAGA	TAATGAAGAT	GTACTTAAAG	1020
15	ATTTTITACTA	CATTTCAAAA	GAAGACATCT	CATTAAAGA	ACTTGATTAT	AGATTAAGAG	1080
	AACGTGCGAT	TAAACAACAC	GGCTTGTTAT	CAAATGGTCT	TAAACAAGGT	CAAATTACAA	1140
	TTACAATGAA	TGATGGCACA	ACACATACAA	TCGATTTAAG	TCAAAAACCT	GAAAAAGAAC	1200
20	GTATGGGTGA	GTCAATCGAC	GGCACTAAGA	TTAATAAAAT	TCTAGTAGAA	ATGAAATAAT	1260
	ACTTTCTAAC	AACAAAGCGC	TATGTTGAAT	AGTGCTTGTT	ATGGAAATAT	ATGGAAGTTA	1320
	AGCGACGTAC	TGTTGCTTAG	CTTCTTTTTT	TGAGGGGAAA	AGTTACAAA	CTCACACAAA	1380
25	CAGTCGCACC	ACGCATTATC	TTTTGCTTAA	ATAGCTTAAT	CATATTTTAT	GAATAGTTAA	1440
	AAACAGGTTA	ATGTGAATAT	CCGAATACAG	CTCCTATAAT	ATGGGTGTAT	GATTCAAATT	1500
30	ACGTAATAAA	ACAATCTAAT	TATAATAGAT	TGGAGCATA	AACTATGAAA	ATGAAAAATA	1560
	TTGCAAAAAT	AAGTTTGTTA	TTAGGAATAT	TAGCAACAGG	TGTAAACACT	ACAACGGAAA	1620
	AACCAGTTCA	TGCCGAAAAG	AAACCTATTG	TAATAAGTga	AAATAGCAAA	AAATTAAAAG	1680
35	CTTATTATAA	TCAACCTAGT	ATTGAATATA	AAAATGTGAC	AGGTTATATC	AGTTTCATTC	1740
	AACCAAGTAT	TAAATTTATG	AATATCATAG	ATGGTAATTC	TGTTAATAAT	ATTGCTTTAA	1800
	TTGGCAAAGA	TAAGCAACAT	TATCATACGG	GTGTACATCG	TAATCTTAAT	ATATTTTACG	1860
40	TTAATGAGGA	TAAGAGATTT	GAAGGTGCAA	AGTACTCTAT	TGGGGGTATC	ACGAGTGCAA	1920
	ACGrTaAAGC	TGTCGACCTA	ATAGCAGAAG	CAAGAGTTAT	TAAAGAAGAT	CATACTGGTG	1980
45	AATATGATTA	TGACTTTTTTC	CCATTTAAAA	TAGATAAAGA	AGCGATGTCA	TTGAAAGAGA	2040
	TTGATTTTAA	ATTAAGAAAA	TACCTTATTG	ATAATTATGG	TCTTTACGGT	GAAATGAGTA	2100
	CAGGAAAAAT	TACAGTCAAA	AAGAAATACT	ATGGAAAGTA	TACATTTGAA	TTGGATAAAA	2160
50	AGTTACAAGA	AGACCGTATG	TCCGATGTTA	TCAATGTCAC	AGATATTGAT	AGAATTGAAA	2220
	TCAAAGTTAT	AAAAGCATAA	CACATATACT	TGATGACGAA	ATAAGTTGAA	ATTGAAATAG	2280
55	AGAGGTTAAG	TGACGATCAA	ACGTTGCTTA	ACTTCTTTTT	AATGCTTAAA	AATTATTTCA	2340

	TTAATAATAC TTCAATAATT GTTAAAAGGG GTTTAATGTG ATTATCTTAG AACGCCATCT	2460
	ATAATGATGT TGTATGATTC AAATTACGTA AAAAGACAAT CGAATATAAT ATAGATTGGA	2520
5	GCATACAATT ATGAAAATGA GAACAATTGC TAAAACCACT TTAGCACTAG GGCTTTTAAC	2580
	AACAGGCGCA ATTACAGTAA CGACGCAATC GGTCAAAGCA GAAAAAATAC AATCAACTAA	2640
10	AGTTGACAAA GTACCAACGC TTAAAGCAGA GCGATTAGCA ATGATAAACA TAACAGCAGG	2700
	TGCAAATTC A GCGACAACAC AAGCAGCTAA CACAAGACAA GAACGCACGC CTAAACTCGA	2760
	AAAGGCACCA AATACTAATG AGGAAAAAAC CTCAGCTTCC AAAATAGAAA AAATATCACA	2820
15	ACCTAAACAA GAAGAGCAGA AAACGCTTAA TATATCAGCA ACGCCAGCGC CTAAACAAGA	2880
	ACAATCACAA ACGACAACCG AATCCACAAC GCCGAAAACCT AAAGTGACAA CACCTCCATC	2940
	AACAAACACG CCACAACCAA TGCAATCTAC TAAATCAGAC ACACCACAAT CTCCAACCAT	3000
20	AAAACAAGCA CAAACAGATA TGA CTCTCTAA ATATGAAGAT TTAAGAGCGT ATTATACAAA	3060
	ACCGAGTTT GAATTTGAAA AGCAGTTTGG ATTTATGCTC AAACCATGGA CGACGGTTAG	3120
	GTTTATGAAT GTTATTCCAA ATAGGTTTCAT CTATAAAATA GCTTTAGTTG GAAAAGATGA	3180
25	GAAAAAATAT AAAGATGGAC CTTACGATAA TATCGATGTA TTTATCGTTT TAGAAGACAA	3240
	TAAATATCAA TTGAAAAAAT ATTCTGTCCG TGGCATCAGC AAGACTAATA GTAAAAAGT	3300
30	TAATCACAAA GTAGAATTAA GCATTACTAA AAAAGATAAT CAAGGTATGA TTTCACGCGA	3360
	TGTTTCAGAA TACATGATTA CTAAGGAAGA GATTTCTTGG AAAGAGCTTG ATTTTAAATT	3420
	GAGAAAACAA CTTATTGAAA AACATAATCT TTACGGTAAC ATGGGTTTCAG GAACAATCGT	3480
35	TATTAATATG AAAAACGGTG GGAAATATAC GTTTGAATTA CACAAAAAAC TGCAAGAGCA	3540
	TCGTATGGCA GACGTCATAG ATGGCACTAA TATTGATAAC ATTGAAGTGA ATATAAAATA	3600
	ATCATGACAT TCTCTAAATA GAAGCTGTCA TCGGAAAAAC AAGAAGTTAA GTGACAACGG	3660
40	TTTACATGTT GCTTAGCTTC TTTTATTATG CGTAATGATG TAAAAAGACG AATATTCATT	3720
	TGTTTGTAAG AGTGGCATT CTATGTCTTA AAAGTGACGA AACTTCAAAT GTGCCAAGTG	3780
45	TTGAATCACA TCAAAATCAT TTTTATTTAA CGAACATTAT GGATTTCTTA ATTTACTTAA	3840
	CGATGATTCA AATATAGTTA AACAAGGTTT AATGTGAATG GAGCAATACG CCATCTATAA	3900
	TAAAGCTGTA TGATTCAATG AATGTAATCG AACAAATCTA ATAATTACGA ATGGAGCATA	3960
50	CAACTATGAA AATAACAACG ATTGCTAAAA CAAGTTTAGC ACTAGGCCTT TTAACAACAG	4020
	GTGTAATCAC AACGACAACG CAAGCAGCAA ACGCGACAAC ACTATCTTCC ACTAAAGTGG	4080
55	AAGCACCACA ATCAACACCG CCCTCAACTA AAATAGAAGC ACCGCAATCA AAACCAAACG	4140

	CGCCTTCAAC TAAAGTGACA ACACCTCCAT CAACAAACAC GCCACAACCA ATGCAATCTA	4260
	CTAAATCAGA CACACCACAA TCGCCAACCA CAAAACAAGT ACCAACAGAA ATAAATCCTA	4320
5	AATTTAAAGA TTAAAGAGCG TATTATACGA AACCAAGTTT AGAATTTAAA AATGAGATTG	4380
	GTATTATTTT AAAAAAATGG ACGACAATAA GATTTATGAA TGTTGTCCCA GATTATTTCA	4440
	TATATAAAT TGCTTTAGTT GGTAAAGATG ATAAAAAATA TGGTGAAGGA GTACATAGGA	4500
10	ATGTCGATGT ATTTGTCGTT TTAGAAGAAA ATAATTACAA TCTGGAAAAA TATTCTGTCTG	4560
	GTGGTATCAC AAAGAGTAAT AGTAAAAAAG TTGATCACAA AGCAGGAGTA AGAATTACTA	4620
15	AGGAAGATAA TAAAGGTACA ATCTCTCATG ATGTTTCAGA ATTCAAGATT ACTAAAGAAC	4680
	AGATTTCCCTT GAAAGAAGTT GATTTTAAAT TGAGAAAAA AACTTATTGAA AAAAATAATC	4740
	TGTACGGTAA CGTTGGTTCA GGTAAAATTG TTATTAAAT GAAAAACGGT GGAAAGTACA	4800
20	CGTTTGAATT GCACAAAAA TTACAAGAAA ATCGCATGGC AGATGTCATA GATGGCACTA	4860
	ATATTGATAA CATTGAAGTG AATATAAAT AATCATGACA TTCTCTAAAT AGAAGCTGTC	4920
	ATCGGAAAAA CAAGAAGTTA AGTGACAACG GCCTACATGT TGCTTAGCTT CTTTGTGTAT	4980
25	GTTTCGATGAT TTGAGAACCC GAATTTTCGA TGGGTCCAAA TATGACGTGG AAGAGACCTG	5040
	AATTTATCTG TAAATCCCTA TCTATCGGGT GTGAAGCACA ACGGGATCAG TTTTATTTAA	5100
30	CGAACATTAT AGATTCCTTA ATTTACTTAA TAATGATTCA ATGATTATTA AACATGGTTT	5160
	AATGTGAAAG GTCAAATACG CTAAGTATAA TAAAGCTGTA TGATTCAATA GACGTAAGCG	5220
	AACAAATCTA ATAATTACGA ATGGAGCATA CACTATGAA AATGACAGCA ATTGCGAAAG	5280

35 (2) INFORMATION FOR SEQ ID NO: 545:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 886 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 545:

45	AGTAAAATTG CCGGTATGAT GGACACAAAC GGTGACCTTG GTCAAGGTGA ATTAGCGATT	60
	AATCCACCTA AATCAGATTT GAACGAATTA CCTTGGGCTA CACGTAAAAA TAAACAGCCA	120
50	GCTTCATCCG AAAAAGGTTT AAGTGGTCAT CATGGTAATG CAGCAATGCC TCAAACCAm	180
	TTAGATTATC AAATATCTAT TGATAAGGTC GTTGAACAGG CGCAAAAAGC TGGTATTAAA	240
	AAGCCGTTTT CAATCGTATA TCCAAGTGAT AAAAATGGTA CCTTTATTGT ATCTAATACT	300

55

GATCAATATA GCGGTAAAAA GCTAGGTACG ATTAAATATG ATGACTACGG TATTATTGCT 420
 AAATGGTTTA CATGGGGCAT TCCGCTTCAC GAAGGTCATT TATTCGGCAT TTAAATAAAA 480
 5 ATCATTAAATT TATTTGTATG TATCGCTTTA TTAGTAGCCA TTGGCATGGG GTTTGTCTCT 540
 TGGATAAAGC GTACAAAAAA TACTGCAGTA AAAGTACCAC ATCGCGTAAA AAAACCAGCA 600
 TCTATATCAC TCATAATATG TTTAATTGTA TTAGGATTAT TAATGCCATT ATTTGGATTA 660
 10 TCACTTATCC TTGTATTAT AATTGAATTA ATATTATATA TTAAAGATCG TCGTGCTAAA 720
 CAATAATGCA CTAAAGTTT TGAAGTACG AAATTTACAA AATGgATTCT CGTCTCTCTA 780
 ATTACTTAAA ACGGGgTtCy AaTAATAAAT CgTACTGaTG GgAAAGTTTT TACTTTTTAT 840
 15 CTGtCCGaTt TTTTnGAAwT TGAAGATAAA AAAGCATCTA AAACGC 886

(2) INFORMATION FOR SEQ ID NO: 546:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4336 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 546:

GGCATTGTG TCCTTATATA AGGAACTGTG tTAAATACAT TACTGTTGTT AAGTTGTTTT 60
 30 TGTAATTCOA AGAGCAGAAC AGAGTAACAT CATCAGTTGT AGTAAACGAT AATCCGGTAA 120
 AACAACTAAA TGAAATAATG AAAGTCATTT AACCTGAACA TTAAATATA TTTGTTTTTC 180
 35 ATTAAGAATA ATTCAAGTAT ATTTAAATCG AGGTTAATTA TCGTATGAAA CGATGCACGT 240
 TATAATAAAA ATGTATGATT CAAATTACGT AATGAAAACA ATCCAATATA TTAAGATTGG 300
 AGCAAATAAA TATGAAATTT ACAGCATTAG CAAAAGCGAC ATTAGCTTTA GGAATTTTAA 360
 40 CAACAGGAAC TTTAACAACA GAAGTTCATT CAGGTCATGC AAAACAAAAT CAAAAGTCAG 420
 TAAATAAACA TGACAAGGAA GCATTATACC GATACTACAC TGGAAAGACT ATGGAAATGA 480
 AAAATATTAG TGCTTTGAAA CATGGTAAAA ACAACTTACG TTTTAAGTTT AGAGGTATTA 540
 45 AGATTCAAGT TTTACTGCCT GGAAATGATA AAAGTAAATT TCAACAGCGT AGTTATGAGG 600
 GGTTAGATGT TTTCTTTGTT CAAGAAAAAA GAGATAAGCA CGATATATTT TATACTGTTG 660
 GTGGTGTAAT ACAGAATAAT AAAACATCTG GAGTTGTCAG TGCACCAATA TTAAATATTT 720
 50 CAAAAGAAAA GGGTGAAGAT GCTTTTGTGA AAGGTTACCC TTATTACATT AAAAAAGAAA 780
 AAATAACACT AAAAGAACTG GATTATAAGT TGAGAAAGCA TCTAATTGAA AAATACGGAC 840

55

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATAACCTTGA TTTAAGATCT AAATTAAAAT TTAAATATAT GGGGGAAGTC ATAGAAAGCA 960
 AACAAATTAA AGATATTGAA GTTAACTTAA AGTAAATCAT TACGAATAAT TAAAAGTAAT 1020
 TGAAGCGGCT TAACGGTGAA ATGTAAATTG GTGCGCATAG CTTATACAAA AAGGATGCAT 1080
 CAATCGATAT CGTCGTTAAG CCGTTTTGGT TTGTGTGTCA TGAATCCTAT CCCAATCTCC 1140
 ATAAAGGTAA AATTTCCACC ACCAACATCA AAATTCTCCA CATCGCAACA TAACCAAATG 1200
 TTATAATAAA TCTATTACAC AAAGAGATAA ATTACTTATT CAAAGGCGGA GGAATCACAT 1260
 GTCTATTACT GAAAAACAAC GTCAGCAACA AGCTGAATTA CATAAAAAAT TATGGTCGAT 1320
 TGCGAATGAT TTAAGAGGGA ATATGGATGC GAGTGAATTC CGTAATTACA TTTTAGGCTT 1380
 GATTTTCTAT CGCTTCTTAT CTGAAAAAGC GGAACAAGAA TATGCAGATG CCTTGTGAGG 1440
 TGAAGACATC ACGTATCAAG AAGCATGGGC AGACGAAGAA TACCGTGAAG ACTTAAAAGC 1500
 AGAATTAATT GACCAAGTCG GTTACTTCAT TGAGCCAGAA GATTTATTCA GTGCGATGAT 1560
 TCGTGAAATT GAAACGCAAG ATTTGATAT CGAACACCTG GCGACGGCAA TTCGTAAAGT 1620
 TGAAACATCA ACATTAGGTG AAGAAAGTGA AAATGACTTT ATCGGTCTGT TCAGCGATAT 1680
 GGATTTGAGT TCAACGCGAC TAGGTAACAA TGTCAAAGAA CGTACTGCTT TAATCTCTAA 1740
 AGTCATGGTT AATCTTGACG ACTTACCATT CGTTCACAGT GACATGGAAA TTGATATGTT 1800
 AGGTGATGCA TATGAATTCC TAATTGGGCG CTTTGCGGCG ACAGCGGGTA AAAAAGCAGG 1860
 CGAGTTCTAT ACACCACAAC AAGTATCTAA GATACTGGCG AAGATTGTCA CAGACGGTAA 1920
 AGATAAATTA CGTCACGTGT ATGACCCAAC ATGTGGTTCA GGTTCACGTG TGTACGTGT 1980
 TGGTAAAGAA ACACAAGTGT ATCGTTATTT CGGTCAAGAA CGTAACAATA CTACATACAA 2040
 CTTAGCACGC ATGAATATGT TATTACATGA TGTGCGTTAT GAGAACTTCG ATATCCGTAA 2100
 TGATGACACA TTGGAAAACC CAGCCTTTTT AGGCAATACA TTTGATGCGG TTATTGCGAA 2160
 CCCACCGTAT AGTGCGAAAT GGA CTG CAGA TTCAAAGTTT GAAAATGACG AACGATTGAG 2220
 TGGTTACGGC AAAGTTGCGC CTAAGTCTAA AGCAGACTTT GCCTTTATTC AACACATGGT 2280
 ACATTACCTA GACGATGAAG GTACCATGGC CGTTGTACTC CCACATGGTG TATTATTCCG 2340
 AGGTGCTGCA GAAGGTGTCA TTCGTCGTTA TTTAATTGAA GAAAAGAAGT ACTTAGAAGC 2400
 TGTGATTGGT TTGCCAGCGA ATATTTTCTA TGGGACAAGT ATTCCAACAT GTATTTTAGT 2460
 ATTTAAAAAA TGTCGCCAAC AAGACGACAA CGTACTATTT ATCGATGCAT CCAATGATTT 2520
 TGAAAAAGGA AAAAATCAAA ATCATTTAAG CGATGCCCAA GTCGAACGTA TTATAGACAC 2580
 ATATAAGCGT AAGGAAACAA TTGATAAATA TAGCTACAGC GCGACACTAC AAGAGATTGC 2640

	GATTGATTTA GATCAAGTCC AACAGATTT GAAAAATATC GATAAAGAAA TCGCAGAAAT	2760
	TGAGCAAGAA ATCAATGCAT ACCTGAAAGA ACTTGGGGTG TTGAAAGATG AGTAATACAC	2820
5	AAAAGAAAAA TGTGCCAGAA TTGAGGTTCC CAGGGTTTGA AGGCGAATGG GAAGAGAAGC	2880
	AGTTAGGGGA TCTTACAGAT AGAGTAATTA GGAAAAATAA AAACCTTAGAA TCGAAAAAGC	2940
	CTTTAACAAT ATCCGGACAG TTAGGTTTAA TTGATCAAAC AGAATATTTT AGTAAATCAG	3000
10	TTTCGTCGAA AAATCTAGAA AATTATACAC TAATAAGAA TGGAGAATTC GCGTATAACA	3060
	AAAGTTATTC TAATGGATAC CCATTAGGGG CTATTAAAG ATTAAGTAGA TATGATAGTG	3120
	GTGTATTGTC CTCTTTGTAT ATTTGTTTTT CTATTAAAG TGAAATGTCT AAAGACTTCA	3180
15	TGGAAGCATA TTTTGATTCT ACACACTGGT ATAGAGAAGT TTCTGGAATT GCAGTTGAGG	3240
	GTGCAAGAAA TCACGGATTA TTAAATGTTT CTGTGAATGA TTTTCTTACT ATTCTAATTA	3300
20	AATATCCAAG TTTAGAAGAA CAGCAAAAAA TAGGCAAGTT CTTAGCAAA CTCGACCGAC	3360
	AAATTGAATT AGAAGAACAA AAGCTTGAAT TACTTCAACA ACAGAAAAAA GGCTATATGC	3420
	AGAAAATTTT CTCACAGGAA CTGCGATTCA AAGATGAGAA TGGTGAAGAT TATCCAGATT	3480
25	GGGAAATAG CAAAATAGAA AAATATTTAA AAGAGAGAAA CGAACGTTCT GACAAAGGGC	3540
	AAATGCTTTC AGTAACTATA AATAGTGGCA TTATAAAATT TAGTGAATTG GATAGAAAAG	3600
	ATAATTCAAG TAAAGATAAA AGTAATTATA AAGTAGTTAG GAAAAATGAT ATTGCATATA	3660
30	ATTCTATGAG AATGTGGCAA GGGGCTAGTG GTAAATCAAA TTATAATGGG ATTGTTAGCC	3720
	CTGCATATAC TGTGCTTTAT CCAACACAAA ATACTAGCTC ATTATTTATT GGATATAAGT	3780
35	TTAAACACA TAGAATGATT CATAAATTTA AAATTAATTC ACAAGGATTA ACATCAGATA	3840
	CATGGAACCT AAAATATAAA CAATTAAAA ATATAAATAT AGATATACCT GTATTGGAGG	3900
	AACAAGAAAA GATAGGTGAT TTCTTTAAAA AAATGGATAT ATTGATAAGT AAACAGAAAA	3960
40	TGAAAATTGA AATATTAGAA AAAGAGAAAC AATCCTTTTT ACAAAAAATG TTCTTATAAC	4020
	TTTGATAAAT ACATAGATTG CATAAGAATA AAATTGTAT AATTTAACAT AAAAGTTGTA	4080
	AAAGTAAAGT GAATTAAAAA CGAACATTAA ATTTAGGCAC TGTGAAAGCG CAGTGTCTTT	4140
45	TTTGTGTCGA AATGTGTAC AGAATAAGTA GTTAAATAAA GATTAAAGTTG AGATAAAGTG	4200
	TTATTCGTAA ATAAAAGAGA GTAGATCGAT AGGAATTGAA TGATATTAGT TAACTATTTA	4260
50	TTAAATTACT TAATAATGAT TAATTTTTAG TTAAAGTAAG TTTAATGTGA AGCACGACCA	4320
	TTGCTCATT TAATGA	4336

(2) INFORMATION FOR SEQ ID NO: 547:

55

(A) LENGTH: 487 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 547:

10 TAAGCTATCT GAGATAATTG CTGATAACAT TAAACCGGCA ATTTTCAGGTT TAATTTCAAA 60
 GCCACGTTCT CTAAACATTT TGTATAAAAT TGTAGCTGTA CAACCAACTG GTTCAGCAGC 120
 ATAACATAAA GGACCAGCAG TTTCGAAATT TGCAATTCTG TGATGATCAA TTACATGCTT 180
 15 AATTGTAGCA GAGGCAATCG TATCAGAACT TTGTTGGaAT TCGTTATGAT CAACTAAGAT 240
 AACATCTTGA CCATCTAAAT CATCTGTTAA TAATTCCGGA GCAGGTACAT TAAATGTATC 300
 TAACGCGAAT TGAGTTTCTG CACTCACATC ACCTAAACGG TATGCTTTGG CTCCTGAATT 360
 20 ACCTCGAAGT TGTTCaAATT CTGCCaTAAT AATCGCAGAT GAAATTGCAT CAGTGkCTGG 420
 aTTCTTATGT CCGAAAATAT ATGTTTTAGC CAntTGTCAA TATCTCCCTT GTAAATTGTA 480
 25 TTCTTTA 487

(2) INFORMATION FOR SEQ ID NO: 548:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 871 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 548:

TTGGTGGTGC AGCAGTTATA GCAATTGTTT TAGCATTCAAT TGGTAAGTTC ACTGCATTAA 60
 TTTCTTCTAT ACCTACmCCA GTTATGGGAG GAGTATCTAT ATTACTTTTC GGTATTATTG 120
 40 CAGCAAGTGG CTTAAGAATG TTAGTTGAAA GCAAAGTAGA TTTTGCGAAC AATCGAAATT 180
 TAGTTATAGC TTCTGTAATT TTAGTTGTAG GTATCGGTAA TTTAGTATTT AACTTAAAAG 240
 45 AAATTGGTAT CAACCTTCAA ATTGAGGGGA TGGCATTAGC TGCACTTTCA GGAATTATTT 300
 TGAACCTAAT CTTACCTAAA GAGAAAAAAC AAAACAATTA AGATTTACAA ATTAAGGAGG 360
 GCGCTTTTAT GAATCATTTA TTATCAATGG AACATTTATC TACAGATCAA ATATACAAAC 420
 50 TTATCCAAAA GGCAAGTCAA TTAAATCTG GTGAACGTCA ACTACCAAAC TTTGAAGGGA 480
 AaTATGTCGC AAATTTATTC TTTGAAAATT CTACTCGrAC AAAATGTAGT TTTGAAATGG 540
 CAGAACTTAA GCTAGGGTTA AAAACGATTA GCTTTGAAAC ATCAACATCA TCTGTTTCAA 600

55

TCATTAGACA TCCGTTTAAAT AACTACTATG AAAAATTAGC GAATATTAAAC ATCCCAATTG 720
 CGAATGCTGG TGATGGTAGT GGACAACATC CAACACAAAG TTTACTTGAT TTAATGACGA 780
 5 TATATGAAGA ATATGGATAT TTTGAAGGCT TGAATGTATT GATTTGTGGA GAcATTaAAA 840
 ATTCACGTGT CGCACTAGTA ATTACCAaAG T 871

(2) INFORMATION FOR SEQ ID NO: 549:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 549:

20 TCAATCTAAT ATATACTTCA TGACTTCCCG CCAAAGGCGC AATCGTAGGA TTAGTCTGTT 60
 TATCGATAAG ATCTATTAAT ATTGTTTCTA CTTTAGATTC ACCTATTCCC GCAAATCTTA 120
 ATAGTTCAGA ATGTATAATT CGATTATGGT TTATAAAATG TGACAACAAT TCATTTTTCA 180
 25 CCATTGGTTG CATTTCTTTC GGTGGACCTG GTAATAAAAT AATTGTTTG TTTTCAAAAT 240
 TCACCATCAT TCCTGGAGCC ATGCCATGAT GATTTGTAA TACAGTTGAA CCTTCAATTA 300
 CTAAAGCCTG TTGTCTATTA TTAGGTGTCA TTTCTGTCC TTGTTCTCA aAAwAGCTTT 360
 30 CAATATATTG AAAGAAGGCT CATCAATAAC TAAATCTnTA 400

(2) INFORMATION FOR SEQ ID NO: 550:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1523 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 550:

ATACCTCCAG CTAGAATACC AGCGTATTTT ATAAATACT TCCTCCATT C AACTATATCT 60
 45 ATATTTAATT ATTTAAATTT CGTTGCATTT TCCAATTGAA AACTCATTTT AAAATCAAAA 120
 CTCTAAATGT CTGTGTATTA CTTAAATTA TACATATTTT GCTTATATTT TAGCATATTT 180
 TGTTTAAACC TATATTACAT TATATCAGAC GTTTTCATAC ACAAATAATA ACATACAAGC 240
 50 AAACATTTG TTTATTATTT ATATCACTTA ACTAATTAAT TTATAATTTT TTATTGTTTT 300
 TAAGTTATCA CTTAAAAATC GTTTGGCaAA TTCGTTGTGA CGCTTGTTCA TCTTCTAATG 360

TTTGTTTTAA TGCATCAATG AGTGCTGTTT GATTTTCAAC AATTGGACCT GGCAACTCTT 480
 TTTTATAATC CATGTAAAAA CCTCTAAGCT CATCGCCATA TTTATCTAAG TCATATGCAT 540
 5 AGAAAATTG CGGACGCTTT AATACACCGA AGTCGAACAT GACAGATGAG TAGTCGGTAA 600
 CTAACGCATC GCTGATTAAG TATAAATCCG AAATGTCTTC ATAAtCTGAA ACGTCTTTCA 660
 10 CAAAATCATC ATGTTTCATCA ATACGTGTCA CAACTAAATA ATGCATGCGT AAKAAAATAA 720
 CATAATCATC ATCCAGCGCT TGACGCAAAG CTTCTATATC AAAGTTAACA TTAAATTGAT 780
 ATGAACCTTC TCGAATCGCT TCATCGTCAC GCCAAGTTGG CGCGTACATA ATCACTTTTT 840
 15 TATCTAATGG AATATTTAAT CTTGTCTTAA TACCATTAAT ATATTCAGTA TCATTGCGTT 900
 TATGTGATAA TTTATCATTT CTTGGATAAC CTGTTTCCAA AATCTTATCT CGACTAACAT 960
 GAAATGCATT TTGAAATATC GATGTCGAAT ATGGATTAGG TGACACTAGA TAATCCCACC 1020
 20 GTTGGCTTTC TTTTTTAAAG CCATCTGGT AATTTTGAGT ATTTGTTTCT AGCATTTTAA 1080
 CGTTACTAAT ATCCAAACCA ATCTTTTTTA ATGGCGTGCC ATGCCATGTT TGTAAGTACG 1140
 TCGTTGCGGG TGATTTATAT AACCAATCTG GTGTACGTGT GTTAATCATC CwCGCTTTTCG 1200
 25 CTCTTGGCAT CGCTAAAAAC CATTTTCATTG AAAACTTTGT AACATATGGT ACATTGTGCT 1260
 GTTGAATAT GTGTTCATAT CCTTTTTTCA CACCCCATAT TAATTGGGCA TCGCTATGTT 1320
 CAGTTAAGTA TTCATATAAT GCTTTGGGGT TGTCGCTGTA TTGTTTACCA TGAAAGCTTT 1380
 30 CAAAATAAAT TAGATTCTTG TTTGGCAATT TTGATAGTAA TTTAAAAGTC GTATATATAC 1440
 TATGTTCTAT CAATTTTTTA ATTGTATTTT TAATCATGTC GTACCTCCGA CGTGTTTTTG 1500
 35 TAATTATATT AATATGTATG AGC 1523

(2) INFORMATION FOR SEQ ID NO: 551:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4923 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 551:

CAGTAAGAGA TTTTCTTAAT TGAAAATAAT CTTACTGCTT TTTTAAATTT AATTTGAGA 60
 50 TTCAATATTA GTTATCTCA TTGTGGCATT AATTGATAAA ATTGTTTTAA TGTATAAAT 120
 CAAAGTCTTC TTCAACAATT TCAATGTCTG CATCAGATCT ATGATATGTG AAAAAGCTAA 180
 TTCTTATGCG GTCTAAATGC TCTAAATGGT GTCGATACTC TTCGATGGCT GCAACAATTT 240

	ATGTGGACAA TAAATCTTTT TTCTGCGGAT TATCTAACTC ATAATCAACA TGTGTCACAT	360
	TATAACGTGC TTTTTTAGAA AGGCTAGCTA AAATTTGCTC GTGAAAAGCA GTTAATGAAT	420
5	CTAAATCTAA TTTGATTTGT AATAGGAAAT TGTTATTAAG TAAATATAAG TCGTTTTGAT	480
	AACGCGACAA TTTGTTTAAAT ACTTCATAAG CTTGTCTAGT CGTCTGAACT ACTTCTCTGA	540
	AAAGTATTTT CTTTCTATTC TGTTGGTGAA TATGTTTTTT TGTAATAGGA CGTTCTTCGC	600
10	TATAGTAATC ATAAATTTTC TCTAACTTTT CGACACGTTG TTTTAAATTA TGACTATCTT	660
	GTTTAATATT ATTAAACTCC GTCGTATCAT TTAATACTAA TTTAAACCAC ATAAAAATAT	720
	CTGAGGATAT ATTTAATGAA TTATAGTAA TTTTGTTC AAATTTAGGT GGTAGAAACA	780
15	CAAAGTTAAC TAGAGATGAA CTTATGACAC CAATCATTAC AAGTACAAAC CTGTAAAAGG	840
	CGGTAATATA GAAAGAACCG GTATGTTGTC CCATAATGAT TAATGCTGTT ACACTCGCCA	900
20	AAGTAGCAAC ATGTGCTAAA TTAAATTTAA ATAAAATAGC AATAAGTACT ATGACGGTAA	960
	CACCCATAAT GATAAAATTA TCACTAAAAA TTGTTACCAT TGTAACAGAT AGTATGGCAC	1020
	CTATAATGTT ACCCAATGCT TGATCAGAAA CTGTTTTTAA TGAACGATAA ATACTAGGTT	1080
25	GCATTGCACA ACAGCACTGA CACCAGCTAA GGCTTTCAGA CCAACATCAT CCGGTAGTAA	1140
	AGAAGCGATA GACATAGCTA AAATAATGGC TATACCAGTT TTAAAAATCC GAGCTCCTAG	1200
	TCTCAAAAAT AATGACGCCC CTTTAAAGTT TATTGAATAT CTAATATTCTG TATTCATTAC	1260
30	TGTTATACAC TTAGTAGTTA CAAAATTCAA GCTTATTTAT AGTTGTTAAA ATAAATCATA	1320
	CATAATACTG ATAGCGATGT AAAACTTTAG TCAGAGATTA AAATAGTATA AATTTGTAAA	1380
	ATAAAAACTC ACATAGTGAC ATATCAAGTT AAACGTTAAT AGTTAACGAT ATAAAATGAA	1440
35	TCTACTATGT GAGCATTTCG TTTATTTTAA TTCAATTAAA AATATACTTC CTTAAAAGTT	1500
	ATTTCAATTG ACTAAAAGCA TAGTCTGCAG CTTTAAAGT TTGTTAATA TCTTCTTCTG	1560
40	TATGTTTCAGT TGTTAAGAAC CAAGCTTCAA ACTTAGAAGG TGCTAAATTG ATACCTTGAT	1620
	TTAACATTAA TTTGAAAAAT TTACCGAACG CTTCCCGTC AGAATGTTCA ACTTGATCAT	1680
	AATGTGTGAC TTTTTCATCT GTAAAGTACA ATGTTAAAGA TCCATAAATA CGATTAAATTG	1740
45	TAGCTGTGAT ATTATGTTTT TCGATTAATT TAAGTAAACC TTCTTCTAGT TGTGGCCTA	1800
	AGCTGTCTAA TTTTTCATAA ACACCGTCTT GTTCTAGTAC TTCGAGTAAT GCAATACCTG	1860
	CTTTCATAGA TAACGGGTTA CCAGCCATTG TACCAGCTTG ATATGCAGGT CCTAGAGGTG	1920
50	CTACTTGTTT CATAATATCT TGACGTCCAC CATAGCTCC AATTGGTAAA CCACCGCCAA	1980
	CAATTTTACC AAATGCAGTT AAATCAGGGA TAACACCTAA TAAATCTTGA GCGGCACCGT	2040

55

	AAATTTTCATT AACCTCTTCT AAAAATCCAG GTTGAGGCAT TACCATTCCA AAGTTACCAA	2160
	CAATTGGTTC TACTAATACT GCGGCAATTT CATCACCCCA AAATTCAATT GCTTCTTTAT	2220
5	AGGCGTTAAT ATCATTGAAA GGTACAGTAA TGACTTCACG TGCGACGCTT TCTGGAACAC	2280
	CAGCTGAGTC TGGAGAACCG AGCTGAGATG GGCCGCTACC TGCTGCAACC AATACTAAAT	2340
10	CAGAATGGCC ATGATAAGAT CCAGCAAATT TTATAATTTT ATTTCTTTTA GTATATGCAC	2400
	GTGCAACACG AATTGTTGTC ATGACTGCTT CTGTTCCAGA ATTTACAAAG CGAATTTTCT	2460
	CAAGAGATGG AATTGCATCA CGTAATTTTT TGCTGAATTC AATTTCTAAT TCAGTCGGTG	2520
15	TACCAAATAA AACACCTTTA GCAGCTTGTT CTTGAATTGC TTTAGTAATA TGAGGATGTG	2580
	CATGCCCCGT AATAATTGGA CCGTATGCTT GAAGGTAATC AATAAATTTA TTGCCATCGA	2640
	CATCATATAA ATATGCACCG TGTCTTCTT TCATAACAAC AGGTGCACCG CCTCCTACAG	2700
20	CTTTATAAGA ACGAGAAGGG GAATTGACAC CGCCTAGAAT ATATTCGTTT GAAAGTTGTT	2760
	GTAAACGTTT ACTTTCACTA AAATTCATTT ATATCAACCT CTTTTAATTT AATATTTTCA	2820
25	TCTAATATCG TATCATAAAA TTATTATAAT GAAGAAAAAG GTGATTATAT GTTGCAAAAA	2880
	GGAGAACAAT TTCCAATATT TAAATTAGAA AATCAAGACG GAACTGTCAT TACAAATGAT	2940
	ACATTAAAAG GTAAAAAGGC GATTATATAT TTTTATCCTA GAGATAATAC ACCTACTTGT	3000
30	ACCACAGAAG CTTGTGACTT TAGAGACAAT TTAGAAATGT TCAATGATTT AGATGTTGCA	3060
	GTATATGGTA TAAGCGGTGA TTCAAAGAAA AAACACCAAA ATTTTATTGA GAAACACGGA	3120
	TTGAATTTTCG ATTTATTAGT AGATGAAGAT TTAAATTAG CTAAAGAAAT GCGGTATATC	3180
35	AGTTAAAAAA ATCATTTGGC AAAGAAAGTA TGGGCATTGT AAGAACGACT TTTATAATAG	3240
	ATGAACAAGG TAAAGTATTA GATGTTATCG AGAAGGTTAA GGTAAAAACA CAAATAGAAG	3300
	AACTTAAAAA CATTTTGGGG TGACATATAT GAAAGTTGTT GGGTTAAATC GSTATGCGTGA	3360
40	AGTTGAAACT GAATTACAAC AACGCTTTTC AGATTTAGAT TTTAAATTTT ATAAAAAGC	3420
	ATCAGAAATA CCTGAGAGCG ACTTGGCTGA TTTAGATATA TTAGTTGGTT ATGATGGCGG	3480
45	TATCAATGAG GCATTTTAC GACGTTGCCC GAATTTAAAA TGGATTGCAT GGGTTGCAAC	3540
	GGGTGTAAAT ACATTGCCGT TAGATTATAT TGCAGATCAC GGCATACTTT TAACTAATGG	3600
	AAAAGGTGTT CAAGCTAAAC AATTATCTGA ATACATTTTA GCTTTCATTT TAGATGATTA	3660
50	TAAAAAGATG AACTATCAT ATGATAACCA ACGACAACAT ATATATGATT CGAAAAATAC	3720
	TGGTAAACGC CTATCAGGAC AAACAGTTTT ATTTTATAGT ACAGGTGCAA TTGCTACTAG	3780
55	AACTGCGAAG TTAGCAAAGG CTTTTAATAT GAATTTAATT GGTCTGAGCA AGTCAGGTCA	3840

TGCTGACATT ATTATAAATG CTTTACCAGA AACGCAAGAA ACGATTCaTT TaCTAAAGAA 3960
 AAAACATTTT GAATTAATGA AAGATGAAGC ACTTTTTATA AATATAGGAC GAGGTAGCAT 4020
 5 AGTTAAAGAA GCGCTCTTAA TAGAAGTATT AAAAAGTAAA GTTATTCGAC ATGCATATTT 4080
 AGATGTGTTT GAAAATGAAC CTTTGAAACC TAATCATGAA TTATATGAAT TGGATAATGT 4140
 AACTATAACA GCGCATATAA CTGGTAATGA TTATGAAGCA AAGTATGACT TATTAGATAT 4200
 10 TTTTAAAAAC AATCTAGTTA ATTTTCTCAA TAAGAATGGT CTAATTGAGA ATGAAGTTGa 4260
 TGCTAAAAAA GGCTATTAAA TGAATCATC ATGTAAATAT TGACACGCGC GCAATACTAC 4320
 15 AGTTATATTT aTAGTAAGTt AATaATgATT ATATAAGAA GatGGTgATA TAGATGAGTG 4380
 TTGAAATAGA ATCAATTGAA CATGAAGTAG AAGAATCAAT TGCATCATTG CGACAAGcAG 4440
 GCGTAAGAAT TACACCTCAA AGACAAGCAA TATTACGTTa TTTaATTTCT TCACATACTs 4500
 20 ATCCAACaSc TGaTGaAATT TATCAAGCAC TTTCACTGa TTTTCCAAT ATAAGTGTG 4560
 CGACAATATA TAATAACTTA AGAGTGTTTA AAGATATTGG AATTGTAAAA GAATTAACAT 4620
 ATGGAGACTC ATCAAGTCGA TTCGACTTTa ATACACATAA TCATTATCAT ATTATATGTG 4680
 25 AACAAATGTGG TAAGATTGTT GATTTTCAAT ATCCACAGTT AAATGAAATT GAAAGATTAG 4740
 CTCAGCATAT GACTGACTTT GACGTAACAC ATCATCGAAT GGAAATTTAT GGAGTTTGTA 4800
 AAGAAATGCCA AGATAAATAA TTTAACTTTG GTAGTATGAC AAATTAAAAA AGCGTTACTw 4860
 30 ACTTCATATA AGTAAGCGTA ATATTTAAGA nGTTAACGA CATGaAAGTt GTTTAACTTT 4920
 TTT 4923

35 (2) INFORMATION FOR SEQ ID NO: 552:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 917 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 552:

45 TCCGGCTTTA AAAACTTTTC CCAATTCCAG CTGGGCCTT TGGCATTAAAT ATTAACCTCC 60
 TGGTTCGGAT TAATTGGGAA CCTAACcNTT TAGGCAATAA TTGGTTTAGG CAACTTCCAA 120
 50 ATTGGTGGTT CAACCAACGT CTTTGGATAC CyTGcTCATT TAATTCTAAA ATGgTyrGAA 180
 CGCATTTTGG TACCCAAAAt GgTGACGTTT GTTTGCACGG TCTAATAAAT TGTCTAAGTT 240
 GTCGATTGGT TTCATTAAAT GAaCACCATT TTGCTCTTCA GCAAGACCTT TGTCTACTTT 300

55

GCGATCCATT GTTTGTTTGA AAAGAATAGC ACCCAAAATT TTATCTGGTG AGAATGAAGG 420
 TGAAGTTACC ACACGTGTAC GCATATCGTG AACAAAGTTGG AACATTTCTG CTTCAATTGCT 480
 5 ATATTGATCT TCGTTTACAC CATATTCTTT AAGTGCTTTT GGTGTACTAC CACCACCTTG 540
 GTCTAATGCG GCAATAAAGC CTTTTCATT TTTTCATTTTT TCTAATTGCT CTTTATTCAT 600
 ACTTTCCTACT CCTTAACCTT TCAATACACC TCCAGTATGA TAAAAATGAG AACATTTCTC 660
 10 AAGTCATAAA CCTTGAAAAG TGTATAAAAT GTGAAAAATA ATTGTCACTT TAATTAAAAA 720
 TATTATTTTA TTCTAGGTAT GACTAACGCC ATTAATGACA TAAAGAAAAT ATGTGTAATA 780
 15 ATCCAACCGA TTAATTCTGT CACACTAAAT TGAAAAATTG GACGTTGCGC AATAAATACT 840
 AAAAAGGGAT ACAATGCTAT AAATAAGAAA AATAAAGGGA TATAACATAG ATAGTAAAGC 900
 CTTTTAGAAG TATGAAA 917

20 (2) INFORMATION FOR SEQ ID NO: 553:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 432 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 553:

ACTCCCATAT CCACTGGCAA GATTATAAT TTGTAATTCT AAAGTATAAT AGCCTATATT 60
 TTTCAAATC TCTTTTTTTA TGATTAACAT CCCTCTAAGT GCACCTAAAG GTTTTTATGT 120
 35 AATGTATTAA ATATTTCCAT TTATACATAT ACATCTTGTA TAAAAAGAAA GAACTCCATA 180
 TATACTCAAA AGGTATACTG AAGCTCTATA TATTATATAA AATTATTATA CTATTTTGAA 240
 AAACATAAGT AAATCACTAA ACATGATTTT TTTCACTTTT AGAAAACCTT TAATACTATA 300
 40 AAAGCACCCA CTCAGTCACT AGTTTGGGCA GTTATTGTAT GCCTATTGAA CTCAATGCgT 360
 ATATTACAAT ACCTTTTtCG CATATTCATA TAAGacTTTG CATCTTTAAG CTTAATTGCT 420
 ATCTCTTTCT CT 432

45 (2) INFORMATION FOR SEQ ID NO: 554:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1374 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

	TTTTGCTTTA TGTAAATGG ATTATTaTTA GAATAAAAAA TCGGTGATGA GCTAAAAAAG	60
	TGTGTAGGAT GTTTTCCKAA CCCAATTTTT ACATCCGAAG ATATCGAACA ATATCTTCCT	120
5	ACTTCTACAT TATTAAAATC ACTACCRAAT CCAATATAAC TGTATTCACC AATGTGAGAA	180
	TTCCTGATTT TACACCATCT ATCTATATAG TTATTGCCAT CAAATTTTGA GTTTGTAATA	240
	TACGCCAAGC GATGAATCTT AACATTGAT TCTTTAGAGG ACTGGTTTTT CAGCAAACCA	300
10	ATTATCTTTT CAATCGCTAT CCTCATCGTC ATTTCTCTCA AGTATTTTGA TTGATAATAT	360
	CTTTATAGCT TTGAATAATT TTAACCTACCT TTGTCGAAAC GTTAGTGTCT TTATAATCAA	420
15	TAGCATCAAT CATCGGTTCC TTATTGTTTT GCATCTCTCT TGCTAGTTCA ACGGATTGGA	480
	TTAGATTGTT ATAGGTAATA CCACCTACAA TAACCGTACC TTTATCTAGT ACTTCCGGTC	540
	TTTCTGTGGA AGTTGGAATA AGGACACCAG GGAACCTCAA AATAGACGAC TCTTCTGACA	600
20	ATGTTCCACT ATCTGATAGC ACAACAAATG CATCTTTTTG CAATGCATTA TAATCAAAGA	660
	AACCAAATGG CTTTAACTGT TTAACCTAATG GATCAAATTC AAATTTACTT TCTTCAATTT	720
	TCTTCCAAC TCTTGGATGC GTTGAATAAA TCACAGGCAT TTTATACTTT TTGGCAATAT	780
25	CATTTATCGC ATTCATTAAT GATTTAAAT TCTTTTCATT ATCGATATTC TCTTCTCTAT	840
	GCGCAGaTAC TAAAATGTAT TGTGCGGTT CTAATCCTAG TTTATTTAAA ACGTCACTGT	900
	GATTAATTTT ATCTCGATGC GCTTCTATCA CTCTGTCTAT CGGTGATCCT GTyACAAAGA	960
30	TATTCGCTTT ATTGAAGCCT TCATCTAATA AATAACGTCT GCTATGTTCC GTATAAGGTA	1020
	GATTCACATC ACTGACATGG TCAACAATTT TACGATTGAT TTCTTCAGGT ACATTCTGAT	1080
35	CAAAGCATCT ATTACCCGCT TCCATGTGGA ACACAGGAAT CTTTAATCGT TTAGCAGATA	1140
	CTGCTGCTAA ACAACTATTT GTATCACCAA GAATTAAAAG TGCATCTGGT TGTTCGCGTA	1200
	ATAAAACATC ATATGTCTTC GCAATAATAT TCCCCaTcGT TTCTCCAnGt TACTTCCAAC	1260
40	TGCCTCTAAG TAGTGGTCCG GTTGTCTTAA TTCCAAATCA TCAAAGAAAA TTTGATTCAA	1320
	TGTATAATCA TAATTTTGAC CAGTGTGTAC TAATATCTGA TTAnAAATAT TGAT	1374

(2) INFORMATION FOR SEQ ID NO: 555:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1472 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 555:

TTACCTCTAG GCATGTCCCT TTCACGGTTT GCTTATGATA ACGTTATCGA CATTGTTAAA 120
 TTGTCCTTTT TGAAAAAATA ATTCTGGTGA GATAGACGCT ACTTGGaTTT CaTCaGTTTG 180
 5 TAAtAACGCA GtATAAtTAC CaTTACTAAA tTGaGTTAAT CGTTCaTATA AAGTACTAAT 240
 AGGATAATAA ATGTTATCCG TTAAGCGCGC CGTATAGTTC ACTTGATACG TTTCGCCTTC 300
 AACAAATTGCT TGTGGACAC GTTAAATATT AGTCATCATA ACCTCAGAAG ATTCAACAAA 360
 10 TGAAAAATGA TACTTTGATA CATAAGAAGT TTGATGTTCA TATGTTGAAT TTATGCTTTC 420
 CGCTTTTTCA AAActATAAG CTGCTGCATA AATATCATCT TTAGCTAATG AATGTGTACA 480
 15 CATAGCATGA TTAAATACT TTGCCGCTTC GTAActTAAA TATAACGAGA CATATCTACC 540
 TTGTCGtKgt GCGCTTGtGC AAAGTGTATC ACTTCTCCCA CATCAGCCAA CTTAGTAGCA 600
 ACATACTTCT TTATAAATCC CTTAATTGA ATATGGTATT GCTTATATTC ATTTTCAGTT 660
 20 AAATAGTAGC GATAATTATA TTCTATTCTC ACAGTAATCA CCTACCTTCG ATAAAAATAA 720
 TTCAACTTGT CGATAACCGT ATTCACTCAA AATAGATTCA GGATGATATT GCACACCAAA 780
 AACCGGAAAT CTAATATGCT CAAATGCCAT AATAATCGCT TCATCGTTTT TTGCTGTAAT 840
 25 CTTTAAGCAA TTGGAAAAG TCGCTCCGTC AGCAATTAAT GAATGATAAC GCATTACATT 900
 GAAATTTTGA GGCAGTCCTT GAAAAATACC TTCATTGGTA TGGCGTAACT GTGTAGTATG 960
 TCCGTGTACA GGATGATAGC CGTGAATGAT ATTTCCACCA AAATAAGACA CGATACATTG 1020
 30 AAATCCTAAA CATACACCTA GTATAGGTAC ACGCTGATAA AATTGTTCTA ACACTTCATT 1080
 CAAGATAGGA TAATCATCCG GaTTACCCGG CCCAGGCGAA ATAACAATTG CTTTTGGCTT 1140
 35 CATATTAATG ACGTCTTCTA TCAGCAGATT ATCAATACCA ACAACTTGAA CTGTTAGTTT 1200
 CGTTTGAGTC TTAATATAGT CTATTAAATT ATATGTAAAT GaATCATTAT TATCTATGAC 1260
 TAGAATCATT GTATACTCCG TTCTAAATGT GTTTTATTTT TATAATATGT ATTGGATGTA 1320
 40 GCTAAAACTT TAAAAGCATT GTCATTATCC TGACCTTGAT TTAACTAAT ATATGGTATA 1380
 TTCTATTGAT CGTACATAAA TGAATATCAG AGGTTCTAG CTGAAACCCT CTATAAAAAA 1440
 CTAGGCCATT GAAATTTCAA ACATTCGTTG GG 1472

(2) INFORMATION FOR SEQ ID NO: 556:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1054 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

	AGAACAGCAA GGATTACTTA CTGAGGAATT AAAGAAGGAT ATTTTAAAAC AGAACAAATT	60
	ACAACGTGTT GAAGACCTAT ATAGGCCTTT TAAACAAAAG AAAAAGACAA GGGCAACTGA	120
5	GGCGAAACGT AAAGGGTTAG AGCCATTAGC GATATGGATG AAGGCACGTA AACATGAAGT	180
	CTCAATTGAA GAAAAAGCAC AACAAATTTAT AAATGAAGAA GTGCAATCGG TTGAAGATGC	240
	TATCAAAGGT GCACAAGATA TTATTGCGGA ACAAATTTCA GATAATCCTA AATATAGAAC	300
10	AAAAATTTTA AAAGATATGT ATCATCAAGG TGTGTTAACT ACATCTAAAA AGAAAAATGC	360
	TGAAGATGAA AAAGGTATTT TTGAAATGTA CTATGCATAT AGTGAGCCAA TTAAACGCAT	420
	TGCTAATCAT AGAGTTTTAG CTGTTAATCG TGGTGAAAAA GAGAAAGTAT TATCTGTAA	480
15	GTTTGAATTC GATACGACAT CAGTAGAGGA TTTCATTGCA CGTCAAGAAA TCAATCATAA	540
	TAATGTAAAT CGCAGTTATA TTTTAGAGGC GATTAAAGAT AGCTTGAAAC GCTTAATTGT	600
20	CCCTTCGATA GAGCGTGAAA TCCATGCTGA TTTAACTGAA AAAGCTGAAA ATCATGCAAT	660
	AGATGTTTTT AGTGAAAAC TAAGAAATCT ATTACTGCAA CCTCCAATGA AAGGTAAACA	720
	AATATTAGGC GTAGATCCAG CATTTAGAAC AGGTTGTAAA TTAGCAGTCA TTAACCCATT	780
25	CGGTACTTTT ATAGCAAAAG GTGTGATTTA TCCGCATCCA CCAGTTTCTA AAAAAGAGGC	840
	AGCAGAGAAG GATTTTGTAC AAATGGTTAA AGCGTATGAT GTGCAATTAA TTGCAATTGG	900
	CAATGGTACT GCAAGTCGTG AAACAGAACA ATTTGTTGCA GATTTAATTA AAAAGCATCA	960
30	GTTGCCAGTA CAATTCATCA TTGTCAATGA AGCGGGCGCT TCAGTATACT CAGCATCAGA	1020
	AATTGCTAGA GATGAATTC CTGATTTTCA AGTG	1054

(2) INFORMATION FOR SEQ ID NO: 557:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1057 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 557:

45	AAATGTCAGA ATACAAGAAA AAAATAATTG AATTAATTGA AAGTAATTTA ACAGGATATG	60
	AAATTTCTAA AAAAAGTGA GTTTCTCAAT ACGTACTTTC ACAATTAAGA CAGGGCAAAC	120
	GCGAAGTAGA TAATCTAACC CTGAATACAA CAGAAAAATT ATATGAATAT GCCAATAAAG	180
50	TTTTGTAATT TAACTAATGT ATAAATTAAT CAAGCTATGT TTATTTGATT TAACTATTAA	240
	TAAAAATCAT ATGGTGAATG GATATTATAA TAATTAAAAT ACAAAAATAG TAGATTCCAA	300

AAGGGAAAAT AAGTGTTAAG TTTTAAATGA TAAAAAAGAT TGGAATGGAT CGTCTTGAAA 420
 TGCTCCCTTC AAAGTTTTCa TTTTTTCAAT GTCGACTTCG AAGGGGGCAT TTTCAATTAAA 480
 5 TTGTTATAGC TTTTATATT TGTATAATGA ACATATAAGT TTAAGAAGGT GCGAGTGAAG 540
 GAAATAAAAA AGCTCAAATG TACCAAATTG TTAATCTTAA TAAATCTCTA CTTTATAAAG 600
 ATTGAATGGA CATTGAGCG TTAATCAGTC AGGAGGGACT TTCCCTCCTA CAATTTAATA 660
 10 ATAATACTTG CTTCAACCACT ATACAAGGAG TGAGTTGTTA TGTTCAAAGT GAATTATTCCG 720
 ATTTTAAGTT ATTATCCAGA ATATAATATC GCAGTAAGTT GGCAACGTTT AAGAGAAGGA 780
 15 AAAACAATAA AAAACAAGAT TTAATACTGC TGCCTCATGA GCGCCTTGAA CATTATTTGA 840
 TGAATAAGTA TAATTTCAAC TATGATTATG CACATAAAAT TGTATCAAAA AAATACGATT 900
 ATTCAATTTT TATAAAAAaAG AAGGTGGATT AAATGCTTAC ATTAATAAAA TTGGAAAGAA 960
 20 GATGaACAGG tTATAATATA TGraTATATa CCTGaAGATG aTATAAGTAC CGGGTAAAGG 1020
 GTCCCGTACC TTTTAAATTA AAAAAGTTCC AGGGGGT 1057

(2) INFORMATION FOR SEQ ID NO: 558:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3754 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 558:

35 CTGTGCTGTA TTTACtTTAA TTTGACGACA TTTGAGAAGT ATTATGATGG nTGAGTTAT 60
 TTTATGAAAG TAATGTATTA ACAATCGAAT TACTAAGTCT ATATTGGTAG GAATATCGAT 120
 GTTTAGTTTA AATGGAATGC ACTATTTAAG TTTTAAATAT GGAGATGTTT GTGACTTTTC 180
 40 GATGATTAAG ATTTTATAG GTGTGCATCA TTTCAAATAA ACTTTGTGTT TAAAATTGAG 240
 CTTAGGAAAT CGATAGGTTT AGATGAGGAT ATTGTTGAAG TTATGTGTCT TGTATCCTTA 300
 GTTGTATATA AAGCGCAAAA AATAGCACCG CTTTCTCTTT ATCTGTGTAG AAAGGATGCT 360
 45 ATTATTGTAA AACAATAGTT TTAATTTnAT TTTCTGATaT ATCATATGTr ATTcTACCTG 420
 TATCAATTTT TATCGAATTA TAACCATCAA AATTATCAAC TTTATCATCA AAGTCTATCA 480
 50 CTTTTCCAAT TAATATTTTA TTATTAGTAA GCGTTAATTT GACTAATTG CCTATGTATG 540
 ACTGTAAATT CATATTTAAT CACTCCTTTT TAATATACGG AACTACATGA AAACCAGTTT 600
 TAGAATAATG AACCTTACCT AATTCGTTT CAATATATTT ACCATTCA CA TAAGATTTAC 660

55

	AAACTGGAAT ATCAATAAAAT TCTTTTTTAG TCATCTTTTC ACTTTCATTA GAATCTATTA	780
	TAGTATAACT TGGTAATAAA GATGAATTAT TTTTCTTATT ATTATATTCA TTTGTGCATA	840
5	AAATATGACG CTTTTGCTTA ACATTATTCA ATTTTCATCGT CATTTTACCA TTACTTATCA	900
	TTTCAAGCAT TTCTTTTTTA GCTTTTTGTA GAAAGGATGC TATTTTGTA TTTATTTAAT	960
	AATCTAATAC TTTTATCTC AGTTTCATCA AACGAATAAA CAGCAAAATC AGTCTGTATA	1020
10	TCTACAACAA GATTTCCTGT TTCGCTCTCG AATTCATTTT CATAGTCAGT TACAAACCCCT	1080
	TCGTATTCTT CATTATTCAA AAGCGTGATA ATAACATCTT TACGATATGC ATCTnCAATT	1140
15	CTCAATACTT TTCACCTACT TATCAATATA AGGTACTATA TGaGcACCTG GTCTTCGAAT	1200
	AATGGCACTT cCCTTTCTGG TTTCAATATA CACATTTTCG ATATGTATTT TTCCAATAAT	1260
	TTGATTAAAA TTAATAATCT CTTTCAAATC AAATCGCTCA TCACTTAATA TTAGATTGCC	1320
20	TGTTGACATT TTTTCTCTTA ACAATTCATT CAATAAATCT ATAGAAAGTA TTGTATAGCT	1380
	AGGCAATTTT TTATTATTTA AAATGGCTCT TTTTTTATTT TCATTATATA GGTGATGACC	1440
	TAACATATAT CTATTTTATT TTCCGCAATT TATTTCTATT TTTATTTTAC CATTTTTAAT	1500
25	CATTTCTTTC ATCTGGTTTT TAGCTTTTTc CTGtAATTAT GCTTCTTTTA CTCTACTTG	1560
	ATATTTACCT TCACGCTCTT TAAAGAACTT GTCCCGCCAA TTGCCAACAT GTGGCACTGT	1620
30	GGTACTTCTA CACCAAGGAT GCATAGGTGG CGCATTCACTA CCTGGTATCA TATCTTTAAC	1680
	TTTAAATATT TTTCCGTTAA GTGAATGACA TAATTTAGAT GTTTTACTAT CTATTTTGCC	1740
	AACATATTTA TATTCGCCAT CTTACCAAG TTCTTTTAAA TATGTTAACT TTTGTGCTTC	1800
35	TGCATTTTCA GTAAATAGTT AAAAAAGCGT ATAAAAATAG CACCACTTTC TCTTTAKCTG	1860
	TCTAAAAAGG ATGCTATTTA TCTTTTGAAT TTGAATTCIT TTTGCTTTT TCTATACTTT	1920
	CAAATCTTTC AACTAATTCT TTAAAAGATT CACTCAATTC TTTTGCAGTT ACATTTCCAT	1980
40	CTAATTGTGA ATCTAACATA ATTAAAATCA TCTCACTTTA TATTTAATCa TATTTATACT	2040
	ATAAAGTTTT TTCAATAATT TTTCAATATG GCTATCATTA ATGATATTAA TATGCGTAAA	2100
	ATATTTAGCA CAAAATTTAC TCACTATTTT ACCATGAAAC CTATTTGACT TGGTAATAAA	2160
45	TTTTACTTGT CCCTTATTAG TAACGATTGT CATTGATTTT ATTGATGGAT GCTTAAAAAA	2220
	TGTAAATAAA TCATATTCTG AAAATCCTGA CTGTCCAGGA TGGTTATGTA ACATAACAAT	2280
50	TGAATTCGGT TTAATGTTAA ATAATAATTC GGTGCTTGT TACCCTGGCA CAAAAGATAC	2340
	ACTATCTTGA TTGACATATA CTTTTGTAAA TTTACCATCT TTTAACAAAT AAGCTACTTC	2400
	ATTGCTATCA TTGTTTTCCA TGAGATATAC CTCCGTTTAT AGtCTGTGCA CTGATATTCC	2460

55

GAAGGTAAAA ATTTTAAAG TAACTTTCT TTATCCCAT GCTACGAGTT CAGATTCAGG 2580
 AAATAGCCCT TTACTAGTAT TTATGTATAT TCTGTCTATG GCATGAATAA AATAATTATC 2640
 5 TCTTGATTTT tTTTCTAAAC TAGATTTTTC AGCATTGATA ACTTCAAGAC TATCTATATC 2700
 CATTTGAATA ATACCAGGCT TAATATTTTC ATCATTATTA GGAAAATATT TATATGTAAC 2760
 ACTTTTATCA TTAATTTCTT TTATTTTAA TATTAGCAAT CATTTCCACC TCTAATTAAT 2820
 10 TAAAATACTA TAATTATATT TTATTTCTGT AAGTTTATGT GCCTCTATAT AGTGTAAT 2880
 ATACTTATTC ATTAGATAGT GTTCAAGAGC TTCATGTTTC TACATTATTA TATCCATTTT 2940
 TTTAATATTT TTCCCTTCTC TTAAACGTTG CCAACTTTGA GCCATATAAA AGTCAGGATC 3000
 15 AAATTGTTTA AATCCACTTT CTAATAAATA CTTATTTTCA AATATATGTT CATAAACTCT 3060
 TTGAATTA AAA TTTTATTTA TATTAgTATT TTTAGCAATT TTAGAAATCT CTATCTGTTT 3120
 20 ATCTCGATTT CTAAGTGAAT TATAATAAAT TTGAGCATGT CTGTTCCCTT TGATACCGTA 3180
 TTCATCACTT TTATTATTAA GTGCACCTGA TTCAATAAAA CAACCTTCTA CTTGATATTT 3240
 ACCTTCACGC TCTTTAAAGA ACTTGTCTCG CCAATTGCCG ACATGTGGCA CTGTGGTACT 3300
 25 TCTACACCAA GGATGCATAG GTGGCGCATT CACACCTGGT ATCATATCTT TAACTTTAAA 3360
 TATTTTCCG TTGAGTGAAT GACATAATT AGATGTTTTA CTATCTATTT TGGCAACATA 3420
 TTTATATTCG CCATCTTCCA CCAAGTTCyT TTAAATATGT TAACTTTTGT GACTCtTCyT 3480
 30 TTTCAACGAA TAATGAAAA AGCATATAAA AATAGCATCG CTTTCTCTTT ATCTGTGTAG 3540
 AAAGGATGCT TTAATACCaT GCTATTTTAT AATTTTcGGG AAATTcTTGC TTCTCGATAA 3600
 35 AGTCTCTTAC TACAGAAAA GACTTATTAC GATATAACAT AAAATATTCT TCATTTTCTA 3660
 TTTCAGATAA TGAGAAATCT ATAACCTCTG CATCTTTTTT ATTAAAAGTT ACTGAACCTT 3720
 TACCGTTACT TATATCATCT TCAGGTATAT ATTC 3754

40 (2) INFORMATION FOR SEQ ID NO: 559:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 815 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 559:

50 ATTTAACTAA AACTATAaAT AATCAAATGA TATTGGAAGA TATTaGCATA GATATCGAAA 60
 AAGGTAAATT GACTTCTTTA ATTGGACCTA ATGGTGCGGG TAAGAGTACT TTACTTTTCAG 120

55

CTGATTATAA AAATAATGAC TTGTCGAAAA AAATATCTAT ATTAAAACAA ACAAACCATA 240
 CTGAAATGAA TATTACGGTA GAGCAGTTGG TAAACTTTGG ACGATTCCCT TATTCTAAAG 300
 5 GTCGTTTGAC GAAAGAGGAT CATGATATTG TCAATGATGC GCTAGATTTG TGCAACTAC 360
 AAGATATCAG AAATCGTAAT ATTAAGTCAT TATCTGGTGG ACAACGTCAG CGTGCATACA 420
 10 TTGCAATGAC AATAGCACAA GATACTGAAT ATATTTTGCT AGATGAACCA TTAAATAATT 480
 TAGATATGAA GCATGCTGTT CAAATTATGC AAACGTTAAA AATGTTAGCG CATAAAATGA 540
 ATAAAGCGAT TGTCATTGTG TTACATGATA TTAACTTTGC GTCCTGTTAT TCAGATCAGA 600
 15 TTGTAGCATT GAAAAACGGA CAACTAGTTA AGTCAGATTT GAAAGATAAT GTCATTCAAA 660
 GTAGTGTTTT AAGTGATTTA TATGACATGA ATATTCAAAT TGAACATATA AGAAATCAAA 720
 GGATTTGTTT ATATTTTAAG GATTGATAAT TTGGAGaCAC TTTAAAGGGG TGATGCGCCA 780
 20 ATTAAAGAAG GGTAAACGT AAAGCATTTA TTTAT 815

(2) INFORMATION FOR SEQ ID NO: 560:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 919 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 560:

GAAACGAATA ATAAATTTAC TGAGTTATTA GTTGAAAAAG CTAATAAACA TGATGATGTT 60
 35 CTCGATAmGT TGATTAATAT TTTAAATAA GCGATACACA CTACTAAAT TGTATTATTA 120
 TTATGTTnAA TTGACnCTC CTAAATTTGC AAAGATAGCA ATTTAGGAGG CGTGTTTATT 180
 TTTATTGACG TCTAACTCTA AAAGATATAA ATTAGACATT TACAAATGAT GTAAATAACG 240
 40 CAATTTCTAT CATCGCTGAT AACAAATTCAT GGTTTAATAT GCAATGAGCA TATACTTTTT 300
 AAATAGTATT ATTCAGTAGT TTTAACAATC AATTAATTGG TATATGATAC TTTTATTGGT 360
 TATTTTTATC CCATAGTGTG ATAATTACTA TTTTTCATTC ATAATAAAGG TTTAAAGCAT 420
 45 GTTAATAGTG TGTAAGATTA ACATGTACTG AAAACATGT TTAACATAAT GATATAAGGA 480
 GTGACGTACA TGATCCGTCT AGGTAAAATG TCAGATTTAG ATCAAATCTT AAATCTAGTA 540
 GAAGAAGCAA AAGAATTAAT GAAAGAACAC GACAACGAGC AATGGGACGA TCAGTACCCA 600
 50 CTTTTAGAAC ATTTTGAAGA AGATATTGCT AAAGATTATT TGTACGTATT AGAGGAAAA 660
 GACAAAATTT ATGGCTTTAT TGTGTCGAC CAAGACCAAG CAGAAATGGTA yGnyGACATT 720

TATAAAGGAG CTGCTACAGA ATTATTCAAT TATGTTATTG ATGTAGTTAA AGCACGTGGT 840
 GCAGAAGTTA TTTTAACGGA CACCTTTGCG TTAAACAAAC CTGCACAAGG TTTATTTGCC 900
 5 AAATTTGGAT TTCATAAGG 919

(2) INFORMATION FOR SEQ ID NO: 561:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 518 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 561:

ATCATATAAA CCGGCTGCTT CTAAAAACGA AAATACTGTT ACTGGACCTA AAAATTTAAA 60
 20 CCCGTATTGT TTTAAATCTT TAGATAGTTG TGTGTCAGTA TCATCAACTG TGATACGATC 120
 AGAAGCATGT TCATACTGCA AATCTTTAGG CTTACCATTT ACATATGACC ATAAAAATTT 180
 ACTAAACTA CCATATGCTT GTTCAATTTT TAAATACCCT TGAGCTTGAT TAACAATTGC 240
 25 TTCTAATTTT TTACGATGAT GAACGATATT TGGAAAAGTC ATTAAGCGGT CGATATCTTG 300
 AGCGGTCATT TGTGCTACCT TTTCTGGTTC GAAATCATAA AATGCTTCTT CATAGGCTTC 360
 TTTCTTTTTT AAAATAGTTA ACCAAGATAG CCCAGCATGT TGTGATTCTA ATGCTAAAAG 420
 30 TTTAAACAAT GCCTTGChAT CATAGAGCGG TTGTCCCCAT ACATGGATCm TGATAGTCTA 480
 AGTAGACTGG GATCTTTAGT ACCAAATGCG CATTCAAT 518

(2) INFORMATION FOR SEQ ID NO: 562:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 562:

CTTATTTAAA AGTAAATCAA TCAATGTATT ATAATCCGAA TAGTCCGCAT AAAGCTGGTT 60
 TGCaGGCAAA tCAATTACTA CAACAAGCAA AAACCCAAAT TAATGCAATG rTTaATTCAA 120
 50 AAACAAATTA TGATGTTGTA TTCACTAGTG GTGCacTGAA TCCAATAATC TTGCTTTAAA 180
 AGGTATTGCC TATCGTAAAT TTGATACAGC GAAGGAAATA ATTACATCCG TGTTAGAGCA 240
 TCCGTCCGTA TTAGAGGTTG TAAGATATTT GGAAGCACAC GAAGGATTTA AAGTTAAATA 300

CAAAGTCGGT TTAGTAACAT GTATGTATGT AAATAATGTA ACTGGACAAA TACAGCCTAT 420
 TCCACAAATG GCTAAAGTTA TAAAAAATTA TCCTAAGGCA CATTTCATG TAGATGCGGT 480
 5 TCAAGCATTG GGCAAAATTT CAATGGATCT CAATAACATA GATAGTATTA GTTTAAGTGG 540
 ACACAAGTTT AATGGTTTAA AAGGACAAGG CGTCTTACTT GTAAATCACA TTCAAAATGT 600
 TGAACCAACT GTCCATGGTG GTGGTCAAGA ATATGGTGTT AGAAGTGGAA CAGTTAATTT 660
 10 GCCAAATGAT ATTGCAATGG TTAAAGCGAT GAAGATAGCT AATGAAAAT TTGAAGCATT 720
 GAATGCATTT GTTACTGAGT TAAATAATGA CGTCCGTCOA TTTTAAATA AATATCATGG 780
 15 AGTTTATATT AATTCTTCAA CTTCAGGTTT ACCATTGCTT TTAAATATTA GTTTTCCTGG 840
 CGTAAAAGGT GAAGTATTAG TTAATGCTTT TTCAAAATAT GACATTATGA TATCTACGAC 900
 AAGTGCTTGT TCATCTAAC GTAATAAATT AAATGAAGTA TTGGCTGCAA TGGGATTATC 960
 20 AGACAAATCT ATTGAAGGTA GTATAAGATT ATCATTGTTGG GCTACTACAA CTAAAGAAGA 1020
 TATAGCGAGG TTTAAAGAAA TATTTATCAT CATTATGAG GAAATTAAGG AGTTGCTAAA 1080
 ATAATGAAGT ATGATCACTT GCTTGTTAGA TACGGGGAGT TAACATTAAA GGGTTCAAAT 1140
 25 AGAAAGAAAT TTGTAAATCA ATTAAGAAAT AATGTAAATA AGTCaTTAAA AGGACTTGAT 1200
 GGGTTTGTCTG TTAAAGGCaA ACGAGATCGT ATGTATATTG AACTTGAAGA CCATGCaGAT 1260
 ATAAATGAAA TAACATATCG ATTATCAAAA ATTTTCGGTA TTAAATCTAT TAGTCCAGTA 1320
 30 TTAAAGTAG AAAAAACAAT AGAGGCAATA AGTGCAGCGG CAATTAAATT gCGCAGaATT 1380
 TGAAGaAAAC AGCACATTTA AAATTGATGT GAAGCGTGCC CGATTAAAAT TTCCCCAATG 1440
 35 GATACGGTAT GGAATTACAG CGTGAATTG GGGTGGTGCC AGTATTGGAG CACTTCGCCA 1500
 TATTTCCAGT GGATGTCCAA CGTCCCAGnC CCAGGAATT 1539

(2) INFORMATION FOR SEQ ID NO: 563:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 968 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 563:

50 ATAnCGTATA CATGTGTTCT TTAAATTGT GATAAGGAGT TTAGGATGGT TTATTTAAAA 60
 TCAATAGATG CCATTGGATT TAAGTCTTTT GCAGATCAAA CCAATGTTCA ATTCGATAAA 120
 GGTGTAAGT CAATTGTTGG TCCAAATGGA AGCGGTAAAA GTAATATTAC AGATGCTATT 180

55

ATCTTCTCAG GTGCAGAACA TCGCAAAGCT CAAAATTATG CTGAAGTACA GTTAAGATTA 300
 GATAATCATT CTAAAAAGCT CAGTGTGTGAT GAAAACGAAG TTATTGTAAAC AAGAAGATTG 360
 5 TATCGAAGTG GTGAAAGTGA GTACTACATA AATAATGACC GTGCAAGATT AAAAGATATT 420
 GCCGATTTAT TTTTAGATTG TGGATTGGGA AAAGAAGCGT ATAGCATTaT CTCGCAAGGT 480
 AGAGTTGATG AAATACTAAA TGCTAAACCA ATTGATAGAC GTCAAATTAT TGAAGAATCG 540
 10 GCTGGTGTAC TTAAATATAA AAAACGTAAG GCTGAATCAT TAAATAAACT TGACCAAACA 600
 GAAGATAATT TAACGAGAGT AGAAGACATT TTATATGATT TGGAAGGTCG CGTAGAACCT 660
 CTAAAGAGG AGGCAGCTAT AGCTAAAGAA TATAAGACAC TTTCACATCA AATGAAACAT 720
 15 AGTGACATTG TAGTTACAGT sCAGCATATT GATCAATATA CAAATGACAA TAGACAATTA 780
 GATCAACGTT TAAATGATTT ACAAGGCCAA CAAGCAAATA AAGAAGCTGA CAAGCAACGT 840
 20 TTAAGCCAAC AAATTCAACA ATATAAAGGT AAACGTCATC AACTTGATAA TGATGTTGAA 900
 TCgCTTAATT ATCAATTAGT AAAAGCTACG GAAGCCTTTG AAAAATATAC GGGACAATTA 960
 AATGTTTT 968

25 (2) INFORMATION FOR SEQ ID NO: 564: ...

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 436 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 564:

TTGTGTGTAT GCATTCAATG TGCTCTGTTT GTAAATGGCT AGCTATATAA TTTAgGATTC 60
 GAGGATCGTC ATCGACAACA AGACATTGCA CCATAGCTAT AACTCCCTT ATCTTTTTCA 120
 40 TTTATTATAC ATGTAAAATA TTTTTCGTA AAAAAACAAT TGTTCAATATT GAGTTCATAT 180
 TTCAACCTTA TACTGACGCT AAAGAAGAAA TAGGGAGAAG TGAATCGATA TGAAATTAGC 240
 GATAAAAGAG ATTATGTTTT ACAAAATTCG TTATATTTTA ATCACATTAA TCATTCTTTT 300
 45 ATTAAGTATT ATGGTGTTAT TTATTAGTGG TTTAGCTCAn GGGCTTGCTA GGGaGAATAT 360
 TTCGTWATTT GAACACTTTG GATAATGATG aaTATGtTGT TCaAAAAATG AAAGAGCCGC 420
 aAATTGaGAA ATCGCA 436

50 (2) INFORMATION FOR SEQ ID NO: 565:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2554 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 565:

	ATGTTTTGTC AATATCAATT GTTTGAGATA AATCCGCTTG TATAAACGA ACTTTATCAT	60
10	CATTAAATTT GCTAGTTAAT TCATTTATAT CAGTACGATA ATATTGCACA TAAACTTCAA	120
	ATCCATCAGT TAATAATTGT TTGACTATCT CAGAACCAAT TGAACCAGAA CCACCTAATA	180
	CTAATGCTTT CATTACTTTT TAATCTCCAA ACGACTATCG ACTTGCTGAT CTAAATTTAA	240
15	ATATAACGAT GACGTTTCGT TAATACTATC TAACGTGATA TTTTCAACAA TGTCTAACAT	300
	GTCAAACACG CTAACACCTT CAAAATACAA TTTAGTATAT TGATTAGCAA TATATTCAGG	360
	TGAGTTTAAA CTTGATATGA ATTCACCTAT AAATTGCTTT TTCAAAAGTT CAAATGCTTC	420
20	TGCATCTTGG AAAATGCTT TTTATCAGC CAACTCATCT AATAATAATT TTTTAATTT	480
	ATCTGGTTCT TCAGTAGCAC TTGTCACGAT TGAAAACTA TACGTCGGCT CTAGTACAAA	540
	TTGATAACCA AATGTATCAT CGATAAGTCC TTCGTTTAAAT AAATTCTGAT AAAAATCTGT	600
25	TTCTTCCCCA AAAATTAAC CAAAGAATAA TGACATTTCT AAATCACGTT GTACATATTT	660
	TTGAGGCGCT TCTTGTAATG GTTTATTTTT AAAACCAAGC ATTAGTCTTG GTGATTGAAT	720
30	TTTCATAGAT TCAGTAACAA ATGCTTCTTT AACATCCTCC GGTTCATCAA CAAGTCCTCG	780
	TTCGATTTTG GGTGGTTAA CTTTATTACG AGCATCCTCG TGTGTTTTTA CTATTCgACA	840
	TATTGCTTCA GGATCCACAT CGCCAACAAC AAATAAAACC ATATTTGATG GATGATAAAA	900
35	CGTTTCATAA CATAGATACA AATCATCTTT TGTAATATCG TATATACTTT CTACACTACC	960
	GGCAATATCA ACACGTATTG GATGTTGTTG ATACATTGCA CGCAATGTAT TaAACATTAA	1020
	TTTATATCCA GGTGCTCTT GaTACATTTT TATTTCTTCT GCAATAATAC CTTTTTCTTT	1080
40	ATCAACAGTT TCTTTTGTA AATAAGGCGT TTCaACCATT GTAAGTAAAC GTTTAATGTT	1140
	GTTTTCAATA TLaTCAGTTG CACTGAACAA GTAGCTTGTA CGATCAAAGC TTGLAAACGC	1200
	ATTTGCTTGT GCGTTATCTT CAGCAAACGC AGTAAATAAG cTTCTTCTTC TTTTTCAAAT	1260
45	AATTTATGTT CcTAAAAAGT GAGCAACTCC ATCAGGTACA GTAACAAATT GGTCTTGTC	1320
	AAGGGGTTTG AATTGATTAT CTAATGAACC AAATTGTGTA GTGTAAGTGA CAAATGTCTT	1380
50	TTGAAAACCT GGTtkGGGGA TAATAAATAA TCGTAAACCA TTTTCTAATt CTTGTTGAA	1440
	tACTCTTTTCG TCTATTAATT CATAATAACG CTCTTTCATT ATTTATCCCC TCCTTTTGTC	1500
	AACACATAGA TtGTATCTAA AAATGCTTTT TCAGCAACAG AAACAATATC TTCGCGACTT	1560

55

TTATGCATAA TCTCTATAAT ACTTTTCGGA CGATCTTCAG ATTCATATCG ATGAGAAAATG 1680
 ATTACTTTTT TAGCTAACTC TAATTTTTCT TCAGTGAAAT CTCCTGCTTT TATTTTTTCA 1740
 5 AATCACTTA TAATAGTGC TTTTGCAGTT TCGTACTTAT CACTTGAAAC CCCACTCAAA 1800
 ACAAATAAAT AGCCATTTTT GCCATCAATT TGTGAATGTA TAGAGTACGC TAAACTTTGC 1860
 10 TTTTCTCGCA CTTCAATTAAA TAAACAGAT GAAGGATCTC CTCCAAACAT CATGTTAAAT 1920
 ACAACAAAGG CAGCATATCC ACTTTGTCCA TATTGTGTTG GAAAACGGTA TCCCATATTT 1980
 AATTTAGCTT GATCCACGTC ATCATATTCA ACAATATAAT CAACTTCTTC ATCGTGATAA 2040
 15 TGATGAGTAG AATGTTGGAA TTGATGTTTA TCGAATGGTT TAAGTGCAA TTTTTCACGT 2100
 ATTTGTTTCT CAACACTTTC AGGTTCTACA TTGCCGACAA CATAAACAGA ACATTGATCA 2160
 TTATTAATCA TTGATTGATA TGTATGATAT AGTGTTCAG CAGTAATATG TGGGATTTGT 2220
 20 TCTAGTTGTC CTGTAGATAA GTATTTATAT GCTTCATTTT CAAACATATG GTCGAGTAAT 2280
 TTTAAAACG AATATTGTGC TTTATTATCT ACCATTGCTT CTATTTTTTT GGCTAATAAT 2340
 GTTTTCTCTT GGTTAACAAA ATTATCATTG AATGCTTTAT TTTCAATTAA TGGATTCCAA 2400
 25 ATGATTCTT GtAATAAATC TAATCCTTGa TTAAATAATG AwtCACCgkT TCyTAAATAA 2460
 CGkkCaTTAA caATTyCTAA tGaAAATGtA ATgACaTGCT GaTCTTTGAA TTTTGAAATT 2520
 30 GTACTATTCA CATACGCACC ATATAAATCG GCTA 2554

(2) INFORMATION FOR SEQ ID NO: 566:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1424 base pairs
 35 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 566:

TnTCGCTATT TtnAATTGGT TTTGTATGGT TTAAGTTATA TCAATATACA ACAnACCCTA 60
 45 AAGCTGATAT CCCAGGTATC ATTTTtagTA CGATTGGTTT TGGTGCTTTG TTATATGGTT 120
 TCTCAGAAGC TGGCAACAAA GGTGGGGTT CAGTAGAGAT AGAAACAATG TTTGCGATTG 180
 GTATTATCTT TATTATTCTA TTCGTTATTa GAGAATTAAG AATGAAATCA CCAATGTTGA 240
 50 ATTTAGAAGT ATTGAAATTC CcAcATTTA CATTACAAC AATTATTAAT ATGTTGTAA 300
 TGTTAAGTTT ATATGGTGGT ATGATTTTAT TACCGATTa TTTACAAAAT TTACGCGGAT 360
 55 TcTCAGCATT AGATTCCGGA TTGTTATTAT TACCTGGTTC TCTAATTATG GGTCTACTAG 420

	TTGCTGTAAT GACTTATGCA ACATGGGAAT TAACTAAATT AAATATGGAT ACACCATATA	540
	TGACAATCAT GGGTATCTAT GTACTTCGTT CATTTGGTAT GGCATTTATA ATGATGCCAA	600
5	TGGTAACTGC AGCTATTAAT GCGTTACCGG GACGACTTGC CTCTCATGGT AATGCTTTCT	660
	TAAATACGAT GCGTCAATTA GCAGGCTCTA TAGGTACAGC AATCTTAGTT ACTGTAATGA	720
	CAACACAAAC TACACAACAC TTATCAGCTT TTGGGGAAGA GTTAGATAAA ACGAATCCTG	780
10	TTGTACAAGA TCATATGCGT GAATTAGCAT CACAATATGG CGGACAAGAA GGCGCAATGA	840
	AAGTGTTACT ACAATTTGTA AATAAACTAG CAACGGTTGA AGGTATTAAT GATGCATTTA	900
15	TAGTTGCAAC GATATTTAGC ATCATCGCCT TAATTTTATG TTTATTTTTA CAAAGTAATA	960
	AAAAAGCAA AGCTACAGCT CAAAAGTTAG ATGCAGATAA TAGTATCAAT CATGAATAAA	1020
	TAAAATAAAT TAATTGAAGT GTGACTAATC AAAAATTATG TTGTGGGGAC ATGATTTTTA	1080
20	AAGTATCGGT GCCAAATATG GTTATCGATA CTTTTTTTAT TTGTTGATTT ATAGAATGTT	1140
	AGAGGAATTA TATTAAATTT TGGCATTGAC GTAGTAGGTC ATTAATAAAG AAAAAGCAGG	1200
	AAGTGGGTCA ACGAAATGAA TTTTGTGAAA ATAACATTTT TGTCCCAATC CCTACTATAT	1260
25	AACATTATTT TAAACGAGGC ATGCGATTAC GGAAGAATAA GCTTATAACA AGTAAACCGA	1320
	TGCTACAGCC AAGTAAATG ATGCCGTTAT GAATAGCGTC ACTTgCTGTA ATCACTTGAT	1380
30	CTGGTGGTAC ATTTAAATAA TATTTTTTGA AAACATCTGC AATT	1424

(2) INFORMATION FOR SEQ ID NO: 567:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 676 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 567:

	TCTCCCATGT TCTGCTAAAT GACGCATCAC TTTTACTTCA TGAGGCGTCA ATACACGTCC	60
45	TTCACCAGCA TTCAAACCGA CAACATTTAA AGGCCATAT TCAATACGAG ACAGTTTCGT	120
	CAC TTGATGA CCAAAATGTT CGAACATTCT TCTGACTTGG CGATTACGAC CTTCTGTAAT	180
	TGTAATTTCA ACCAATGTTG TGTTTTTATC TTTATCTTGT TTCTTAACTT TCACTTCAGC	240
50	CGGTTGCGTC ATACCATCTT CTAATTCAAT ACCTTTTCT AGCGCTTCA CTTCTTCTCT	300
	CATTAAATAA CCTTTTAATT TCGCAACATA TTTTTCCTTA ATTTGATATC TTGGATGTGT	360
55	CATTAAATTA GTAAATTCAC CATCATTTGT GAGTAATAAC AATCCAGAAG TATCATAGTC	420

ACGTCCTCTA TCATCAGATA CACTTGtKAT CACTTGAGTT GGkTTATGGA AAAAAATGKA 540
 AAAtTTTGTCT TCTAGTTCTA TTTTAATACC TTCAACTTCA ATCGTATCTG ATGGCTTCAC 600
 5 TTTTGTTCCT AATTCAGTGA CAGTCGTACC ATTCACTTTC ACTTTTCCTT CAGAAATTAA 660
 AGTTTCTGCC TTACGT 676

(2) INFORMATION FOR SEQ ID NO: 568:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 454 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 568:

20 GAAACGGTTC TACCAAAAAA CAGTAAGGGC TAAACCCAAT CATGGTAAGA CAAAAAGTAC 60
 AAATAGCCAT GCCCAAGTTG AACTCGCTGT ACGCCTATTT CTTTCTAAAA AGATAATAAT 120
 AAAAGCCAAT ACTAAATTAA TGATGAATCC AATGGCTAAA ATAATAGTAA ATAACGTTCC 180
 25 TAAATCGTTT GAAAATGTAA ATCGCATAGT CTTTTCTCCT ATAAAGAAAG GCACAAAAAA 240
 ACATTTTGCA CCTTTCACGT CATATTATTT ATTCACAGAT AAAGTTAAAA TTGcATTGAA 300
 30 TTCTTCTTCA TTATTTGGGA ATGTTCTTTC TTCTATTCTT TTAATAGTAA TATTTACTAA 360
 TTTTAAATT GTAGCTTCTT CAGAACTTAA AAAAGCATTa ATGTTTTTTTT CTAATAACTC 420
 KAAAGTCTCA GCTGtAAAaG TTTTAAGTTT AATT 454

(2) INFORMATION FOR SEQ ID NO: 569:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 894 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 569:

45 ACGATATTAC CACCCTCTGA AATAACATCA ATCATGTGAT CTGTAAAGTC CCAAGGATGA 60
 CTTGTTGTGA AACGAACTCT TGGAATCGCT ATTTTAGAAA TTGCTTGTA AAGATCTCCT 120
 50 AAGTCATATT CTATATCCTG TAAATCTTTA CCATAAGAAT TTACATTTTG ACCTAAAAGC 180
 GTTATTTCTT TGTAACCTTC ACGAGCAAGT TCACGTACTT CATCTATAAT GTCTTCAGGT 240
 CTACGGCTTC GTTCTTTACC TCTTGTAAT GGAACAATAC AATATGTACA AAACCTTATCA 300

TCAATAACGT CTCCTTCTTT AGACCATACT TCAACAACCA TTGCTTTAGA TAAGTATGCT 420
 TCTTCTAAAA TTTCTGGTAA ATGATGAATA TTATGTGTAC CAAATATCAT ATCTACATTT 480
 5 TGATACGATT TTAAAATTTT ATTCACTACT GACTCTTCTT GTGACATACA ACCACAAACA 540
 CCGATTAAAA TATCAGGTCG TTCTTTTTTC AAATrCTTCA AATTACCTAT TTCTACTAAC 600
 ACTTTGTTCT CGGCATTTTC TCTAATCGCA CATGTATTAA TTAAATAAC ATCTGCAGTG 660
 10 TTAATATCAG TCGTkGCTTG aTAGCCTAAT GCyTcmAGTA TACCAGCAAT GACCTCAGTG 720
 TCATGTGCAT TCATTGACA TCCATATGTT TTAATTAAAA ATGTACGCTC GTTCCCCATA 780
 15 CCGCGATATT TTyCATCAAT TtGGgAAATC nCTATTATAA CGAACTTCTk GTtTACCnCC 840
 TTTTTTnCGC TCCTTTAAAA TTAAGGCGGC TGATAAACAG GTCCAAAATA TTAC 894

(2) INFORMATION FOR SEQ ID NO: 570:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 441 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 570:

30 TATCAATCCC ACAGCACATG CTGAACAAGA TCAAACATGG GAGAAGATTA AAGAACGCGG 60
 TGAACCTAGA GTGGGTCTTT CTGCAGATTA TGCACCAATG GAATTTGAGC ATACAGTTAA 120
 TGGTAAGACT GAGTATGCAG GTGTAGATAT TGATTTAGCT AAAAAAATTG CGAAAGATAA 180
 35 TAATTTAAAA TTAAAAATCG TCAATATGTC ATTTGATAGT TTGTTAGGAG CTCTTAAAC 240
 TGGAAAAATT GATATTATTA TTTCCGGAAT GACTTCAACG CCTGAACGTA AGAAGCAAGT 300
 TGATTTTTCA GATTCATATA TGATGACTAA AAATATCATG CTTGTAAAGA AAGATAAAGT 360
 40 TAATGAATAT AAAGATATcm AAGACTTTAA TAATAAAna GTnGGGGCAC AAAGGGACTG 420
 AACCAGAAAA AATCGCTCAA C 441

(2) INFORMATION FOR SEQ ID NO: 571:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1205 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 571:

55

AAGTCGTTCA TCTTTTCTA CAAGGGTGTA AAAATAATCA ATCATATCGT ATAACGCTTC 120
 TTTACCAATG ATGTCATATG GTGTTGTTGT CATTTAATCA CCCATTTTCA AAAATTTACT 180
 5 GTTACGAACT TAAGTTAATA TATACTAAT ATAACATGAT TTAAACATT TGAAAGAAAT 240
 ATGCATATTT GCCAATTTAA TTTATATTGT TTGAAAGTGT TTCTTTTTTC TTGAAAAAAC 300
 GTTGAACTTT ATTTAAAGGT tGATGATGTT CGAGGTTTAG TTCGTTtAAT AAAGATtGGA 360
 10 ACTTTTGTA ACCTTGATTA TAGTCTTTAA CTTGGAAGTC TAACTCATAA TCCGTAGTAT 420
 CGAAATACTC ACTTTTATCT AAAACCACTA AATCACCTTT ATATTTAGTT TCTTGGCGAT 480
 15 ATGTCGTTAA TGCACCAAGT ATTGATAAAG TTGTATCTTT TACACCAAAC TGTTCAACTA 540
 TAATTTGACG AATGTCATCT GGAAGATTGT CGTTTGAAAT AATCAAGTTC ATCTCTGGTT 600
 TAATGTCGAC GATATAGTTG TATTCTAATA GACCAACCTT TGCTGGTGTC TTTAAAGTCA 660
 20 TTTCATATTG ATTGTCTTTA ACTCTTATGC GTAGTGCAGA GCGATGTTCC TTTAATTTGA 720
 AATCGGGTGT ATCAATATAG TAATTGACTT GCTTAmAAAG CACACTGTCT TTAAAATATT 780
 TCTCTTGCAA TTTATTATAG ATTGAtGCAG TTATCATTTG TtTAAATTCT AttTCATGAT 840
 25 TTGTTGCCAT GATATGTATA CACCTCGTAT CAAATTCAAT TTATCTTAAC TATATTATGA 900
 ATGACAAAGT TGAATTTTAA AAGTAATTTT CTTTATCTAT TATCAATGTT AATTTGACCA 960
 30 TTAAAAATAG TGTTGTAAG TGTTTTGTAT TATTGAtTTG TGTTAAATG TTATGGAATA 1020
 AGAGGAGGAT TAAGCATGSG TTTWTATATT AATGAAATTA AAATTAAAGA TGACATACTT 1080
 TATTGTTATA CAGAAGATTC TATTAAAGGA TTATCTGAAG TAGGACAAAT GCTCGTTGAT 1140
 35 AGTGATAATT ATGCCTTTGC GTATACATTA GATGATGGTA AAGCGTATGC TTATCTCATT 1200
 TTCGT 1205

(2) INFORMATION FOR SEQ ID NO: 572:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 570 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 572:

50 TGAAGAAATA GCAATGATGA AATGCCGTAT GTTATATGAG ACGGGGTCAT TTCTTGAATT 60
 AAGAGAAGAA ACAATTGTCT TATTGAAAAC TGGCATACAA CAATATGATG CATTGATGAT 120
 TTATTACGTG AAAAGTTTGA TTGGTTTGGG ACAATATTTT GAAGCGGTAG AAGTAATTCA 180

ATTTGCTAAG TCAAAATTAA TTGAAGATGA AAAACGATTG ACTCAGTCAT TAGCTGATTT 300
 TGTTACGTTA TCAATGAGGG AACAGACGCA CTTGATTTTG AAGTTAATAG ACAATGGTCA 360
 5 TTTTCAATTT CAAGAAACGG TATTATATAT ATkAAaAyCT AATaCGTACa GTtATAACCT 420
 CATTAGTTTA ATGATTGAGT ATTTAAGGTT CGCAAATTGT ACACAAGAAC TGACAATTGA 480
 10 AAAGTATGGT ATGGATGTAA CTTTTGTACC AGCTAATTTA AAAGGGCTAG AACATACAAC 540
 ACTTAAAGAA AAAGTTATAC CTAACGTTAT 570

(2) INFORMATION FOR SEQ ID NO: 573:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 939 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 573:

GTTGAATGGT TAGCAGCTGC AGTTGTATTA TATTTCTGTG GTGTAATTGT TGACGCTCAT 60
 25 GTATCATTCA TGTCTTTTAT TGCAATATTT ATCATTGCTG CATTATCAGG TTTAGTCAGC 120
 TTTATTCCTG GTGGTTTCGG CGCTTTCGAT TTAGTTGTAT TACTAGGATT TAAAACTTTA 180
 30 GGTGTCCCTG AGGAAAAAGT ATTATTAATG CTACTTCTAT ATCGTTTTGC GTACTATTTT 240
 GTACCGGTAA TTATTGCATT AATTTTATCA TCATTGAAT TTGGTACATC AGCTAAGAAG 300
 TACATTGAGG GATCTAAATA CTTTATTCCT GCTAAAGATG TTACGTCATT TTTAATGTCT 360
 35 TATCAAAAGG ATATTATTGC TAAAATTCCA TCATTATCAT TAGCAATTTT AGTATTCTTT 420
 ACAAGTATGA TCTTTTTTGT AAATAACTTA ACGATTGtKt ACGATGCTTT tATATGATGG 480
 AAATCACTTA ACGTATTATA TTCTATtGGC AATTCATACT AGTGCTTGTT TATTACTTTT 540
 40 ACTGAATGTA GTTGGTATTT ATAAGCAAAG TAGACGTGCC ATTATCTTTG CTATGATTTT 600
 AATTTTATTA ATCACAGTGG CGACATTCTT CACTTACGCT TCATATATTT TAATAACATG 660
 GTTAGCTATT ATTTTGTTC TGCTTATTGT AGCTTTCCGT AGAGCGAATA GGTTGAAACG 720
 45 CCCAGTAAGA ATGAGAAATA TAGTTGCAAT GCTTTTATTC AGTTTATTTA TTTTATATGT 780
 TAACCATATA TTTaTTGCTG GAACGTTATA TGCATTAGAT ATTTATACGA TTGAAATGCA 840
 50 TACATCTGTA TTGCGCTATT ACTTCTGGCT TACGATTTTA ATCATCGCTA TCATCATAGG 900
 TATGATTGCA TGGTTGTTTG ATTATCAATT TAGCAAAGT 939

(2) INFORMATION FOR SEQ ID NO: 574:

55

(A) LENGTH: 1059 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 574:

10	GAATTAATTA AATATTACAC ACAGCCTCAT TTTTCATTTT CAAATAAATG GCTATATCAA	60
	TATGATAATG GAAACATTTA TGTGAACTT AnGAGATATT CATGGTCAGC ACATATATCT	120
	TTATGGGGCG CTGAAaGTyG GGGAAATATT AATCAGTTAA AAGATCGTTA CGTAGATGTG	180
15	TTTGGACTAA AAGACAAAGA TACTGATCAG TTATGGTGGT CTTATAGAGA GACATTTACA	240
	GGTGGCGTTA CACCAGCCGC AAAACCTTCT GATAAACTT ATAATCTTTT TGTGCAATAC	300
	AAAGATAAAC TACAAACGAT TATTGGTGGC CATAAAATAT ACCAAGGCAA TAAACCAGTA	360
20	TTAACATTGA AAGAAATCGA TTTCCGTGCA CGAGAAGCGT TAATAAAAAA TAAAATATTA	420
	TATAACGAAA ATCGTAATAA AGGTAAGCTT AAGATCACCG GTGGCGGTAA TAACTACACT	480
25	ATTGATTTAA GCAAAAGATT ACATTCAGAT CTAGCAAATG TTTATGTTAA AAATCCTAAT	540
	AAAATAACTG TTGACGTCCT CTTTGATTAG TATATGAAGG TGACTTATAC TTCATGCACT	600
	TTAATTCCAA ATCAGATTAT TTAAATGATA ATTTTAAAG TGTATGATGT ATATAATAGG	660
30	TAAAATTTTC TATATATTTA AATGGAATTG GGAGTAGGAA TGTGACAGAA ATAGTATTTT	720
	ATAAAATTTA TTCtTGTCAC TCCCCAACTT GCACATTATT GTAAGCTGAC TTTCCGCCAG	780
	CTTCTATGTT GGGGCCCCGC CAACTTGCACT TGTCTGTAGA aTTTCTTTTT GAAATTCTCT	840
35	ATGTTGGGGC CCCGCCTATA ATTGAAAAAT GCTTGTTACA TGGGCATTTT CATTCGGTCA	900
	ACTACTACCA ATATAATATT GtAGaGCCTA AGACATTGAT TTATTATGTC TTAGGCTCTA	960
40	TTCCTTCATT TAATGATTAA nTTATTATAG CAATACTTTA TTGTCCCATG ATTAGTGTTT	1020
	TTTTAATGAG ACATAGTAAC TATAAAGTTT AATAATCGT	1059

(2) INFORMATION FOR SEQ ID NO: 575:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 574 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 575:

55

GTTTGCTTTA GGTCTGTTT CATATTTATA CTTCGAAGGA TTTACCTTTT TGaAGTCTGG	60
--	----

TTCAACTKGC TTTTATTCT TTTGAAATC AGCTGGTTGA GTAGTTATGA GTTCATTATT 180
 TTTATTAGmA TAAATCTTAC CATTAAACATA TTTATAATCT TTTGTTATAA AGTCACCATT 240
 5 TCTGAATGGA ACTACTTGAT TATGACCTTT AGAGAATAAA TCAGTACCGA ACATTAAATA 300
 GTTCTTCGTA TCTATACCAG CCAAATGTAA AATTGTTGGC ATTACATCGA CTTGACCAGC 360
 ATATTCATTA TTGATACCAC CAGATTTACC AGGGATTTTA ATCCAGAAAC CAGTTCTGTT 420
 10 TAAATCTGTA AATTTAGCCG GTGTGATTTT TTCACCTAAT AGTTTTTCCA TGGCATTGTT 480
 ATGGTTTTCA GAGATACCAT AGTGGTCACC ATAAATCATA ATCACTGAAT TGTCATATAA 540
 15 TCCTTTTTTC TTCAAGTCAT TAATATATTC TTCT 574

(2) INFORMATION FOR SEQ ID NO: 576:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 796 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 576:

CAATGTTTTA TAGTACAATA TATTTTAAAT AATACTCGTT AAGGAGAATG ATATGATATC 60
 AATTCACGCA ATTCGACAG GAAAAATCCA AGATTTCCT TATAGCTCGA AAAGACCGAT 120
 30 GCGCTCTGCT TTAGATAAAA CTAAGATTTC ACAAACAACA TGGTTATCTT CAACTGGTTT 180
 CACTGGTGAT GAACAGGCTT ATAAAGATCA TGGTGGACCA CATAAAGCAG TTTGTGGGTT 240
 TAGTAAGCAT AATTATGCAC TGTATCAAGA TGATTACCT AACTACCTA CTCATGCGAT 300
 GTTTGGAGAG AATTTAACAT TTGATTATT AGACGAATCT GATGTTTACT TTGGTAATCA 360
 ATATCGTTTA GGTGAAGCGT TAATTGAGGT TTCTGAAATT AGAGAACCAT ACTGGAAAAT 420
 40 TCAAGCAAAA TATAATATTC CTGATTTAGT GAAGCGCATG TCTACATCTG GTAAACAGG 480
 TTTCTATTTT CGGGTATTAA AACAAGGCTA TGTATCTCCA AATGATCAGC TTTACTTAAT 540
 ACAAGAAGCA CCAATCGAAC ATCGTTTATC TGTACAACAG CTTAATGACC TTTATTATAA 600
 45 TGATAGACAA AATCAAGaTA TGTTACGATA TGCACTAAAC AATCCATTTC TGTCACCAAC 660
 AAGACGCGAT AAACCTTCAA AAATGTATAA CAGAACATTG GAAATAATTA CCTTTCATTn 720
 50 ATAAGTGTTA AATGAACCTT TCAAAACaNa AAGGAATCAA CTTACACAT CGTTTGTATG 780
 AATAGTCTTA TCTATA 796

(2) INFORMATION FOR SEQ ID NO: 577:

(A) LENGTH: 1095 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 577:

```

10 AGAAATTATG ACAAATATA AAGATGGAAA GTTAGTTTAT GCATCAGTCG AACCAGGATC      60
   TTACGTAATA CATAAAGATG ATGCAATTAA ATATGACGAT TATTCTAAGT TAAAAAATT      120
   AAGTCAGCTA ACTAACTTG ATCATCCAAA ACCAGTTCCA TATAGCGTAc TCAAATCAAA      180
15 TCTTTCGGAG TACCTTTAAC AAGCGTTTCA TTTATGACAC ATGGATCAAA GGATACTAAA      240
   GATGAAGTGT TGCCGGCATT GGCCTATTTT ACTTTTTCAC CAAAAAATTA TGAAGACAAG      300
   TCTAATCCAG ATCCAAAAGT TTAAATTTA GTACATATGG ATTTCTTAAA TGCATCTAGT      360
20 GATTTTGGTA ACGCACATTT TGTGTTTTTA AGTAAATATA TTAAAGAGTA TGAATCAAAC      420
   TATGAAACAG CGTCAGATGA TTCTTTAAAA TAGTATTTAC TGTGTGAAAA ATAAATAGTG      480
   TACTACATTA AATAATCGCA ATAATAATCC CGATAAACAA TCAGCATTAC TGCTTATCAC      540
25 ATAGAGTTCT TAATAACTAT AACTCTATGA TTCGCAAATA ATAAATGATT GTCATCGGGA      600
   TTTATTTTTT TCAATTTATA AAGTGACATT ACCTTGTTCA TCAGCAGGTT TGAAAACAGT      660
30 AATCACTGCA CTAATAATTG CTAAAATGTG TGGGATACCT GTCCAACAGA ATATTAAGTG      720
   TAGAATACCT TGCATATTCT TGCCGGCATA AAATTTATGA ATACCAAAC TACCTAAGAA      780
   CAATGCTAAT AAAATATAAA TAACTTTGTT TACTTGCAAT TCTTCCCTC CAGTTGAATT      840
35 GCTTATAATG ACATTAGCTT CTCTTTTTAT TATACCCACT TTTAGTTCAA ACATTCTAGT      900
   TTAAGCATT CCAATCATCT AAATTTAGT TATTCAATCC TTACAATAAA TTTAGGATTA      960
   CATTTAGTT GCATTGTATT ATTTTACGTG TGAAATATAC GTAATGAATC ACATGACAAY      1020
40 CTYCAAATTG AAAAATATAC ATTCTATGAT GTAAGGTCGC ATTTTAAATA TATTTACGTh      1080
   AAAATAGTTT GGATG                                          1095

```

45

(2) INFORMATION FOR SEQ ID NO: 578:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 489 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 578:

55

ACTGCCAAAA TAATCATAGC CATGTTCTAC AGCTGCTTTC GCTACAATAT CCAAACGCAT 120
 TTCAAAACAA GCGGTACAAC GTAAGCCGCC TTCTTTTTCa TCAGCTAATT CTTTATCCTT 180
 5 CACCATTTC ACAAACTTAT GTGGTTCATA AGGTGCTTCA ATATACTTCA CATTGCGACC 240
 AGTCTTGCGA TTAATATCTT CCACAAATTG TTCTTGCACT TTAGCACGTC GTAAGTACTC 300
 ATTTTTCGGA TGAATATTTG AATTCGCGAA ATAAATTGCA ATGTCTGCAT ATTGTGTTAA 360
 10 AAATCTAAT GTATATGTAC TACAAGGTGC ACAACAATA TGAATAAGa TTTTAGGTCT 420
 GATTGCTTCT CTTTCCCACT GsCCGATTAA TnTCTTCAAC ACCTTGTTCA TAATTAATTT 480
 15 GTnGATTTn 489

(2) INFORMATION FOR SEQ ID NO: 579:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1287 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 579:

CGTACTGCAT CATTGTGTTGA AAAATTGCA AACTTCTTCA AGATGTTATG GCTTAGACTT 60
 AAAGCGATGA AGCACTACAA AGCCTTAAAT AAAGAATCTA AGAAGCAAGA ATTTGAAAAT 120
 30 TCATTCAAAG ATGTTCAAAA AATTATGCGT ATTGTGAATC ACAATATTAT TTTACGCTTA 180
 AAAGAAGAAC AAAATAGTAC AAATGTACTT GAGGTAGCT TAGTCATTAA TCATTACTAT 240
 35 GATATGAGTC GCTCATTAAA GTGGCGTGCA CAACGTCGAA AAGAACGTCA AGAAAACAGC 300
 AATCAAATCA TACCGCAAGC TATGTTCCAT AACCACAAAT TGGAAGCATT GTACTTACAA 360
 CGTCATCTTT TAGATGAATT AATTCGCAAA AATAAAATCA ACAATATCGT TGCAGCTCAA 420
 40 ATTCGAGAAA ATATCAATTa CAACGAAATT GTCTTGCTT TACAGTCCAA ACATTaAGCA 480
 AGaCwTaCaw TmCCCCCGTA CATATGAGAC AAAGTCATTA TCATCTCATA TGTACGGGGT 540
 TTTTATATTC AACATCAAAA AATCAGATTG ATGAAAAGTA AATAACCTTT CATCAATCCG 600
 45 ATTTGATTAT AGAATCTATT TTTTAAGTTT AAATGGAATT GTACATACGT TAACATTCTT 660
 TTGATAAATT AAATACAATT TCATACGCAA ACTAGTTTGA TTGTGTAATA AATTATGCCA 720
 50 ACGTTTCTTA GTAATAAATT CTGGTATCAC TACTGTAATC ATATAGTTTT GATCGTTGGC 780
 TTTACGATTA ATCTTATCGA TAAAACGTGA AATTGGTCGT ATAATACTGC GATATTCCGA 840
 ATGTAAATA ACTAATCTTA CATCTGGGAA ATGACGTTTC CATTTCTCTT GGAATGCTTT 900

55

TCGGTAATAA ATAGACTTAT CAACTGCTGT TGTAATACTT GTTATCGGCA CAATTGCTAA 1020
 ATTACGATCT ACCACGTCCA CATTAGAAGC ATCAATGTCA GAACGTAATT GTTCTGCGAT 1080
 5 ATCTCGATAA TGKTTGKTAA TTTTCAAGAA GAAAATCACC ACGAACGGCA AGAAAATAAG 1140
 TATCGGCCAT ACTTGGCTAA ATTTAGTTAT GAGTAAAATC cATAAnaACA ATAAATGTCA 1200
 10 CGATACCACC AAGTAAGTTC ACAGACAAC TACTTAACCA ATTCTTAGGA CGTTCATGAA 1260
 TCCATTTAAT AACCATACCG AATTGTG 1287

(2) INFORMATION FOR SEQ ID NO: 580:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1223 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 580:

GCCTCCTTTC CTATATCCAT GGGTGTCTTA GTTGGGAATG GCTTATTCCC TATCCCAAAC 60
 25 ATCTGTCAAT TCAAGACTAT CACTGTATAC TAAATCGCCT nTCTATACnT CCATnCATTT 120
 CATTTTCCAG TTGGAACAAT TTGTTTAATT ATTTTAAATA TCATTATCAA TCCAGAAGTA 180
 30 TTTACTATTC ACTTTTACAA TAATCAATCA TTAACTACA CATGGGTAGT CGGTGGTTTG 240
 CTTGGGGTTA GCTTTTTAAC TGGCAATTTA TTGTTATTGC CAAAATTAGG TGCAACATTA 300
 ACTGTAATTG CAACAGTTGC GGGTCAAATT ATTATGGGTG TCATTATTGA TACATTGGA 360
 35 TTATTTGGCG CTACAATTCA TGATTTTAAT TTAATTAAAG CAATTGGAGT ATTGTTACTC 420
 ATTGTCGGCA TCGTCATAAT GAATCAATTT AACAAGAATA ATTTATTACT AACTGATCAA 480
 AAGTATTTAC TGTTTTGGCT TCTATTAGGA TTTATTTTTG GTTCTTTTCC ACCTATTCAA 540
 40 ACGACAATTA ATAGTGCTTT AGCTAGTCAT ACTCATTAC CAGCCTTTGC ATCATTAGTA 600
 TCATTTACAA TTGGGTCAAT AGCGCTATTG ATTTAACC GATTTTTTAA TCGTTCCTTA 660
 AACTAAAAA CAAGTCATTT AAAATTCGGT AAATTAAAGC CTATCTATTT TACTGGCGGT 720
 45 ATACTTGGCA TGGCTTTTGT AACAGCTAAC ATTATCTTAA TGCCTCATAT GGGTGCAGCA 780
 TTAACAACAC TTATTGGGAT GTTTGGCCAG ATTCTAATGG GCATATTGAT AGATCACTTT 840
 50 GGATTATTTG GTTCACCTAA AATAGCAATG ACATCCAGAA AACTATTGG TCTATTATGT 900
 ATTTTGACAG GCATTATACT TTAAGATTA TTTTAAATTA ACTTTTAGCT TATCATTTTA 960
 ACTTGTAATT ATTTTAAAAA GTGATAAGCT ATTTTTTTGT GGTCTAAAAA TCTTTAGAAA 1020

55

CAACTCATT C TTAAGACCTA AATTAATGTT ATnTTTTTAAT AATTTACACC AAATTAATAG 1140
 CAAAAATTAT GTTATTTCGTG CTAATATTTT ATAGTTGGTT ATTCAATTAA TTAAAAATAA 1200
 5 GTCAAAATGC ACAACTTTTT ATn 1223

(2) INFORMATION FOR SEQ ID NO: 581:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 454 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 581:

CACCTTTTTA TCATGCTTAG TTATCAATAA ATCTATATTT TGCTGTTTTA CAATTTTTTT 60
 20 AACTTTATCA ATCTCATTAT CTTGGACTAA ATAAATATAT GATCTTGCAT CTGTTGCTAG 120
 AGCTTGTTTCG TGTTTTTCTG ATAAACATA TGTGATGGAA GCGTGAATAA TAATGCCTAA 180
 TGTAACAAAA CTGATAATTA ATATACTGCT TATCAATAAC ATTAAGCGGT GGTGAAACTT 240
 25 CATCATTGTT CTTTAGGTCT TTCCAATTTA TAGCCTAAGC CACGCACAGT TTTAATAAGT 300
 TGTGGCTTCT TAGGATTATC TTCTAATTTA TCTCTTAAAT GACTGATATG TACATCAACA 360
 ATTCTTGAGT CTCCTGCAAA TTCATAATTC CATACCGTAT TTAACATATG CTCTCTCGTA 420
 30 ATGACTCTGC CTTGTCTTTC TATCAAATAA AGCA 454

(2) INFORMATION FOR SEQ ID NO: 582:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 452 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 582:

TTTATAAGAT TTTATTTACA AATAATTGGT TTTCATATGT ATAAACACTT TTGACTTTCA 60
 45 AATCTTAACG ATAATTCTAT TACAATACAA TCCCCTATTA GAATGATTTA TGTAATAAAA 120
 AAAGCGGAGT TTCCCCCAGC TTTTCTAAAC GACTACATAA AATATAAGAT TGCAATTAAA 180
 TGCAATAGTG ATGCTATTAC AATAAAAAA TA GCCAAATCA TATGAAAATA TGGTCTATTC 240
 50 TTTTGTGCAT AAAACCATGC ACCAATTGTA TAAGACACAC CACCTAAGAA AATGAATAAT 300
 ATGAATATCC ATGATGTGCG AATAAAAAA ATTGGTAACA AGATAATACC TACCCAGCCC 360

AAAATCCCCC AAAGTGTCTG TCCCCATAAT AA

452

(2) INFORMATION FOR SEQ ID NO: 583:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1472 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 583:

15	CGCTTTTTCG AATAAGTCAT TAGCCGCTTT TAAGCCTTCT TCTTTTCGAT CTACAACAAG	60
	TAAAATAAAT GGCTTTAACG CTTCTTCTTT TTCACCTTCA AGCATATCTG GTTTTGAAC	120
	CATTTCAAAT GGAGATTCA ATCCATTATT ATCGCTCATT TCAATAATTG CATCATACTG	180
20	TGCTTGTGAC ATACTTGCAA TAGCCTGTTT TGCATTTTCT TGAAGGAAAT ATAAGTTTTT	240
	CAATTAGGA TGCTTATTTA ATGTACTTAA TGTAATCGGT GTAATGTCTT TCTCATAAGA	300
	CACITCAATC ACTGTACTAT TTGTTCTACC AGGAATTGGT GGTTTTTTCAT GAATATGCTT	360
25	TGATACTTCT CCAATTCCAA CGACAGATTG ATTTTTCGTT CGATTATAAA AAATAATATT	420
	GTCGCTTCT TCTAACTGAG TATAAAAATG ATAACCATTG CGTTTAATAC CGTTGTACGT	480
	GTGCGTATAA ATCGTATATT GGTTCACAGG TTCAAATTCT TCAGTTTCAG CTAAAAAGAA	540
30	ATAACGCGGT ATCTTAATTT CGCCTTTACC AAGACCACTT ATTAAATCAA ACTCTTCTGC	600
	AGTGATTGGA TTGAACAATG TCTCTTTCAT ATnACTTATA CGAAATTCCA AAGCTTCACT	660
35	ACGCTTTAAA TAATCTGCTG TTAATGGTTT CAATTGTTCA TTAAAACGAA ACTGTACACG	720
	TATTTTATTT TGTGCACCTG TTTCAACACT AATAATTTCA CCACATCCAA GTAGTCCAGT	780
	ATCCGTCTGA ACTTGATAAA AGATGACTTG ATCTCCTACT TTAGCCTTTT TAAACGCTCT	840
40	AAATCCTTGA GATGGGTAA AATGTGCGCC TGATTCAAAT AAAGCTGTTT GTCCTACTAA	900
	CGGTTCAATTA TGATTCCAAC GGTTATATCC ACAATTCAAC CAAAATAAT TCGTTTCTGC	960
	TGTCACTTTA ATACTCCTTA ACCTGAATAA ATTTTAGAAA CACTATGAAT TACATTCTTT	1020
45	tAGTGTTCCT TATGCAGTTG GACGCGTATG CGAACAACTG TATACCCTTT GTTCACTGCG	1080
	ATTTTAATCG CATTTCTTAT AACATTGTAG CGCCAGGAC ATTAATTTAC GTCCAGACC	1140
	CTTATCGTTT TCACTTCTAA GTAAGTCGAA CTATTTTGCT TTACAACAAG TGCGACTCTA	1200
50	AATACAGTTG GACACACATA CGAGCAACTG TATACCTTTT AATCAGTTTT CTATATTTTA	1260
	TTTATTATAT CTGTCTTAAT GATAAAAATT GTTACAAACA GTTTAACATA TTTAGCTACC	1320

55

ATGATACCAC TATGCTTGCh TATCTCTATA GCGCCATTGA TACACATTTT TAAATATCTA 1440
TACTGCCGTT AGAATTTTAT CATGTCTnAA TT 1472

5 (2) INFORMATION FOR SEQ ID NO: 584:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 787 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 584:

GGTAGTGAAT GGGGTTCAAG ACAACAGTAT TGGAAGTACG AATGAATCAC AGTTTTTtagg 60
 AAATTATATT GTAATTAAGC ACGCAGAAAA TGAGTATAGC TTAATAGCTC ATTTACACCA 120
 20 ATATTCAATC ATTGTGAATG AGGGGCAAAA TGTTAAATAT GGTGATATCA TTGGGAAGGT 180
 TGGGAATTCT GGCAATTCTA CTGAACCTCA TATACATTTT CAAGTAATGA ATGATAAGAA 240
 TATTGAAGCA TGTACATCTT TAAAAATTCTG ATTTATAAAT AATCGAGAAC TTATCAAAGG 300
 25 GGATGTGGTC TCGCGATTAC AAGCTGAATG ATGGCGATAC TTATAAAATC TCGACACTAT 360
 AAAAATGGTA TAGTGTGAG ATTTTCTTGC TTATTTAGTT AATTCAAAGT GCACGCCGGA 420
 TTCATTAGAA GTCGACGTAT TTTTGTGTTGT AATAGAGTAA CCGGTCATTG AAATTTTtagA 480
 30 TTCAATATCT GAAGCGGAAT TTGTAGATTC AGGATTATAG AAGCTACATT CATAAGTGTT 540
 ATCATCTTTC TTTTAAAGTA TAAACATACC TTTGGCTTTA ACTTCGACTT TAGTGTGTTT 600
 AATGTCAAAA GTTTGAGTGC TATTACTATA ATTAACACCA GCCCAAACCG ATTCATTATC 660
 35 TTTCACAACG GGGAAAGTCAT CTTCTTGCTT AACGACGThA CTTCATCTTT CTCTGTCTTA 720
 AAGACATCTT TAGATAAGCC TGGATACAAC ACATATCCAT ATTTATTGTC AGAATTAGAA 780
 40 TGCTTTT 787

(2) INFORMATION FOR SEQ ID NO: 585:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 830 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 585:

ATGCTAAATT AGGGGGAATA TTTACGATAA AGAGACCAGA AAAATAATAA AACACATGCT 60

55

TAAAATGAAG ACTATTTTTT ATTACAAGAA AATGTATCTA GTAAACTTAA AGTAGCAAGA 180
 CCTAATAAAT TTAATGCATG TTGTGCACCT TTTTACCTT GGCCAGCTTC GAAATGTTTG 240
 5 TAAGCAGCTA CACTTAAAAT GCCTATCGTT GATAGTGATG CAAGGCGAGA AATGTTTTTA 300
 TTGATAAAGC TAGCTGAGTA TAAAGCAGCA GTAGTTGCTT CTGCAATGCC GACGTATTTT 360
 10 ACAAGTTCTT TTTGCAAGCC AAAAGTATGT TCAAACAGTT CAATCATACC CTTATCTTCT 420
 TGCAATTTAG GTTTACTGGC TTGGTATAGC TCTTTCGCAA GTTTTAAATT CGTTGCGTAA 480
 CGCAAAATCA TATTTAATTC CTCCCAATAT TTGATTTTTT GTGAAAGATG ATTACTTTAT 540
 15 CATTTTTACC CGTTTCTATA AAAATGAATC AATTATGTAA CGTATGTGTA GTTTAGGAAT 600
 GTTTGCTATG GAAATATAAT TCTGTTCACT CAAAATGTAT GAAATTAATG TGTAGTTTTG 660
 TCGAGTTGCT CTTTTAATTT GGTTAGATTG TTTTITAGAG AAGCGGTACT ATTTTAAAGT 720
 20 GCATCAACAG ATTTACCTTC GTTTTGAGAC ATTGAGTTTA TTACAGCACG AAGTTCTGTT 780
 TCTAGTATGT CaGCGTCGct TTAGCATTAG AACTTAaTat TTatAcTCTT 830

(2) INFORMATION FOR SEQ ID NO: 586:

25

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 412 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 586:

35 TTAGGACGTT TTTACAATCA GTACAATGAG CTCATTGTTA TTAGTCCTTT AACGGCGTCT 60
 TTTAATGCTG GCGCTACATT TGGGCGATTT CATCATTTAA TTGATACTGA AACTTTAGCA 120
 AAATTAGAAC ATGAAAAAGG ACATTATTAT CAGAAGATGA TATGTGATGA CAATGTAGAA 180
 40 ATGATTTCTA TAAATAACAT ACCGAAATAT CCGAGAAATC ATAATGTATT AACTAATCAT 240
 GACTCATACG AATATTCATT GAATTTAGGA AGTAGTAATA GTTATTCAAA GTATGAGCTT 300
 ACCTTAGATG ATATTTATGT TGGTGCTACC TTTtAACAAA TTATATTTAT ATTCTAGCCm 360
 45 ACTAAATAAA AGGGkaCtat TTGaATCaAA CmATaTgTAT TAACCTTTTT TA 412

(2) INFORMATION FOR SEQ ID NO: 587:

50

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4709 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 587:

	TTCAGTTTGA AATTAATCAT ATAAATTTCT TATGGGAGGG TTGATATCTT AATGATTAAC	60
5	ATTATTTTCAG CTATAGGATC TATTGGAACA TTTATTATGG CTTTATTTTA TTTTGTATCA	120
	GTTTCAGTTC AACTTTTATCA AATGAAAATT AGCTTTCTGC CAGCTTTAGG TTTTAACCAA	180
	ATTTTATTAG AAAGGGAGGA GGATCAACTT AATATAATGA ATTCGGCAAC AGAAGAGCAT	240
10	CATCATAAAG ATTATATTAA ACTATATAAT TTAGGTGGCG GTGCTGCTAA TAAAATTGCA	300
	ATAGAGGTTT TATTGGGGAA GGATATAGTC ATTCAGAAAA AATACGTGCA TATTkTACCT	360
15	AGTAAAGAAG GGTACATGTT ACCAATTAAT AAAATGTGT ACGAAGAATT AGAAAGAACG	420
	ATTGAsAACA ATGGTCATGA AGCTGATTTG AATGTACGTA TGACTTATTA TCATAATGTA	480
	AGTCGCAAAC AACAGGAAGT TATATTAAAA GGTCAAATCG ACCGTTTTAA TACTTATAAT	540
20	AATAAAGAAA TTTATGATTT GCAGTTTATC TAAAAATTGA TTTAAGAGGG TAGTTGTTTA	600
	TTGCGAAAAA TATCATTCAA TTTAATGAA ATAATGGCGT CATTACTATA AAATATTACT	660
	TTATGTTGTA ATGCATTTTT CTATAAGATA GAACTAAAAG GAGGGGCAAA GATGCAAATT	720
25	AGACAAATAC ATCAACATGA CTTTGCTCAA GTGGACCACT TAATTAGAAC GGCATTTGAA	780
	AATAGTGAAC ATGGTTATGG TAATGAATCA GAGCTAGTAG ACCAAATTCG TCTAAGTGAT	840
	ACGTATGACA ATACCTTAGA ATTAGTAGCT GTTCTTCAAA ATGAAGTTGT AGGGCACGGT	900
30	TTACTAAGTG AAGTTTATCT TGATAACGAG GCACAACGGG AAATTGGATT AGTGTTAGCA	960
	CCTGTATCTG TTGATATTCA TCATCAAAAT AAAGGTATTG GGAAGCGATT GATTCAAGCA	1020
35	TTAGAACGAG AAGCAATATT AAAAGGATAT AATTTTATCA GTGTATTAGG ATGGCCGACG	1080
	TATTATGCCA ATCTAGGATA TCAACGCGCA AGTATGTACG ACATTTATCC ACCATATGAT	1140
	GGTATACCAG ACGAAGCGTT TTTAATTAAA GAATTAAAAG TGAACAGTTT AGCGGGAAAA	1200
40	ACAGGTACCA TAAATTACAC ATCTGCTTTT GAAAAAATAT GATTTCAAGC TAGGATTACA	1260
	TTAGGTAGAG TTCATATTAA TAATAAAAAA TGTTTGCAAT CAAATCGTAC GTTGTCGTTT	1320
	GTAATTCTTA AAATAGCAAT AAATAAAATG TTTGTTAGTA AAGTATTATT GTGGATAATA	1380
45	AAATATCGAT ACAAATTAAT TGCTATAATG CAaTTTTAGT GTATAATTCC ATTGACAGAG	1440
	ATTAAATATA TCTTTAAAGG GTATATAGTT AATATAAAAT GACTTTTTAA AAAGAGGGAA	1500
	TAAATGAAT ATGAAGAAAA AAGAAAAACA CGCAATTCGG AAAAAATCGA TTGGCGTGGC	1560
50	TTCAGTGCTT GTAGGTACGT TAATCGGTTT TGGACTACTC AGCAGTAAAG AAGCAGATGC	1620
	AAGTGAAAAT AGTGTTACGC AATCTGATAG CGCAAGTAAC GAAAGCAAAA GTAATGATTC	1680

55

	GTCAAACACT AATAATGGCG AAACGAGTGT GGCGCAAAT CCAGCACAAAC AGGAAACGAC	1800
	ACAATCATCA TCAACAAATG CAACTACGGA AGAAACGCCG GTAAGTGGTG AAGCTACTAC	1860
5	TACGACAACG AATCAAGCTA ATACACCGGC AACAACTCAA TCAAGCAATA CAAATGCGGA	1920
	GGAATTAGTG AATCAAACAA GTAATGAAAC GACTTCTAAT GATACTAATA CAGTATCATC	1980
10	TGTAAATTCA CCTCAAAATT CTACAAATGC GGAAAATGTT TCAACAACGC AAGATACTTC	2040
	AACTGAAGCA ACACCTTCAA ACAATGAATC AGCTCCACAG AGTACAGATG CAAGTAATAA	2100
	AGATGTAGTT AATCAAGCGG TTAATACAAG TGCGCCTAGA ATGAGAGCAT TTAGTTTAGC	2160
15	GGCAGTAGCT GCAGATGCAC CGGTAGCTGG CACAGATATT ACGAATCAGT TGACGAATGT	2220
	GACAGTTGGT ATTGACTCTG GTACGACTGT GTATCCGCAC CAAGCAGGTT ATGTCAAAC	2280
	GAATTATGGT TTTTCAGTGC CTAATTCTGC TGTAAAGGT GACACATTCA AAATAACTGT	2340
20	ACCTAAAGAA TTAAACTTAA ATGGTGTAAC TTCAACTGCT AAAGTGCCAC CAATTATGGC	2400
	TGGAGATCAA GTATTGGCAA ATGGTGTAAT CGATAGTGAT GGTAATGTTA TTTATACATT	2460
	TACAGACTAT GTAAATACTA AAGATGATGT AAAAGCAACT TTGACCATGC CCGCTTATAT	2520
25	TGACCCTGAA AATGTTAAAA AGACAGGTAA TGTGACATTG GCTACTGGCA TAGGTAGTAC	2580
	AACAGCAAAC AAAACAGTAT TAGTAGATTA TGAAAAATAT GGTAAGTTTT ATAACCTATC	2640
30	TATTAAAGGT ACAATTGACC AAATCGATAA AACAAATAAT ACGTATCGTC AGACAATTIA	2700
	TGTCAATCCA AGTGGAGATA ACGTTATTGC GCCGGTTTTA ACAGGTAATT TAAAACCAA	2760
	TACGGATAGT AATGCATTAA TAGATCAGCA AAATACAAGT ATTAAAGTAT ATAAAGTAGA	2820
35	TAATGCAGCT GATTTATCTG AAAGTTACTT TGTGAATCCA GAAAACTTTG AGGATGTCAC	2880
	TAATAGTGTG AATATTACAT TCCCAAATCC AAATCAATAT AAAGTAGAGT TTAATACGCC	2940
	TGATGATCAA ATTACAACAC CGTATATAGT AGTTGTTAAT GGTCATATTG ATCCGAATAG	3000
40	CAAAGGTGAT TTAGCTTTAC GTTCAACTTT ATATGGGTAT AACTCGAATA TAATTTGGCG	3060
	CTCTATGTCA TGGGACAACG AAGTAGCATT TAATAACGGA TCAGGTTCTG GTGACGGTAT	3120
	CGATAAACCA GTTGTTCCCTG AACAACTGA TGAGCCTGGT GAAATTGAAC CAATTCAGA	3180
45	GGATTGAGAT TCTGACCCAG GTTCAGATTC TGGCAGCGAT TCTAATTCAG ATAGCGGTTT	3240
	AGATTCCGGT AGTGATTCTA CATCAGATAG TGTTTCAGAT TCAGCGAGTG ATTCAGATTC	3300
50	AGCAAGTGAT TCAGACTCAG CGAGTGATTC AGATTGAGCA AGCGATTCCG ACTCAGCGAG	3360
	CGATTCCGAC TCAGACAATG ACTCGGATTC AGATAGCGAT TCTGACTCAG ACAGTGACTC	3420
	AGATTCCGAC AGTGACTCAG ATTCAGATAG CGATTCTGAC TCAGACAGTG ACTCGGATTC	3480
55		

CGATTCTGAC TCCGACAGTG ATTCCGACTC AGACAGCGAT TCAGATTCCG ACAGTGATTC 3600
 CGACTCAGAT AGCGATTCCG ACTCAGATAG CGACTCAGAT TCAGACAGCG ATTCAGATTC 3660
 5 AGACAGCGAT TCAGATTCAG ATAGCGATTG AGATTCCGAC AGTGACTCAG ATTCCGACAG 3720
 TGA CTGGAT TCAGATAGCG ATTCAGATTC CGACAGTGAC TCAGATTCCG ACAGTGACTC 3780
 AGACTCAGAC AGTGATTCCG ATTCAGCGAG TGATTCCGAT TCAGATAGTG ATTCCGACTC 3840
 10 CGACAGTGAC TCGGATTTCAG ATAGCGACTC AGACTCCGAT AGCGACTCGG ATTCAGATAG 3900
 CGATTCCGAC TCAGATAGCG ATTCAGAATC AGACAGCGAT TCAGATTCAG ACAGCGACTC 3960
 AGACAGTGAC TCAGATTCAG ATAGTGACTC GGATTTCAGCG AGTGATTTCAG ACTCAGGTAG 4020
 15 TGA CTCCGAT TCATCAAGTG ATTCCGACTC AGAAAGTGAT TCAAATAGCG ATTCCGAGTC 4080
 AGTTTCTAAC AATAATGTAG TTCCGCCTAA TTCACCTAAA AATGGTACTA ATGCTTCTAA 4140
 20 TAAAAATGAG GCTAAAGATA GTAAAGAACC ATTACCAGAT ACAGGTTCTG AAGATGAAGC 4200
 AAATACGTCA CTAATTTGGG GATTATTAGC ATCAATAGGT TCATTACTAC TTTTCAGAAG 4260
 AAAAAAGAA AATAAGATA AGAAATAAGT AATAATGATA TTAAATTAAT CATATGATTC 4320
 25 ATGAAGnAAc rCCTTAAAAG GTGGCTTTTT TACTTGATT TTCCAAATAT ATTGTTTGAA 4380
 TATAATTAAT AATTAATTCA TCAACAGTTA ATTATTTTAA AAAGGTAGAT GTTATATAAT 4440
 TTGGCTTGGC GAAAAAATAG GGTGTAAGGT AGGTGTGTTAA TTAGGGAAAA TTAAGGAGAA 4500
 30 AATACAGTTG AAAAATAAAT TGCTAGTTTT ATCATTGGGA GCATTATGTG TATCACAAAT 4560
 TTGGGAAAGT AATCGTGCGA GTGCAGTGGT TTCTGGGGAG AAGAATCCAT ATGTATCTAG 4620
 TCGTTGAAAC TGA CTAATAA TAAAAATAA TCTAGAACAG TAGAAGAGTA TAAGAAAAGA 4680
 35 TTGGATGATT TCAATATGGT CChTTCCCA 4709

(2) INFORMATION FOR SEQ ID NO: 588:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1554 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 588:

CTTTTTTAAt TAwCgGaAtA TTGtCaTgaT tAcAcTTCGt TAGGGTTTAc gTCGtAATTT 60
 50 CATTTaaTAA gCgCTTCacC ATTAAAtGTG gTAmCCTTTA ATTCGCCAGT AGAAACATCA 120
 CAGTAACTAA GCGCAATTC AGGTTGATTC ATAACAAAC TTAAATATA GTTATTTTGT 180

55

CGTCTAACCA TACCTTTTCGT TTGTTTCGGA TCTTCCATCT GTTCACAAAT AGCTACTTTA 300
TATCCATTAT TAACAAGTGT ATCTATATAA CTATCTGCAG AATGATACGG AACACCACAC 360
5 ATCGGAATTG GATTTTCTTT TTTAGCATCT CTTTtagTTA AAGTAATTTC AAGTAyACGT 420
GATGCCTCCT TGGCATCTTC ATAAACATT TCATAGAAAT CACCTAGTCT AAAAAATAAT 480
AAGCAATCTT GGTATTCTGA TTTTATTTTT AAATATTGCT GCATCATTGG TGTAACATTA 540
10 GACATATTAT TTCTTCACAA CCTTGTCTC TTTTAAAAT TTGTCTTTAC AATATATTCTG 600
TTTGTAAGyT TTTTAATTAT TAATTATTTA ACTTATACAT TTTAACATAC TTACTTTTAC 660
AAACCTATTTC ATACCATATA ATCACGAAGC ATCTTAAATG TATAAGAAAA CGCCTCAAAC 720
15 CTAATAAAAT GTGTCAATAG CATGTTTAgA ATTAAATTAA AATTCTAACA TTCAAGACAT 780
TTAATTAAGT AAGGGCGTTC AATATTAAAA TGAACAATGA CTCTGTTTGA AATCATATAT 840
20 CATAAAATTA TTTTATAAAC CTTTGAAGAA TACCACGTTT TTTTAGAGTA ATTAATAAGA 900
AATAACTTAT AATAGATCCG ATAGCACTTG aGACTATGaA CGTAATCATT AACGGTTTTAA 960
TGAAGAAGTC TTGAAGCCCA AGGaAATATG CTAATGGtAT aCaAATTAAA cTTCCgATGA 1020
25 CaCCAGTTCC aAGTACTTCa CCGACCGCGG CCaTAAATAT ATGTTTACGA TATnygTAAA 1080
ACATACTAGC CAATAAACT CCAATCATAC TACCCGGAAA TGCAAAAGst GTACCAGTAC 1140
CAAAAAGAAC TCTTAAATTT GATGATATAA GCGCTTGAGC TAATCCATAC CAAGGACCTA 1200
30 CTATGACCGC ACTTAATACA TTTACAAAAT GCTGTACTGG TGCTGCCTTA ACTGGTCCTA 1260
GAGGAATGAT GATAAACTG CTTAATACAA CATTtATTGC AATTAAAAGT GCAGTTATAG 1320
CCAGTTTTCT TGATTTcATA TGATTGTTCT CCTTTTTGTT TGTAATTAAT CACTATGCTT 1380
35 GGCTTTATTA TGGTCATTTA AACGTGTTTC CATTGTTGAT ACAAACATTT TCAATAATTG 1440
ATTGCTTCA TATTGTGAAG TTTGAACTG TTCAACTATG GGCAATGTAT TTATTTCTGC 1500
40 TTCTaACTC TGAATGGtAT GTTCCGACTG nTCCAGCGCA TTTGTTTCCC GnAA 1554

(2) INFORMATION FOR SEQ ID NO: 589:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 638 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 589:

TTTTGACGAA AAAAGTTGAT TTACAAATAT ATAAACGTTG TGATTTCAAT GTTTGTATAG 60

TATATACATG ACAGCAACTT GGGAAAAAAA GGAAGGTAAC GAAGGTTTAT TAACTGTTAC 180
 TGTTCCTGCA GAAAAAGTAA ACAAAGCrCT GAGAGATCCC CTCATAATTT CCCCAAAGCG 240
 5 TAACCATGTG TGAATAAAATT TTGAGCTAGT AGGGTTGCAG CCACGAGTAA GTCTTCCCTT 300
 GTTATTGTGT AGCCAGAATG CCGCAAAACT TCCATGCCTA AGCGAACTGT TGAGAGTACG 360
 TTTCGATTTT TGA CTGTGTGT AGCCTGGAAG TGCTTGTCCC AACCTTGTTT CTGAGCATGA 420
 10 ACGsCCGCAA GCCAACATGT TAGTTGAAGC ATCAGGGCGA TTAGCAGCAT GmTATCAAAA 480
 CGCTCTGAGC TGCTCGTTCT GCTATGGCGT AGGCcTAGTC CGTAGgCAGG ACTTTTCAAG 540
 TCTCGGAAGG yTTCTTCAAT CTGCATTGCG TTCGAATAGA TATTAACAAG TTGTTTGGGT 600
 15 GTTcGAATTk CAACArGTaA GTtAGtTGCT AGAnCCCA 638

(2) INFORMATION FOR SEQ ID NO: 590:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1242 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 590:

AAAATATTCC CGTACATTTT GATGTCTGTA GGGGCTTTTT TGA CTTTAGG ATTTGTcATT 60
 30 TTTTCAATTC ATAAAGGGAG ACGAACGAAA AATGAATCAG CACGTAAAAG TAACATTTGA 120
 TTTTACTAAT TATAATTACG GCACATATGA CTTAGCAGTA CCAGCATATT TACCGATAAA 180
 AACTTAATA GCTTTAGTAT TGGATAGTTT GGACATTTCa ATATTTGATG TCAATACACA 240
 35 AATTAAAGTG ATGACGAAAG GTCAATTACT TGTTGAAAAT GATCGACTCA TTGATTATCA 300
 AATCGCTGAT GGAGATATTT TGAAGTTACT ATAGGAGGGA AAATAGATGG TTAAAAATCA 360
 40 TAACCCTAAA AATGAAATGC aAGATATGTT AACGCCTTTA GATGCTGAAG AAGCAGCTAA 420
 AACAAAATTA CGCTTAGATA TGAGAGAGAT TCCTAAGTCT TCAATTAAAC CAGAACATTT 480
 TCATTTAATG TACTTATTAG AACAAcATTC TCCATATTTT ATAGATGCTG AATTAACCTGA 540
 45 ACTACGTGAC aGTTTCCaAA TACATtATGA CATTAAATGac AATCATAcAC CTTTTGATAA 600
 TATTAAATCA TTTACTAAAA ATGAAAAATT ACgTTACTTA CTCAATATCA AAAATTTAGA 660
 AGAagTAAAT CGTACACGCT ACACATTTGT GTTGGCACCA GATGAATTAT TTTTCACAAG 720
 50 AGATGGATTA CCCATTGCTA AAACAAGAGG GTTACAAAAT GTTGTGTGATC CATTACCTGT 780
 GTCAGAAGCT GAATTTTTAA CAAGATATAA AGCGCTGGTT ATCTGTGCAT TCAATGAGAA 840

55

AACTAAAGTT ATTGAAGCGG CAACGTTAGA TTTACTAACG GCATTTTATAG ATGAACAGTA 960
 TCAGAAACAA GAACAAGATT ATAGTCAAAA TTATGCATAT GTACGCAAAG TAGGACATAC 1020
 5 CGTTTTCAAA TGGGTGCTA TCGGTATGAC AACGTTAAGT GTTTTATTAA TTGCATTCTT 1080
 AGCCTTTTTA TATTTTTCAG TAATGAAGCA TAATGAGCGC ATTGAAAAAG GATACCAAGC 1140
 10 ATTTGTAAAG GATGTATAC GCAAGTACTA AATACGTATG ATGATTTAGA TGGTAAAAAT 1200
 TgATAAGAG GCACTTTACA TTTATGCCAA AAGTTATATC CA 1242

(2) INFORMATION FOR SEQ ID NO: 591:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 744 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 591:

TTCCAGATAG AGCCTTACTT GCTGCCGCTG AATTGTCTGA TCGTTACATC ACTGATCGTT 60
 25 TTTTACCAG ATAAAGCGAT TGATTAGTT GACCAAGCAT GTGCAACAAT TCGTACGGAA 120
 ATGGGATCAA ATCCAATGA ATTGGATCAA GTTAATAGAC GTGTCATGCA ATTAGAAAT 180
 30 GAAGAAAGCG CACTTAAAAA TGAATCTGAC AATGCGAGCA AACAGAGATT ACAAGAACTA 240
 CAAGAAGAGC TTGCCAATGA AAAAGAGAAA CAAGCAGCAC TTCAATCTCG TGTAGAATCA 300
 GAAAAAGAAA AAATAGCAAA TTTACAAGAA AAACGTGCGC AACTAGATGA AAGTAGACAA 360
 35 GCGTTGGAAG ATGCACAAAC AAATAACAAT TTAGAAAAAG CTGCTGAACT ACAATATGGA 420
 ACAATTCCTC AATTGGAAAA AGAATTAGA GAATTAGAGG ATAATTTCCTA AGATGAGCAA 480
 GGTGAAGATA CAGATCGAAT GATTCGTGAA GTTGTAACAG ACGAAGAAAT TGGCGATATT 540
 40 GTCAGCCAAT GGACAGGCAT ACCAGTTTCA AAATTAGTTG AACAGAACG TGAAAAATTA 600
 CTTCACTTAA GTGACATCTT GCATAAACGT GTTGTAGGTC AAGATAAAGC GGTGACCTG 660
 GTTTCAGATG CAGTAGTTAG AGCAAGAGCA GGTATTAAAG TnCAAACAGA CCTATTGGTA 720
 45 GTTTCtnATT CCTAGGTCCn ACTG 744

(2) INFORMATION FOR SEQ ID NO: 592:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1449 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 592:

	TTGTTATAGT GTCTGAATCT GTGTAAGTCT TGCCCATGTG TTCTGAGTAA AGCACCCACT	60
5	GTTTATTTAC TTTCGTTGT AGTCTAGCTT CGTGTAGTAG TTTGTTTAAAC TTTTGTGCTG	120
	ATATACCGTA GTCTGCCGCG ATTTGAGTTG TGGCTAATGT GCCAGTTGAT TTTAAGATTT	180
	CATCAACATA ATCTGCTTTG GGTTTTAGCT CTCCAATTC TTGTTGTAAA AGTAAGTTT	240
10	GCTCTTTTTT TTTCTTATAC TCAGTCAACA CTGTAATGAT GTAGTCTGGA TCTTTTAATG	300
	TTTGTTCAAT TACATTGTCT GTTGCCTATA TACCGTGTTC GCGAATAGCT GGTaGGACAT	360
	CTGATGTTAC CCAGCGTTTG AATTTTCTAG CGGTTTCTCT AATTTTTTCG TTTTGTCTTT	420
15	GTTTAGAAGC ATCGAAGATT AGACTGTATA ATCCTGATTC GTTGATAATG ATCATATTTT	480
	TGTTTTGACC TGATGCACTA AATTGGTGCG TCAGTTTGTC CTCGCTATCA ACATGATTTT	540
20	TAATGGCATT GTCTGATCTT GCATATCCTA AAATCTCAGC AATATCTTTT CCTACAAAAT	600
	AAGGTTGCTT TTCAATTTCT ACTGTTCTTA CTGGTAGCTC TTTAAAATTA AATGTTTGTA	660
	ATGCTTGCAT TTGAGTATCC TCCTTTTTTC TCAACACCCA CATTGAGCAG ACGGTTATCG	720
25	CAATGACTAT CGAATGTATT TAAACGCGGC TCATATCATC GCCAGCTCTC GCTCACATCT	780
	GCTCAATGTG GATGTTGATA AGCGTGGTTA TATTAAGAAG TGAATGTTAC TGATTCACCT	840
	TCCGCCACTC TGTTAAATCA GTAACCTTGT TATCGCTTTC AACACCGTTA AGCTTGTCTA	900
30	ACGCTTTCAC TACTTTTTGG AACTCTTTGA TAGCACTtCG TAGCTTTTTTA GTAATTTTAT	960
	CTTCTACCAT TTCCAAACCA GCAAAATGCGT CTTCGTTATT CATGCTTAGA TGTTTGTGTA	1020
	AAAGATCTCG AGTGATCTTT ATTTCTTTAA GTGATTTATC ATAAGCTTCA ATTTGTCCTG	1080
35	AAAGGTTATG ATATTTTAGT TGTAGTTTTA CTAATTTTAA TGATTGGTCT TGCATTTGTT	1140
	ATGTCTCCTT TAAGATGTTT GTTTGCGTTT CGTGTAATTT GTGGGTAAAA AAATATCTCC	1200
40	AATATTTTCG TCAAAAAAAT CAGCGATAAT AACATCTCA TCATTCTTAA ATTGATGCTT	1260
	TCCTAATTCT TTTAAACGAT AACCTTCAGT TGATATATTC AAGAGGTTTG CTAAATCTTC	1320
	TTGAGTACAC TTTCTTTCTT TTCTCAACTT TATTAAATTC CATTGCATGT TGTCACCTCC	1380
45	CGCTTTACAA AACCTACTAT ACACGATACG GTACTTGnGT CAACATAAAA GTTTGCTTTT	1440
	CGTGTATTT	1449

(2) INFORMATION FOR SEQ ID NO: 593:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1170 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 593:

5 ACCAAAAACA CTTTCTAAAG AAATTGTTTC AGAACCATTT GCTAAAAACC AATTACGTGA 60
 AAAAGCTCGT CTAAC TAACA TTAAAGGTTT AGAGATTCCA AGAGTTTCAT AACTTTTAGA 120
 CGATGATGAT TTCATTACAG ACGTAGAAAC AGCAAAAGAA TTAAAGCAA AAGGTGATAC 180
 10 AGTCAAGTTC ACTACTAATA AATTCAAAGT ATTTGCTGCA ATTCAGATA CTGTAATTCA 240
 TGGATCAGAT GTAGATTAG TAACTGGGT TGAACGCA CTACAATCAG GATTAGCAGC 300
 TAAAGAGCGT AAAGATGCCT TAGCAGTAAG TCCTAAATCT GGATTAGAAC ACATGTCATT 360
 15 TTATAATGGA TCTGTTAAAG AAGTTGATGG AGCAGACATG TATGATGCTA TTATTAACGC 420
 TTTAGCAGAT TTACATGAAG ATTATCGTGA TAACGCAACA ATTTATATGC GATATGCAGA 480
 20 TTATGTCAA ATTATTAGTG TTCTTTCAA TGAACAACA AATTTCTTTG ACACACCAGC 540
 AGAAAAAGTA TTTGGCAAAC CAGTAGTATT TACAGATGCA GCAGTTAAAC CTATTGTGGG 600
 AGATTTCAAT TATTTTGAA TTAACATGA TGAACAAC TATGACACTG ATAAAGATGT 660
 25 TAAAAAGGC GAATATTGT TTGTATTAAC AGCATGGTAT GATCAGCAAC GTACATTAGA 720
 CAGTGCATTC AGAATTGCAA AAGCAAAAGA AAATACAGGt CCATTACCCA GCTAAGCCCC 780
 AAAAGGTAA TGTACAGCT AAGGCTAAAT CAGCTGTAAT ATCAGCCGAA TAGGGGTGAT 840
 30 GAAATGAGTT TAGAAGAAK TAAATTGTGG TTGAGAATTG ACTATAATTT CGAAAATGAT 900
 TTAATTGAAG GTCTCATTCA ATCGGCTAAG TCTGAATTAC TATTAAGTGG GGThCCAGAT 960
 TATGACAAAG ATGACTTGGA ATACCGCTT TTTGTACAG CGATTAAATA TATCATTGCA 1020
 35 AGAGATTATG AAAGTCGTGG ATACTCAAAT GACCAATCTA GAAGCAAGGT GTTTAATGAA 1080
 AAAGGATTGC AAAAAATGAn TTTGAAATTA AAAAGTGGT AGGTGATTTT TAAATGGAAT 1140
 40 TTAATGAATT TAAAGATCGC GCGTATnTTT 1170

(2) INFORMATION FOR SEQ ID NO: 594:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1393 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 594:

TGAAGGGTGT TTTGTTTATG GCCAATTGCT GTGTTATTAA AnCGAATGTT TCGAATGGAA 60

ATCATTCCGG AAGAGGACAT GTCTTCAGCT TCTGATTGTA AAGCACAGTT TTATCGCGGT 180
 ACTTTGCAGC GCTTTTACCA AtCnTTGCAC GCAGAAAAGC TTACACCTTA TGTTATGTCT 240
 5 TATGACGATA TCATTTTATT TTGTAAAGAA AACAAATATCT CTGAAGTAGT GACTGCGGGT 300
 GATATTATGA GTTATCATCT TGAAGAATAT GATATTTTAC ATCAACGTTT TTTATTCAAT 360
 GAAGCACGCA TTGCCGTTAC TTTGATACGT GGGAAATCATT ACTTTAAAGC GAGTAAAACA 420
 10 ATGAATCAAC AAGGGGAGCC ATACAATGTT TTTACTAGTT TCTATAAAAA ATGGCGACCT 480
 TACTTGAGGC ATAGAGACGT ATATCACTAT GATTTAAAT CATTGAAAA CTTTGTCTT 540
 GCATCACCTG ATGATTTAGT GTTGATGAC ATAGCATTG GATCCTCACA aATAATTGAA 600
 15 CAGatAAATG GCAACATTTT TTAGATCAAG ATATACAGAA TTACGAAAGC GGAAGAGACT 660
 ATTTACCTGA AGTATTAACA AGTCAGCTAA GTGTTGCTTT AGCATATGGA TTATTAGATA 720
 20 TTATTGAAAT TTTAATGAT TTATTGGCGC GTTATGATGA AGATGAGGCA AACTATGAAG 780
 CATTTATACG TGAACCTATT TTTAGAGAAT TTTATTATGT GTTAATGACA CAGTATCCTG 840
 AAACCTCATA CCAAGCTTTC AAACCTAAAT ATCGACAGAT AAAATGGTCG CAAAATGAAG 900
 25 CGGATTTTAA TGCATGGTGC GAAGGGCAAA CAGGATTTCC AATCATTGAT GCAGCAATAA 960
 TGGAAATTGAC ACAAACCTGGT TTTATGCATA ATCGAATGAG AATGGTTGTG TCGCAATTTT 1020
 TAACCAAAGA TTTATTTATA GATTGGACAT GGGGAGAAAA ATTCTTTAGA AAGCACCTTA 1080
 30 TTGACTATGA TGCAGCATCA AATATTCATG GATGGCAATG GTCTGCTTCT ACAGGTACGG 1140
 ATGCAGTGCC GTATTTTAGA ATGTTTAATC CAATAAGACA GAGTGAACGC TTTGATGCTA 1200
 AAGCTTTGTA TATCAAAACA TATCTTCCGA TTTTAAATCA AATTGATGCA AAATATTTGC 1260
 35 ATGATACACA ACGCAATGAG TCCAACCTTT TTGAACAGGG GATTGAATTA GGTAGTCAIT 1320
 ATCCAAGACA AATGGTAGAT CATCAAGAAA AACGTACACA AGTTTTAGCT ACATTTAAAG 1380
 40 CGCTAGACTA ATT 1393

(2) INFORMATION FOR SEQ ID NO: 595:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1484 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 595:

TTAACCTTGAA AGATTTCCCC GACATCTATA ACTTGTATAA TGTGTATATG TCGTTTTTAG 60

55

	AAAAATTAAA AAAAGAATAT CTTGATTGGT ATAATCAGAC CTTAGAGTTT TCTAATTTAT	180
	CAACAATGT AGTAAGWATA GATACTCCTT TTAAAGATAA TTCTTTAGAT AATTTAATAA	240
5	TTTACGCTTT ATACGATCAG TCCAGAGACA TGATTACACT GACAGATGAC GGCTATACTA	300
	TATcTGATTT AGAAAATAAT GGTATTTCTT TAAATAAATC AAAAAAACGT AAAAAGATTT	360
	TTGAAGAGCA CCTTTCAGCT TACGGTATTA AATATAACGA TAAAACTCAC GAAATTTTTG	420
10	TTCAAATAA CTTTAAAAAT TTTAATAAAT CGAAACATAA TTTATTACAG TGCCTTATAT	480
	TTGTTAATGA TATGTAATA CTTTCTAATC CTAAGTCACA GAACATATTT ACAGAAGATG	540
	TTGCAAACAA ATTGGATGAA CATAACATTT ATTACGGAAG AGATTTACCT ATTATAGGAA	600
15	GCAGTGGTGT TGTTCATAAT TTCGACTTTT TTATTAGCGC TAAGAAAAAT CAAAAAGAAA	660
	AATTTATCAA TGCTATTTCT AACCCATAA ATTCTATGAT TATTAAGTCG AAAATAACGG	720
20	ATGCTATGCA AGCAAAAAAA ATAAAAAGAC ACAGGCAAAA TGAGTTTATT TTTATTTTAA	780
	ATGACTCAAA AAAAGAAATA AATGAACATA ATAAAAATCT TCTTCATGAA AACTATATTA	840
	GTACAATAGA TTATAGCGAA TTAGATGAAA AGATAGGTTT ATTGATTTAA TATATATAGA	900
25	CGTGATAATA TCAATGTTTA TATTAAATGA AACGAATTTG AAAATTTCTGA AACAGCTTAA	960
	GATAGCAAAT TGAATAGCCT TATTGATAAT GCAGAATCGT CTACACTTAG TTGAACAAAT	1020
	TCTATGAGAA TAGATATTGT TAAACTATTT GGGTAGGCGA TTTTATTTTG ACAAGAGTCA	1080
30	GAAGATTATT TAGTTCATAG TTTAAGTTAC AAATGGCTAA ACTTTATGAA AATGGTATGA	1140
	CTAGAATCGA AATTATACTC GAATACGATT TAACACTCTC AATCTTCTCG AACTGAATAA	1200
	ATACTGGGTC ATTCAATCAT CAAGACAAC TAATAAGTGA TGATAAATGT TCATACCATT	1260
35	TTGCAAATAT AACCAGATAC TCTCAACAAA ATGTAGAGTT CAAAATTAAA AATGTCCCTA	1320
	ATTGGCACAT TGCTGTATGA TAATAAATTC AACTTAGATT TCAACAATAT AACTAATGAT	1380
40	GCTTATATTT ACCGCGATAT CAAACCAAAC AGCCCATCCA CAAACAAATT AGAATATATC	1440
	mTGTTAATAA TAGTAAATGA TATATCTTCC AAAGACCGCA AACT	1484

(2) INFORMATION FOR SEQ ID NO: 596:

- 45 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1435 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 596:

55

CTTTATCaTT AAATGCCaTT TTGaATCCTT CTTGAATTTG TTTGATTTC TTTTCAGTAT 120
 TATTAACCTT GTTCAAGACA GTGCCATTAA TTTTCTTCTT GAGTGGACCA ACCTCATCTT 180
 5 TATTAGTTGA TTCATGAATC GTTAATCCAA TGTGTGGCTG TATTCTATAA CCATCATTCG 240
 CTATAGTTGA AACATATTGT GaTAATTGTA ATGGTGTATA GGTATCATAT TGACCAATTG 300
 ATAAATCTAG ATAATTACCT GGATTATTTG TTAATGGTTC GATTTGACCT CTTGTTTCAT 360
 10 TTGGTAAATC TATCCCTGTT TTCACACCTA AGCCTACTTG ATTTAATCCT CTTCTTAGCT 420
 TTTGGGCAGG TGAACCTTATG TCTGAAGGTA AAGCCATACC AGAATAATAA GGGTCTCCCG 480
 CTAATTTTAA TGCTGTTTTA AACATATATA CGTTTGATGA ATGCATCAAA GCTTGCTTAT 540
 15 CATTAATAGT TACATGCCCG TTTTATTGA AGTATGATCG TTTTGTCAA CCACCTTGA 600
 AATGTAATGG TTCATCGACC ATTGTTTCTC CAACTTTGAT AGCTTTATTC TGATAACCGG 660
 20 CTAATAATGT TCCACCTTTT ACAGAAGATC CAACCGCAA TTGAGAAGTA AACGTACCAA 720
 TGTCATAATC AGTCATTTTA CCACTCTTAT TAATCTGCTT TCCGGCAAGC GCAAGAATGT 780
 CTCCATTTTT AGGATTTTGT ACAACCATCA TTGcATTATC CATATCTTTG GCACCTTGAC 840
 25 TGCGaSTTCT TAATTTGTTT ATCTAATAAT GCTTCTACTT CTTTTTGAAG ATCTATATCG 900
 ATCGTTAATT TCAAATCTTG ACCGCGAGCG CCAGGrTTTA ACACCTTCTGA AGATGTAAC 960
 TTACCAGATT TGTCCGTTGT GTATTTCAAT TCTTTCTTCT TACCACGCAA TACATCTTCA 1020
 30 TATTGATATT CTAGGTAAGA TTTTCCAACA CGATCATTGC GTGAATATCC TTTGGATAAG 1080
 TAATGTTCTG TCAATTCCTT TGGaATACCT TCAGCAGGTG TCGATACATC TCCGAATATA 1140
 CCTCTTAAAG TATCGCCATA TGGATATTTT CTATCCCAAT CCATAGACGT GTTAACACCT 1200
 35 GGTAATTTGG AAAGTTGCTG AGAACTGCT GCATACTCTT TTCACTGAC ATCTTCATTT 1260
 TTTATCATTT GTGGATCTAA AACTGTTCTT GGCATTATC TCTCGAAAA TAGCTAAAAC 1320
 40 TTGGTAAATC TTTAGAAGAC AATTCATCTA ATTGTGGATT TTCCGGATT CGGATAACAG 1380
 TTGGTnTATC CATAATGGAT CCTGGTTTTAA TACCTTCCAn CnGGCGAACA TAGCC 1435

(2) INFORMATION FOR SEQ ID NO: 597:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 516 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 597:

55

5 AGAAAAAGAT TTAGAAAAGA TTACCATTCTG CGACATAACA ACACGCGCTG ATATCAATAG 120
 AGGGACATTT TACTTACATT ACGAGGATAA ATATATGTTA CTCGCAGATA TGGAAGATGA 180
 10 GTATATTTCA GAACTAACGA CATATACTCA ATTTGATTTG TTACGCGGTT CTTCAATTGA 240
 AGACATTGCG AATACTTTTG TAAATAATAT ACTCAAAAT ATTTTCCAAC ATATTCATGA 300
 TAATTTAGAG TTTTATCATA CTATTTTACA ATTGGAACGC ACGAGTCAAT TAGAATTGAA 360
 AATCAACGAA CATATTAAAA ATAATATGCA ACGTTACATT AGTATCAATC ATTCTATCGG 420
 AGGCGTTCCA GAGaTGTATT TTTACAGTTA CGTTTCCGGA GCAACAATTT CaATTAnnTA 480
 15 AATACTGGGG TAATGGACAA ACAGCCCAATT TTCAGT 516

(2) INFORMATION FOR SEQ ID NO: 598:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1955 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 598:

AGTACTTTAC CATGAGGTAA GACTTTATTC CCTGCTTTTA TAGATTGTGC TTTTATTTGT 60
 30 GTGCCAGAAC CAATAGtAct GGTcTAATGA TTTTGCAC TC ACATTATCAA TAGCTTTTTG 120
 TTTGTCTTGA CCTTCAACAT CTGGCACTTT ACTATACTCT GCATTAGATG TGTCACTTTT 180
 TGATTTACCT ACATTTAAAT ATTTCAAAGT ATTTTCCATT ATTGGTTTAA ACGCTTTACT 240
 35 AACACCTAAT TCATAAGCTT CTTGGTCATT TTTTGTGCC AAGCTCATA CAGCGTATAC 300
 AATAACTTTA GGATTTTTCT TCGGCGCGTC ACCCATAAAA CTTACAAAAT ATGGGTTTGG 360
 ACCTTTAACG TATCCACCAC CATTAGGTGC AGCGACTTGT GCTGTACCAG TCTTACCTTC 420
 40 GACCTCATAA CCATCAATAC GATAGTTTGC AGCGTGACTC TTCTTACTAT TCACAAC TAA 480
 ATCCAATTGC TTTTCAACTT TTCAGCAGT ATCTTTTGT ATTGGTTTGC CTGCGATTG 540
 TTTTGCCTT TTATAAAATT GTCTTTTACT AACAGGATTT TCAACGCTAT TCACAAACCA 600
 45 TGGTTTTAAC ATATTACCAT CATTAAAGAA CGCTGATTGC GCTTGTAACA TTTGAACAGG 660
 TGTACTGTT GTCGATTGAC CAAATGATGA CGTTTTTTG TGCAACTCAT TACTCCATCC 720
 AATTTGaCCA GGTGCTTCTC CATCAAACAT ACCTTTAGTT GaTTTTCCaA ATCCaAATCG 780
 50 TTCATACCAA GATTTCAATT TGTCTGCACC AACTAAATCT TGTAATGCA TCATCAATGT 840
 ATTAGATGAA TAAGTAAATC CGAGTGACAT TGGGATTTCA CCCCACCGA CTCTATTCCA 900

55

AAAAGCACCT TCTTGAATAG CAGCTGCTAA CCCATATGAT TTAAATGTTG ATCCAGGCTC 1020
 GTATGTGTTT TGATAAAGGT CATTGCCCCA CTTTTTACCA AAGTCTTTAC CAGTTTCAGG 1080
 5 ATTAATGTT GGTGCTGAC TGTATGCTAA AATTTCTCCA GTTTTGGCAT CCATGACAAC 1140
 CGCAAATAAA TCTTTCGGCT GGTATCTTTC AACCATGCCA TCTAAAGCTT CTTCAACAAA 1200
 TACTTGAATA TTTGAATCGA TTGTTAAATG GACATCATCA CCACGTTTAG GCTGCTTCTC 1260
 10 TTTTTTAGTA TTTGGTGCGA TATATCCCCA AATATCATGA ATATATCTCA ATGATCCTTT 1320
 AGATCCACTT AAATAACTAT CAAAAATCTT TTCAACTCCA AGTGCACCTT TAAGTTCACC 1380
 15 AGTATCCGGA TTTTCTGAG CTCTACCAAT TAAGTGTGAT GCAAAATTGC CATTGGGATA 1440
 AAAGCGTTCT GTTTCAGGCA ATAAAGAAAT ACCAGGCAAA TTCATTTTCT CTATTTTCAA 1500
 TTTGTCCTGA TACGTAAAT TTGTTCTTTT GCGTCCAAAT TCAATTTGGA AAGCTTCTTT 1560
 20 TTGACTAAGT CTCTTTTCAA TTTCTCTGG CTTCATATTA ATGACTGTAG ATAATTTCTT 1620
 TGCAGTCTCT TTTTATCAA CTACATGCCT AGGTTTTTTA GAATTGGCAC TCGCCTTTTT 1680
 ATCTATTACT GCAACAAGTT TATATCTTTC TACATCTTCT GCTAGCACTT TACCATKACG 1740
 25 ATCATATATC TTTCTCGTT CTGGTTGTTG TGCATTCTTA ACTAAATACT TTTCAATTGC 1800
 CTTCATGACT AAATCTTGAC CATTAGAATG TCCAGTAATC ATGATATATG AAATTCTTAA 1860
 AACCAATATA AAAAAGAGCA GTCCGAATAA ACCAACAAGT AGGACTGCCC CTATTTTATT 1920
 30 TTTTTTAATT TTAATTTTTT TGGTCGCCAT TACTA 1955

(2) INFORMATION FOR SEQ ID NO: 599:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 704 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 599:

TAGCAAGAGT AGGTGATAAC AGTTCTTTTA ACGTGTGTTG TGCTTGCTCA TTTCTGGCT 60
 45 CATTGTAGAC TAAAATATAA TAAGCGTTTT CAAATAAGTG CTTTTAGCA TTTAGTACAC 120
 CAGATTATG ACTACCAGCC ATTGGATGAC CACTGACTAA ATGAATATTA TGCTTTAATA 180
 AATTGCATTC GTGTTGCTGT ATCATTGCTT TAGTACTACC AGTATCAGAA ACAATAACAC 240
 50 CAGGTTTAGT TGGCATATCT ATAAGCTCGC TAAGATATTT ATTTGTGATA GCAACAGGTG 300
 TTGCATAAAT AATTACATCG GCTTTTTTAA TAGCTTCACT ATAATTTAA CATTTTTCAT 360

55

TATTAGTATT AGGGTTATGG TATTTTATAT TGCTAGCAAG ACTTCCACCA ATTAATCCAA 480
 GCCCAACAAA TAAACTGTG GTCATATAAA TCACCTTATT TCGAAATTTT CAGAATAATA 540
 ACATTGTAAA TGAGCTGTTG ACACAGTGCA ATAGTAAATA AAAATCGATA ATAGCATTAA 600
 TAGATAAACG GAGATAAATC ATCTACAATA AAGAGTATAG TAACACAATG GCAACGGAGG 660
 GGTAATCAA TGGAAACCAAT ATTAGAAATG ATCAAACAT TAAC 704

(2) INFORMATION FOR SEQ ID NO: 600:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1158 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 600:

CCTCTAAAGG ATCACGAACC TTTTTCATCG TGACTAAGCC TATAATACCC TTAAACCTAT 60
 TATTATTAAC TTTTACCTCT GTGTATTCCC TATCAATTAG GCGACGCCAG TGACGTTTAT 120
 CTATATACTT TACTTTCACA GTCACCAACT CCTGTGCATT ATTATATAAA ACTACATAAA 180
 TGATGTCATG TCATGATACA GTTTTATAGT TTTTGTTAAT CATATGAAAA TATTTATTAT 240
 TATTTTACTA CAACTCGCTT CAATTTACTT AAAATAGACA ATATTAATTA GATAGTACAC 300
 ACATTTCTTC ATAAAAGTGA TTTTCAAAA ATATAATAA CACACTCTTA TCGTTTTCAA 360
 AATCATTTAA TGCTATTTTC ATTAAAAACA GCTGAAGCAT CAAATCTATT CTGATTCAAT 420
 CAAGAATACA TATAAAAATG AAGTGACTCA AAGGTTTATT AACAACCTTC AAACCACTTC 480
 ATTGATCACT TTTATTTTAA AGCATATTTT TCGATTACTG ATTTAAGATG CGGATATTGC 540
 GTAATTAATT CAGATTGCTT AAACAATTCA AACTGCTTAA ACTCAAATGC CGGTTGACAC 600
 ATACAACCTA CTAAACTAAA AGTATTTGAT ATTTTCGATTG AAGAAGCAAA AATTGTTCCCT 660
 TTAGGCACTA CATATTGCAA TACATCTCCA TTTTGGATAT CAGTACCCAA TGTTGCAGTC 720
 GTATATTCCC CATCCGATT TATCATATGA ATTGTTAGAG AATCGCCAGC ATGATAGTAC 780
 CATACTTCAT CAGCATCAAT TCGATGAAAA TGCGAAATAT TGTCATCTGT AAGTAAAAAA 840
 TAAATACTAC TAAACGGCGC TCTGCGTCCA TCTTCAATA CTTCTCGAAT TGTCTCTCTA 900
 TAGAAACCAC CTTCAAGATG TGATTCAAGT TGCAATTcAT CaATCCATTG TTCTGCTGAT 960
 TTCATTATTT CAAATCCACA TTATGGAATA CGTTTGTAC ATCTTCTAAA TCTTCTAATG 1020
 CATCGATTAA TTTTTCAAAT GTTACTTGGT CCGCTTCAGA AAGTTCAATA TCTGTTTGAG 1080

GAACGACTGC AAATTGAT

1158

(2) INFORMATION FOR SEQ ID NO: 601:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 827 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 601:

TGATACTTAT	TTTTCAATAT	TATTTGGAAT	AAATCTTnT	AATTGTGGTG	TGAAATAAAC	60
ACCGAAATTC	TTTTCAGATT	CATCTTTTcC	ACTACCTGCA	TATAAAATAC	CAATCAGTyC	120
ATGTTTtagAA	TTCAATACAG	GAGATCCTGA	ATTACCTGGT	TGTGCATACG	CATCAAATTC	180
CATAAACGTT	CCACTGATAT	GGTTAATCGT	TCCTGTCGAT	TCAAACATTT	TATATTTTGT	240
TTGTGCACCC	TTTGGATAAC	CAATAACAGA	AATTCTATCT	TTCACTTTTG	CTCCGTCTGC	300
AAATTTTGTA	TAAC TAACGT	TCTTATTAAA	ATTCAAACCT	TCTGTACTTG	TTTCATGAAC	360
ATGAACTATC	GCAAGGTCTT	CTTTTCCGGG	ATATTCTACA	ATGTCTTTAA	CGTCGTAGTT	420
TCCTCCGCCT	TTACCTTTAC	TCGAATGATG	TGCTGATACT	CTATTTTTTAA	AAATATCATT	480
ACTTTTtagCG	ATATGTTTGT	TAGTTACGAT	TGTATTTTTTA	CCAACAAC TA	CACCAGTACC	540
ACCCACAAAT	GCTACCACTG	aATTGTATGG	TTCTTAGTT	GCATCGGTAA	TTTCTTTGAC	600
ATTCTTTTCT	GCTTTGGCAA	TTGAATGArG	CTGATTAGAA	ATATTTTCAG	CAAAACCAAG	660
AGATGTTAAA	ATAGTTAAAG	CAGTTAAACC	TTTAACCATT	ACATTTTTTAT	TCATATTTTA	720
CCTCCTTCCA	AATTTATTTT	nATAGTTAAT	GTATCTAGAT	ACACATTTAC	TATATACATT	780
TTAAATAAAA	ATATTnATTA	AATATTAATT	AACCACTTAT	naACGTT		827

(2) INFORMATION FOR SEQ ID NO: 602:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 796 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 602:

nGGnACGTAA AGTATTTATG TTAAATGGCA nAAATATATG AATGATGAGT TACCTGGATT 60
ACCAATGTTT CAAGGTAAAT CGATAACTAT TGTTAACGAT AAAGTACGAA ACTTAGACAT 120

CGACGATTTT GTCGAGGTTT ACTTATATGG GGGGATGTTA TGAATAATGT ATTGTTAGAG 240
 GTTAAAGATT TAGAAACATC ATTAAAAATA AATAATGAAT GGTTAGCAAC TGTTGAAAAT 300
 5 ATTTCTTTTG AATTATCTAA AGGAGAAGTT TTGGGTATAG TAGGGGAATC TGGTTGCGGT 360
 AAGTCCATAT TAAGTAAGTC AATTATTAAA TTATTACCAG AAAAGATATC TAAACTAAGT 420
 10 AATGGAGAAG TTATATTTGA TGGTAAACGA ATCGATACGC TCAATGAGAA GCAATTGCTA 480
 GATATTCGAG GAAATGATAT TGCTATGATT TTTCAAGAAC CTATGACTGC TTTAAATCCT 540
 GTATTTACCA TAAAAATCA ACTTGTGGAA TCTATAAAAT CACATAAAAA AATTTCTAAA 600
 15 AAAGAAGCAA ATAAATTAGC AAAAGATTTA CTAAAAAAG TTGGAATTGC TAGACAAGAT 660
 GAAATATTAA ATAGCTATCC TCATCAATTA TCTGGTGGTA TGAGACAAAG AGTAATGATT 720
 20 GCAATGGCCA TTTCATGTTT TCcTAAATTA TTAATTGCTG aTGAACCTAC AACAGCATTG 780
 GATGTCACGA TTCAAG 796

(2) INFORMATION FOR SEQ ID NO: 603:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1618 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 603:

GGATGTGTTT ATCCCAAATT GaAAcAgAaA TTTATGATCA AAtGATGTGG ATTGGTTTTTA 60
 35 AaCCATTCAa AATTaCCAAT ATTAmACAAG AATCTGAAGA CATTAAATCA TTTtACAGTT 120
 GAAACTGAAG AATATGACTT TAGTGAATTT ACACCAGGCC AATACATCAC AGTTGATGTT 180
 40 TCTAGTGATA AACTTCCATA TAGAGCTAAA CGTCACTATT CTATCGTATC AGGTGAAAAA 240
 AACCATTTAA CTTTTGGCGT TAAACGTGAT GTCACAACAG AACATGAAGG CGAAGTTTCA 300
 ACAATTTTAC ACGATGAGAT TAAAGAAGGC GATATGATTA ATTTAsTGCG CCTGTAGGTG 360
 45 GATTTCGTATT AGAGAATACG ACTGAACCAC AACTTTTCTT AGGTTTCAGGT ATTGGTGTTA 420
 CACCTTTAGT AGCTATGTAT GAAGCTGCCT CTGCCAAAGG TTTAGATACA CAGATGGTTC 480
 AAGTTGCTGA AAATGAACAA CATTtACCTT TCAAAGACAA CTTCAACAGT ATCGCAAGCC 540
 50 ATCATGACAA CGCTAAATTA TATACACACT TAnAgATAAA CAAGGCTATA TTGGTGCTGA 600
 AGAATTACAA GTATTTTTAG CAAATAAACC TGAAATTTAT ATCTGTGGTG GTACAAAATT 660
 55 CTTACAaTCT ATGATCGAAG CACTTAAATC TTTAAATTAC GATATGGATC GCGTACACTA 720

ATATGAACT TAATCAACGT TACGTAAAAG TATTTGCATT ATATTTTCGTA AGTATTGTTA 840
 CTGCAAAATAT TATTGTTAAA AATAATAATT TAATTAAAAC AtTGATACAA ACCATAGCCG 900
 5 GGTACACGGT CTTTGCAGTT GGTTTGAAGT ATTTAACTAA ACGTAAAAAT AAATGACATC 960
 TATACCAAAA CAGCTATGAC TTTACTTTGT CATGGCTGTT TTTTAATTGG GAGTAGGACA 1020
 10 GAGATGATAT TTTTCGCAAAA TTTATTTGGT CGTCCCACCA CAACATGCAT TGATGTATGC 1080
 TCACTGAATT TCATAAGAAA GGAGTTCACA AGATGACCGT AGATATTGGA CGGATTTATG 1140
 ACAATAAAGA TAATACCGAC GCTATTCGTA TCCTAGTCGA TAGAGTCTGG CCGAGAGGTA 1200
 15 TTTTCGAAAAG AACTGCTAAC CTAGATTATT GGTTAAAAGA CATTGCCCCCT TCTACTGAGT 1260
 TGCGACAATG GTTCCAACAT GATCCTAAAC TTTTGGAGC TTTTAAAGAA AAATATGAAA 1320
 AAGAATTACG TGATCAGGAT GCGCAAAAAG ATGCTTTTGA AAAATTAAAG GATATTGTAA 1380
 20 ATCAGCATAA TCATGTTCTA TTGTTATATG CAGCAAAAGA TACTAAACAT AACCAAGCTG 1440
 TAGTACTACA GCAGTTGCTC AATACTTAGT TATATATGTT TATCGTGAAT CACTATAAAT 1500
 25 GTCGCTAACT TCATTTCTAA TATTTATTTT TAAAAAGCAT AATTACTACA ATTAATTGGA 1560
 ACTTTAATAA TTACTAACT TGAACAATA TTTTACTTTA ACAAATAAA GTTTTAAA 1618

(2) INFORMATION FOR SEQ ID NO: 604:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2115 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 604:

40 ATTATAATGC AAATATTGAA GCATTTAAAA CAGTCGCTAA AGCAGTAGGC AAAGAGAAAG 60
 AAGGCGAGAA GCGTCTGGAA AAGCATGATA AAATATTAGC GGAGATTAGA AAGAAAATTG 120
 AACAGAGTAC GTTAAAATCT GCATTTGCAT TCGGTATCTC AAGAGCAGGT ATGTTTATTA 180
 45 ATAATGAAGA TACATTTATG GGACAATTCT TAATTAAAAT GGGTATTCAA CCTGAAGtCA 240
 mAAAArAmAA AACTACGCAT GTTGGTGAAC GCAAGGGTGG TCCTTATATA TATTTAAATA 300
 ATGAAGAACT TGCCAATATC AATCCAAAAG TTATGATTTT AGCCACTGAC GGAAAAACGG 360
 50 ACAAAAATAG AACGAAATTC ATTGATCCTG CAGTTTGGAA ATCATTAAAA GCTGTGAAAG 420
 ATAACAAAGT TTATGACGTT GACCGAAATA AGTGGTTGAA ATCAAGGGGG ATTATCGCAA 480
 55 GTGAAAGTAT GGCAGAAGAT TTAGAAAAAA TTGCAGAAAA AGCAAAATAA AAATACAGCG 540

TGTGATGATT TACCTGAAAG TTTATTTTCGA ATAAATTTAA TTACATAACC GACAAGGATT 660
 GTTTTAACAG TTCTTTTAAT GAATTGGCGC ATCGTTACAT ACCTCATTTT TCTATATCTT 720
 5 ACGAACTATA TACCCATTCA TATATGCTTT TTAACGTC TGTGACAAT TTAATTTTTA 780
 GGAATATAA TATAACCATC TTTATCTGCT TTTTGTAGTAA AAATGACAAA AATTGCATGT 840
 10 ATTATTGaGA TGATGGTAGG GATACCTGTC CAGAAAAATA ATAAGTGAAA AAGACCTTGT 900
 CCAAATTTAT CAGCATAAAA TTTATGaATA CCTAACCTC CAAGAAATAA TGCAACmATA 960
 ACATAAATGG CTTTATTGAC TTTCAATTTGT AAtCCTCCTT AACTATAATT CTACTTAAAT 1020
 15 TCGTTGTGAA AACCAATATT TCTAACTTTA GAATTTTCAA ACTTTCTAAA ATTATAAGTA 1080
 TATCTTTTTTA AAATAAGCTA GAATTTCTAT ATAATAAATG TTAATAACGT AAAAGGGAAT 1140
 GATGACATAG TGATACGTCA AGCACGTCCA GAGGACCGAT TTGATATTGC GAAGTTAGTT 1200
 20 TATATGGTTT GGGATGATAT GGAATTAGAA TTGGTAAAGC ATCTACCTAA AGACATGGTA 1260
 TTAGATGCAA TTGAAAAAAG CTGTGTTGAT GCAACATATC GAACTTTTTA TCAGCATATT 1320
 TTAGTTTATG AAGTAGAAAA TAAAGTAGCA GGTTGTATTA TTAGCTATAG TGGTGAAAAAT 1380
 25 GAATTGAAAT ACGAAAAAGC ATGGGAACTA CTTGACTTGC CAGAAAAAAT AAAACAATAT 1440
 GGCACGCCAT TACCTGTAAA AGAAGCTAAA GACGATGAGT ATTATATAGA AACCAATTGCG 1500
 30 ACATTTGCAG CATATAGAGG TAGAGGCATC GCGACAAAGT TATTAACGTC ATTACTTGAA 1560
 TCAAATACAC ATGTTAAATG GAGTTTGAAT TGCATATTA ATAATGAAGC AGCATTAAAG 1620
 TTATATAAAA AAGTAGGCTT TATATCTGAT GGACAGATTG AATTATACAA GCACATGTAT 1680
 35 CATCATTTAA TTGTTAAATA AAATACTCGA CAGTTCGATG TAAGTCGATT GCCGAGTAGT 1740
 ATCATTTTCT ATTAAATGCC TGCAAATAAT GCACTAATAT AAATACCTAA TGCATATAAT 1800
 AAACCGAAAA ATGTATTTGT TTTACCAGCA GCAGCCATTG CTGGCATCAT TGTAGGCGGT 1860
 40 GTATCATTTCT TCTTGAAACG TCTGATAACT TTAACAGGCA TTGGGAATGA TAACAACGCA 1920
 AGTAAGTAAA ATAATGAGCC ACCAGGTTTA ATAATGATCG TAAGTACAAT AAAGGCATAA 1980
 45 GCGATAAAGT ACATGATTGC CATAAATGTT AAAGAAGCAT TTTTACCTAA TAGAATGGGT 2040
 AAAGTTTTGC GACCACTTGC TTTATCTTTG ACACGGTCGC GAATATTGTT AGCCATATTA 2100
 ATTAAACCGA TAGTG 2115

(2) INFORMATION FOR SEQ ID NO: 605:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1003 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 605:

5 TATGTCTCAA ACTCAAATTA ATCAAATGTT CCAGCAGAAA AACATGTCTA CTGAATTAAA 60
 ACGTCGTTAT GCACAACGTT TATTACAGTT TCCACATGTA CACAATAAAG AATACTTGAA 120
 ATCTTATGCT AAAAACCCTA AAGAAACTAA AGATAGTTAT ATTTCTGGTT TTAAAGAGAA 180
 10 TCAATTGATT AAAATAGAAG CGATTAAATC ATTGTTTGCA ATGGATAAAT CTCCATTAGA 240
 ACATGTTAAA CCTGCTACAA AACCAGACGC TTCTTGGGAT GAGATGAAAC AAAAAGCAGT 300
 TGAAATTGGT AAAGCTGATA CTACATCGAA TAAATTGGT ATTAGAGATC AATACTGGAA 360
 15 ATTAATTCAA GAAAGTAAGC GTAAAGTTAG ACGTGACTAC GAATTCAATG TTAATTCTCC 420
 AGAATTCCAA GATTTAGAAT TACTTGTAAG AACAAATGCGT GCTGCTGGTG CAGATGTTCA 480
 20 ATATGTAAGT ATTCCATCAA ACGGTGTATG GTATGACCAC ATTGGTATCG ATAAAGAACG 540
 TCGTCAAGCA GTTTATAAAA AAATCCATTC TACTGTTGTA GATAATGGTG GTAAAATTTA 600
 CGATATGACT GATAAAGATT ATGAAAAATA TGTTATCAGT GATGCCGTAC ACATCGGTTG 660
 25 GAAaGGTtGG GTTTATATGG ATGAGCmAAT TGCgAAACAT ATGAAAGGTG AACCACAACC 720
 TGAAGTAGAT AAACCTAAAA nTTAAAATAC AAATAGCACA TAACTCAACG ATTTTGATTG 780
 AGCGTATGTG CTATTTTTAT ATTTTAAATT TCATAGAATA GAATAGTAAT ATGTGCTTGG 840
 30 ATATGTGGCA ATAATAAAT AATTAATCAG ATAAaTAGTA TAAATAACT TTCCCATCAG 900
 TCCAATTTGA CAGCGAAAAA AGACAGGTAA TAACTGATTA TAAATAATTC AGTATTCCTG 960
 35 TCTTTGTTGT TATTCATAAT ATGTTCTGTT AACTTAATAT CTT 1003

(2) INFORMATION FOR SEQ ID NO: 606:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1097 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 606:

40 GCTnTTATAT TTAAAAAATT TATTnGCGGA TAATTACCTT TATTAAATCC CACCCATTAG 60
 50 GAWTaCGAGT AATAGGACTA ATACTGGAAT CACATAATGT AACATAACGT CCCTCCTTTA 120
 ACTTAATTTT AATTGTAATC AAATTTGACA ATAAGTCAAA mCATTAAATAC CTATGATArG 180
 TATCAITTTAT TAACATATGT ATCATATTTT TAATCTTGCG TAATTTTTAT CGTTAACTAT 240

CTTTTGTCAT TTTAATTATG TTAAGATAAT AGTAGTTATA GAAGTTCAAT CTATAGGAGG 360
 CATAGCATGG ATATTCCAAA AATCACGACA TTTTAAATGT TTAATAACCA AGCTGAAGAA 420
 5 GCTGTAAAC TATACACAAG CTTATTTGAA GATAGTGAGA TTATAACAAT GGCTAAGTAT 480
 GGTGAAAATG GACCTGGTGA TCCCGGGACT GTACAACACT CAATATTTAC ATTAAATGGA 540
 10 CAAGTATTCA TGGCGATTGA TGCTAATAGT GGCACAGAAT TACCAATGAA TCCTGCGATT 600
 TCATTATTTG TTACAGTAAA AGATACTATT GAAATGGAAC GACTATTTAA TGGATTAAAA 660
 GATGAAGGTG CCATTTTAAAT GCCAAAACG AATATGCCAC CATAACAGAGA GTTTGCTTGG 720
 15 GTTCAAGATA AGTTTGGAGT AAGTTTTTCAA TTAGCATTAC CTGAGTAAAA GGATTCGCAC 780
 AGCtTGgAAT GATAAAGAAA CACTTTTTCT TATGCATGCG TTTTACCTAT GTATTTCTAT 840
 TTTAAGTACA CATTAGCATT TTAGTTTCGC TCATTTTAA AATCACAAAG ATAATTTGTA 900
 20 ACTTAAATGA TGATACGTTA TTTAAAAACA CGATACTTCG TTTCAATGAA CGCATTAAAT 960
 AATAAATAAA CACCTCACCA TAAGAAAGGA CTACTTTCTT TTGTGAGGTG TTTatTTGTC 1020
 25 GACAACTTGA TTATGATTGT TTCATTTTTT GAATAAGTTC ATAATCAGGT GTAGCATACA 1080
 AAGTTTTTTG ATTGTCA 1097

(2) INFORMATION FOR SEQ ID NO: 607:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2031 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 607:

40 GAGGTCTAGC ATGTCTCGTT CAAAAAATA CTTTACTTA TCTAGCTTAA TGATTATTTT 60
 AAGCTTTTTT TTTAATACAA ATAACGTTTT CCTAAGTGA CTTTTTAATT CTTTTATTAA 120
 ATTAATACTT TTCTGCAGTG TTATTAAGTC AATTGTACTA ATTTGTCTA TAATTTTTGC 180
 45 AGATCGTTCA ATTAAATCAC TAAAGCCTGA TGCAGATTGG ATTAGAATTG CGAGTAAAG 240
 TTTGCCTTGG aTTATTCTAA TTGTTATTTT AGTACATATC TTTTCAATTG TTCGTACATT 300
 CGGTTTTATT TAAAAAGTT AATATGTCAT TGTAGCCTTA ATACAAAACA ATACAATGTA 360
 50 TCATGCTATA ATGAGTAAAA CAATTTGATA ACGTTGTTGC GTATAAAAAA TATTAGATTT 420
 TCGAAATCAT AACTATGCAT CTAATCGCTA TAGTTATACA ACAAGATATA ACATATAATG 480
 55 AGGTTTGATA ATGCATCGAC AATTTTTGTC GTCGCGTTGC CAAAACCTCT TTTTAAATT 540

	TTTAGGCGCT TTATTACGTT ACCTCATTTT TTTTCTGAAT ACTGACGGAG GTTTTCCAAT	660
	CGGAACACTG ATAGCCAATT TGA CTGGTGC CTTTGTAATG GGATTGCTAA CAGCCTTAAC	720
5	AATTGCATTT TTTTCAAACC ATCCGACCCT AAAAAAAGCT ATTACGACTG GTTTTCTTGG	780
	TGCTTTAACG ACTTTTTTCAA CATTTC AATT AGAATTAATA CATATGTTTG ATCATCAACA	840
10	ATTTATAACT TTA CTACTAT ATGCTGTAAC AAGTTATGTC TTTGGTATTT TGTATGTTA	900
	CGTCGGTATA AA ACTAGGTG GTGGTTTATC ATGATATCAA TCATTTTAGT CATGATTGGC	960
	GGCGGTTTTG GCGCAATTGC TAGAAGTGCC ATTACTGATT ATTTTAATCA TAAATTTACT	1020
15	TCAAAGTTAC CTATCGCAAC ATTGATAGTA AATCTAGTTG GTAGTTTTTT AATTGGATTA	1080
	ACTATAGGCT TATCAATTTT AATCTCATGG TTCCCTGCGT TCTTTGTTAC CGGTTTTTTA	1140
	GGTGGCTTAA CAACTTTCTC AACGTTAGCC AAAGAACTTA CACTAATGAT GACGCCAAAA	1200
20	TTTAATATTA ACCTTTTTTCT CAATTATTCA CTTTTACAAT TCATCATTGG ATTTATAGCT	1260
	TGTTATATTG GCTATCATAT TTA AAAATAA AATGCTTCAT TCAGCAAATA GGTAAATTAC	1320
25	GACACCTTCC TGAACGAAGC ATTTTTTAAT TTTTCATGCAA ATTTTTAAGC ACCATATAAT	1380
	GCCTACCAAA TTTCAATAAT CTTTGTGTC GTTTAAATAA TGTGAATGTC AATAAATTCT	1440
	CCAAACTAGT CGAAAATAAA GGGAGTGGGA CATAAATCCC TAAAAAACA GCAGTAAGAT	1500
30	AATTTTCAAT TAGAAAATAT CTTACTGCTG TTCTCTATTT ATACAATACT TCGTATTGaa	1560
	TGGcTTTCGCT TTCCTAGGGT GCCGTCTCAG CCTCGGcTTC GACTGGCACT GCTCCCTCAG	1620
	GAGTCTCGCC ATTAATACTA CGTATTAACA TGTAATTTTA CTTTACATA CTTTAAAAAA	1680
35	TAAGACACTT TGCCCAACTT AACTACCAA TAGAAACCTC TGTTAGAATT CCTCAAAATG	1740
	ATATTTTCGCG ATATGTTAAT GAAATTGTTG AAACGATACC TGATAGCGAA TTCGATGAAT	1800
40	TCAGACATCA TCGTGGCGCA ACATCCTATC ATCCAAAAAT GATGTTAAAA ATCATCTTAT	1860
	ATGCATATAC TCAATCTGTA TTTTCTGGTC GAAGAATAGA GAAATTACTT CATGACAGTA	1920
	TTCGAATGAT GTGGTTAGCT CAAAATCAAA CACCTTCTTA TAAACTATT AATCGTTTTA	1980
45	GAGTGAATCC TAATACTGAT GCGTTAATTG AATCTTTATT TATTCAGTTC C	2031

(2) INFORMATION FOR SEQ ID NO: 608:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 687 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

TCCCCGGGAA TCGAACCCGC GAnTCCTGc GTGACATGCA GCGTGTTAA CCGCTACACT 60
 ACGAGACCTA TTAnATTAAA AACTATGTAT TCGGGGAGGC GGATTTGAAC CACCGACCTT 120
 5 CCGGTTATGA GCCCCACGAG CTACCGAACT GCTCCATCCC GCGCTAATAT TATTTTGAAT 180
 TACCTAATTA ATATACCATA ATCAAAAACC TAAAGTsrrG AACTTTTTGA ATTTAATTIA 240
 10 AATGTTATCT CTTAAATAAT TACTTAAATA TCGTAGCAAC ATGTTCTCTG TTGAACACAA 300
 ATATTAGTAT ATTCATTTTT GTAGTGTAAG TCAACGACAT TTTCAAAGTT TTTTGTGTAA 360
 AAAAACGCTT CTTATTCCCT TTTATCATAT AAGTGTCTAA TAGTTGTCAT AAATAGTGkT 420
 15 AAAGCATTTA AAAAGGTATA GGAGTTATAA AGTTTACAAC GCCTATACCT TCTGAAAAAG 480
 AAATTATAAA AGCTTGTTAC ACCGCATATT CTTTCAGTCA GCGACTACCA ATATAACATT 540
 GTAGCCCTAA GACATTGCTT GACGCCTCAn TTACAACAAT TTTTCAAAT CAGCAGCTAC 600
 20 CTACTGACAC AACATAACAC AACCChTACA CTAnCTATCG TGTCATGTAA TCTTGCATCC 660
 GATCTTGCAA CGCTGTAAAT GTTTCGA 687

(2) INFORMATION FOR SEQ ID NO: 609:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 843 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 609:

35 GTGTTGAATA CTTTGACAAA ATGAAATCAT TTGTTCTTTC GTTTCAAATT TAACTGTTTG 60
 AATTAAATCA GTTCGTTTTT CGTAGTACTT CGGCGTTGTG TTCATATTCA TTTTTTCTAA 120
 40 AAATAAACTA GTAAACAATG CACCTTTAAG ACTCTGACTG ACAACGTGTG GTGCTAAAAA 180
 GAAACCTTGA TACATTTCAA GCAATGCATT TAATGATGCA CCCGCTTCTT TACCAATACC 240
 AGGTGCTGTC AATCTATAAC CACATCGTTC AATTAAATCT TTTCTACCAG CAATGTATCC 300
 45 ACCAATCTTA GCTAAACCAC CGCCAGGGTT TTTAATTAAT GATCCTGCTA TTAAATCGGC 360
 ACCACATTCT ATAGGTTTAC GTCTTTCAAC AAATCCCCCA TAACAGTTAT CCACAAATAT 420
 TAAAATATTA GGATGCACGT TTTTCAACCT AGTAATTACC TTTTCAATTT CATCTAGCGG 480
 50 AATTGAAGGT CTTTGATCAT AGCCTTTCGA ACGTTGAATC GCTATTACTT TGGTGCGCTC 540
 AGAAACCCCa TCTAACACAC TTTCAATATC GATCTTACCT TCTTTAAGTG CAATATCTTT 600
 55 ATACGATmCG CCATGCTCCA TTAACTTTTC AATACCATTT CCGTTTACGC CAATGACTTC 660

ACTTTGTAAT GCAATAGTAA TCGCATGCGT ACCTGAAATA ATTTGCGGAC GAACAATTGC 780
 ATCTTCTGCT TTAAATGCCT GCGCATATAT TTCTTAATTG AACGAATTGA GAACTAATCG 840
 5 TAA 843

(2) INFORMATION FOR SEQ ID NO: 610:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 775 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 610:

GATTATTATA AACCGAATGA AACAAAAAAG TTTTACATA TTTTCAACAA CTTTATTTCA 60
 20 ATAAAGCATC AAGTTGATCT AAAGTTTGAT TCATACCTTG TTCAACACCC ATGTTTATGA 120
 CTGTGTGAGC GGCTTCTTTT GTTGGAATA CCGATGTTGA TGTCACTGTC GTTTTCGTTG 180
 25 TATTGCTCTT GGAAATGAC AAAGTAATTT TCATGCTTGG CATTTTTGTA TCTTTTGCAC 240
 CTTGAGGTGT CGCAAATGAG TCAATATATT CAATTAAATA TGGACGCTTA ACCGTTTTAT 300
 ATTCTGCTAA TGTATAGCTG GTCATTGTAG GTGTTTTAAT CGCATAAAAT GCATCACCGC 360
 30 CTGAAACAGC ATTAAAGCGA AACACTTTGG TACTAGCGTC TTTTGGATGA AACCATTTTT 420
 CAAATAATGC TTTkGTCGTA TAGGCATCAA ACACCTTTTC AATCGGTGCT TCTATCGTTC 480
 TTGAAAAAAT AATTTTATTA TCTTCAACTT TAATCGTCAT CTTCCCACTC CCTATCTTTT 540
 35 ATATTTACTC ATCTTAACAT GCATTGAGTA ACATATGTGA CCAAATCATT TGTATAAAC 600
 TATAGTAATA TTAATCATTC CCTCTGTnGA AATTGATAAC TACGAAATAA AAAACACACT 660
 CTATTCAGCT ATTAACAGTT GAGTAGAGTG TGCAAACCTG TTGAATTAAT GCGCTAACAT 720
 40 TTCTTCTTTG ATTTGGTCTT nATnTAATTT TGAAGGATAA TATGTTGGCC AGTTA 775

(2) INFORMATION FOR SEQ ID NO: 611:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 728 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 611:

ATCAAAATTT GATACTAAAG GGCTCACGCT ATTTTATGTA TTTATTGGAT TGATTATGTT 60

TGTGGCAATG TGTTTATTTA AAGTAGAAAA ACATGTTTCT TCACCATTTT TACCTGTGGT 180
 TGAATTTAAT CGTTCGATTA CTTTAGTTTT TATAACTGAC CTTCTAACAG CTATTTGTTT 240
 5 AATGGGATTC AATTTATATA TTCCAGTCTA CCTTCAAGAA CAACTAGGAT TATCTCCATT 300
 GCAAAGTGGG TTGGTTATTT TTCCTTTATC TGTAGCTTGG ATTACATTGA ATTTTAATTT 360
 10 ACaTCGAATT GAAGCAAAAC TATCAAGGAA AGTTATTTAC TTAATATCAT TTACATTGCT 420
 ACTAGTAAGT AGTATTATCA TTTCATTTGG TATTAAATTG CCGGTACTTA TAGCATTTGT 480
 GTTAATTTTG GCAGGATTAA GTTTTGGATA TATTTATACG AAAGATAGTG TGATTGTCCA 540
 15 AGAGGAAACT AGCCCATTAAC AAATGAAGAA AATGATGTCA TTTTATGGAT TAACTAAAAA 600
 TCTTGGGGCA TCAATAGGTT CAACAATAAT GGGATATCTT TATGCGATAC AATCAGGAAT 660
 CTTTGGTCCA AACTTACACA nTGTGTTAAG TGCTGTTGCT GTAATTAGCA TGGGCTTATn 720
 20 GTTTTATG 728

(2) INFORMATION FOR SEQ ID NO: 612:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 913 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 612:

TATCAGAGCA AATAAGCAAG TTAGATAAAA AGATTTCTGA CTTACAATTA ATTAGACGCT 60
 35 CTGTATGTGA ATTTATTAAA GGAATCTCTC TAATAGATAC CAGCATTTTA AACAGACAC 120
 TACAGTCACA ATATGATAAA GAAGCATCTA TAAATATGG TCATACGAAA GCATATCAnT 180
 40 CATTATTAG ACGTAAAGAC AGCTTACAAT CGCAGGATAT CAGACATAAA TTGACAACTA 240
 TCTTCAATAA ATTTAATCAT ATGTCTTTGA GTCATTATCC AATCCAAGAT TGTAGTGATC 300
 TCGTATTTGA GTGGAAGGCA TTTATGAACA CTATCGCTGA TTTTGmTGAT GAAACATTAT 360
 45 GCTGTATTGC TAAAACATAT GAAGATGATA CGCGTTTCAA AGATTACTTT AATTCATATG 420
 ATAATCAAAA TTTAGCATCA TACATTTAG AAGCTGTTAA TTATTTTTTG AGCAATGTGA 480
 ATAAGAGCGA CAATTTTAA TCCTCATACA GATGCTACAA CACTGAATAC ACCAACTAAA 540
 50 ATAATTGTTT AAAGTAGAAA ACTAATCATT TATTCTCTAA ATCAATAACT ATATTAAAAG 600
 TTATACCTTT GCAAAGCGAA TTAGTATAGG TTACCGAAAG GAGAAAGGAT TAGGTTCCAT 660
 55 TCGATTTATA AAAAAATATA TTTGAAACA TACAAAAGCG CCAGAAGATA ATCATTTCTT 720

TCAATGCATA TTCAACTAGC TCAGGTTTAC TCTTTAAGCC AAGCTTTGTC ATAATATGCG 840
 TCTTATGTGC TTCTACTGTT TTCACAGATA CAAATAATTT CTCTGCAATT tCTTTAATCC 900
 5 CGTAACCTTT GGC 913

(2) INFORMATION FOR SEQ ID NO: 613:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 654 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 613:

ACAAGTGTG ATGTCATAGT AGCGTCAAWT GTGCTTTATT TCTGGGACAC ATTTAAATTT 60
 20 TTCCACCCCA TTAAGTGGAG CTTAATCTTT ATTACAATTT TACTATTATT AAACATTTTT 120
 TCTGTAAAT CATTGGAGA AACTGAGTTT TGTTATCAT TGATTAAAGT GTTAACAATT 180
 ATCGTATTCG TTATTTTTGG CTTTTTAATG ATTTTCGGTA TCTTAGGTGG TCATACATAT 240
 25 GGATTGAAA ACTATACAAA AGGCCAAGCA CCGTTTGTG GTGGTATCTC TGGTTTCTTA 300
 GCGTATTAT TAGTCGCCG ATTTTCGGT GGTGGTACAG AAGTAGTAGC AGTAACTGCT 360
 30 GGTGAATCAG ATGACCCTAA AAAGTCTATG CCTAAGGCAA TTAAACAAGT ATTTTGGCGT 420
 ATTCTTTTAT TCTATGTCTT ATCAATTGCA GTAATTGGTG CAATTATTCC GTACACAGAT 480
 CCATCATTAT TAAGAGCAAG TAGTTCAATA AGTCAAAGCC CATTTACAAT TGTATTGAT 540
 35 AGAGLAGCA TAGCCTTTC AGCATCAGTA ATCAACGCGG TTATTTTAAC TTCATTATTA 600
 TCCGCTGCAA ATTCAGGTGT TTATACAACA GGCAGAATGT TGTATTCCTT AAGT 654

(2) INFORMATION FOR SEQ ID NO: 614:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 642 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 614:

TCGATCTTTA TTGATAATAA TTAAATTGTC GCCTTTAAAA TGTGATATTA ATCCTGCGGC 60
 50 AGGnTGTA ACGAGTGATG AACCTAGTAC AACAAGGGTG TCAGCATGTT CAATTTTATT 120
 TAATGCCCTT ATGATGGTAG GTTGATCTAA CATTTCACCG TATAATACGA TGTCCGGTCG 180
 55

ATAAGATTTA TGACATACAT TACAATAAAA ACGATTTAAC GTGCCATGTA ATTCATCAAC 300
 ATGTTGACTT CCAGCGTCTG AGTGCAAACC ATCGATATTT TCGTGATGA CACCTAAAGA 360
 5 TTGTTGATTA CGTTCTAATT TTGCAATCCA ATCATGAACG ATATTGGGCA TCGTATCGAC 420
 AAATAGTAAG CGCTTATGGC AGAAATTGAT AAAACCTTCA GGATCATCTT CTAAATAATC 480
 10 ACGGCTTAAC AAGTATTCTG GCGAAAGCCC ATCTTTTGaA ATTTTCATCAA ATAAGCCACC 540
 CATTGaACGG AAATCTGGAA CGCCACTTGC GACAGATACA CCAGCACCTG TAAAAAATGT 600
 AATACGATTC GAACTATCTA TAATATGTTT TAGTGTCTCT AA 642

15 (2) INFORMATION FOR SEQ ID NO: 615:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1210 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 615:

TTTCCCCnCC CnCCAAATA TCCAAnGGAA CTTTAATAGT CCAATTGGCA CAGTAAAACT 60
 ATGGCATTITT ATAAGTATAA TATATCTGTA ATTTATGGTC AATTAGTAAA TTGTTTTTTA 120
 30 TTTGAAACAT ATTTACATc AAAATCACAA AGACTTTTAG ATTTTGtTCT AAAAACTCT 180
 TAATAATTTA TTTAATGAGA AGAGTTGCTT ATATAGTAAA TTGTGAAGCC GTTAAAACAA 240
 CGTTACAAAA CCTATATCTT TAATACGGAA CCATATGGTA TGAATCAAGG AATACTTAAA 300
 35 CTAAAACTTC TCTATCAGAT TTATTTGTTG CGAAATCAAC AACTTTAATT GCTTGCCCTT 360
 CATTTAATGG ATAATTTGCT TGCgTAATTT TAACTTTTAC AATTTGACCT ATGAGTGATT 420
 CGTCACCTTC AAATTGTACT TTCATATAAT TATCTGCATA TCCAATAAT GTACCTTCTG 480
 40 TGtCACCTG TTCTCAGGA ATTACTTCAA GCACATCTTG ATCAAATTTA GACGCATATA 540
 ACTTTCCGAG TTGATTGCTT AGCGTAATTA ACTTATGCAC CCGTTCATTT TTAATTTCTT 600
 45 CATCAATTTG GTCATCCATT CTTGCAGCTG GCGTGCCAAT TCTAGGAGAA TAAGGGAAAA 660
 CATGCAGTTC AGAGAACTTA TGCTTTACGA TAAATCATA TGTTTCTTGG AACTCAGCTT 720
 CAGTTTCACC TGGGAAACCA ACAATTACAT CACTCGTAAY TGCCAAGTCT GGTAAAGCTT 780
 50 TATGCAATTT TGTTAATCGT TCTGAAAATC TATCCATTGT ATACTTACGT CTCATACGTT 840
 TTAATACTGT ATCTGAACCA GATTGTAATG GAATATGCAA ATGACGCACA ACTTTTGTG 900
 55 AACGTTCTAA AACGTCAATT ACTTCATCTG TAAGTTGACT TGCTTCAATT GAAGAAATTC 960

CTTTTAAATC TTGACCATAT CCACCTGTAT GAATTCCCGT CAATACAATT TCCTTATATC 1080
 CTGAATTCAC TAGTTGCGTC GCTTGTTCAA CTACTTTTTTC CGGATCTCTT GAACGCATTA 1140
 5 AGCCACGAGC CCATGGAATA ATAcAGAATG TGCAGAAGTT GTTACAACCT TCTTGAATTT 1200
 TTAATGACGC 1210

(2) INFORMATION FOR SEQ ID NO: 616:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 652 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 616:

20 TAATAAAATA CAATACTTTT CAATACAGAG AATCCCGCAA TGTCGGGATT CTTATTTATG 60
 CTGATTTTGT TTTTGTCTAT GATCAGGGAC TTtCAGGGAC TCAATTAATT ATCACATTCA 120
 TATTTTCTTA TGACCAAATT GATACATATG ATGCTATTTA ATGAGAAAAT TAGGCATCAC 180
 25 TTGGTTATTG AATTCCTTC ATTAACTTT CCAGCTCAAT TTAATAGTTA GTCGACTATT 240
 ATTCATTAAA CACTTTTTTAA TCATAAAAAA GTGTTTTTga TAATTCACTa CCaAAACAC 300
 30 CTTCTTACTT ATAATTCTAT TTGTTACCA TTTCTAATCT TATCGGCTAA ATCATTCACT 360
 TTTCTTAATC GGTGATTTAC ACCTGATTTT GAAATTGGAC CAGTTGATAC CATTTCTCCA 420
 AGCTCTTTCA ACGAAATTTT TTGATGTTCT ACTCGAATTC TAGCAATCTC TCTCAACCTG 480
 35 TCTGGTAAAT TTTCAATACC AATTTCTTTA TCAATCAATT TAATGCTCTC AACTTGTTTC 540
 ATCGCAGCAC TAACTGTTTT ATTTAGATTG GcCGTTTCAC AATTAACGAG TCGGTTAACA 600
 GAATTACGCA TATCTCTTAC AATACGTACG TCTTCAAATT TTAATAACGC TT 652

(2) INFORMATION FOR SEQ ID NO: 617:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 617:

50 TnACCACTTT AAAATAGCGC TTAATAAAAtG AAGGGGGCAA GTCTTATGAC GtTTTACaAT 60
 TTCATCaTGg GTTTTCAAAA TGtTAACACA CCATTTGGTA TATTGGCCGa ACACGTTAGT 120

TCTAATTACA CAGATCATCA ATTAATTGAA ACTACAAATA GAGCTATTAG CTTATATATG 240
 GCAAATTAAT TTGAGTAGTA CCAATTATGA TGTATTAGTG CATCCCAAAT ATCTTTTGT 300
 5 TTAAAGTTTA TTTCATCATT TCTTATCGAA AATGGTGTA TAATGTCTTT ATCTAACCAG 360
 GTGTTGATAA GTTCATTTGG TACACCATCT AACACATTT CACTTTTACT AATTATAAAA 420
 10 CATTCCCAGT CAAGTGAAAC ATTTTGTGGA TTCACATAAT TACATTGATT ATGATTATCC 480
 ATAAACACTC ACTCCTTTAA AATTCTGTAC TCTTCATTGC GTTTTACCCC GTCACATTAT 540
 CTTTTAAACT AAAATCATCA TTACTTATGA AAAAAATGTA CATCAAAGC AAAGGTTTTC 600
 15 GCTACCGAAA AAGTTTAAAT AATGGTttaa TATATtGGT ACTCATTTTA ATAAAAAGAG 660
 AATACATTTT GAGCTATCAA TACTTTTTAT TGAAGAGGTG TTATTsyTGG CTAAAACGTT 720
 ATATTTAATG CGCCACGGAC AAACTTTGT TAATTTAAG GGAATAATtC AGGGATTGTG 780
 20 AGATTCGCCG CTAACAGA 798

(2) INFORMATION FOR SEQ ID NO: 618:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1786 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 618:

ATACGCTCAA TTGATAAAAA TTAAATATCT GCCATTTGTA TACGCATCGC TTCGAATGTT 60
 35 TCCGTGCAA TATCAAATGA TAATTGTGAC TGGAACCTTA AACATCGAAT CATACGTAA 120
 GCATCTTCTT GGAATCGTTC CTCAGCTATA CCTACAGTTC TTATTATTCG ATTATTAATA 180
 40 TCTTGTGAC CATCAAATA ATCATACAAT TTGTATGCTG TATCCATTGC TATCGCATTC 240
 ATCGTGAAAT CTCGTCGTTG CAAATCTTCG TATAAATCAC GAACAAATGT AACACCACTT 300
 GGTCTACGGT GATCGACATA ATCTTCTTCA GCCCGGAATG TTGTCACTTC ATAATTTtCA 360
 45 TCATtAAAAa CTACATTtAT CGTGCCAtGT yCTThACCTA CAGGTATCGT ATGACTAAAG 420
 ATAGATTCTA TTTCATCCGG CGTTGCACTT GTTGTGATAT CTATATCATG AATATTTCTT 480
 CCCATGACAT AATCTCTTAC AGAGCCACCT ACATAATATG CTTCAAAACC ATTGTCTTGA 540
 50 ATTTGTTCTA ATATAGGCCT TGCCTGTTCA AATAATGATT TATCCATATT ATTACTCGCC 600
 TTTACTTTTG TTATGCTCAT TTAGCATTTT TTGATAATAA TACTCATATT GATCTGTAAT 660
 AAGTTCTGAT CCAAACGTT CAGCAATATC TGCTAGCATG TTTTCTGAA GTTTGTTGTA 720
 55

ATCTACGACA AATCCAGTTT CACCATGTTT AATAACCTCT TTAATTCCAC CGGCATTGGA 840
 ACCAATTGGA ACGACGCCTG TWTTTCATAGC CTCAAGTAAA GTTAGTCCAA AGCTTTCTTT 900
 5 TTCACTTAAT AATAATACTA AGTCAGATAA TTGGTAAAAT TCACTTACGC AATCTTGTTT 960
 CCTAAAAAT AAAACATCCT CTTCTACGTT TAACTCTTTC GTCAATTGAC GCATTGGCAC 1020
 10 TAATTCAGGA CCATCTCCAA GTAAAATTAA TTTACTAGGT ATCTTTTCAC GTACTTTTGC 1080
 AAATGTTTCT ATAATAGTAT CTATGCGTTT TACTTGTCTA AAATTCGATA CATGTATTAA 1140
 CACTTTTTCA TCTGGTGCTA TACCAAATTG TGATTTTAAT GCTGTGTTAT GTTTAGTTGG 1200
 15 AAATCATTTC TCACGTACAA AATTATAAAT CGGTATAATT TCTTTGTTAG TTTCAATAAT 1260
 TTCATGTGTT TCTTGTGCTA AAGATTTACT CACACTTGTC ACAATATCAC TTTTTTCAAT 1320
 GCCAAATTTA ATTGCACCTT GGAGTGAATG ATCATAGCCC AAAACAGTAA TATCAGTACC 1380
 20 GTGTAGCGTT GTCATAATTT TTATATCTTT ACCTGACATC TCACGAGCTA AAATCCCACA 1440
 AATTGCATGA GGTACAGCAT AGTGCATATG CAACAAATCA AGATCATATT CTTTAATAAC 1500
 25 TTCAGCGATT TTAGTACTTA ACGTAATATC ATACGGTGGA TACTGAAATA CTGCATATTG 1560
 ATTCATTCA ACTTGATGAA AAATCATATT CGGTAATGGT TTTCTTATTC TAAACGGGAT 1620
 ATTTGAAGTG ATAAAATGTA cTTTCGTGACC TCGCTCTGCT AATTTAATTC CTAATTCTGT 1680
 30 GGCAATAATT CCAGAACCAC CCATGGACGG GTAACATGTT ATACCTATCn TCATTGCTT 1740
 GCCCATCCTT TCTTTCTATT TCTCnCTATG ATnCTCGATG CGTAGA 1786

(2) INFORMATION FOR SEQ ID NO: 619:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 844 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 619:

45 ACAGGTATGG ACTTnGCTCA AATGACACGA CATTATTTAT CAAGACCTAT TGCTATAATC 60
 TTTTGGATCA TTGCAGAACT AGCAATTATC GCTACAGATA TTGCTGAGGT TATTGGTAGT 120
 GCTATTGCTC TTAATCTCCT ATTTAACATA CCTTAAATCG TCGGTGCACT AATAACTGTA 180
 50 CTTGATGTAT TTTTACTACT TTTTATAATG AAATATGGTT TTAGAAAAAT TGAAGCTATT 240
 GTTGGTACAT TAATTTTCAC AGTGTTATTC ATCTTTATAT TTGAAGTCTA TATTTTCATCA 300
 55 CCACAGTTGA ATGCTGTGTT AAATGGATTT ATACCACATA GTGAAATCAT TACAAATAAC 360

TTACATTCAT CAATTGTACA ATCTAGAACA TACTCAAGAC ATAACAATGA AGAAAAAGCG 480
 CAAGGaTTAA ATTTGCTACG ATAGATTGGA ACATTCAGTT ATCAATCGCA TTTGTAGTCA 540
 5 ATTGCTTATT ATTAGTGTaTA GGAGCATCAC TATTTTkCAA CTCAAATGCT GACGATTTAG 600
 GTGGTTTCTA TGATTTATAT CACGCCTTAA AAActGAACC TGTACTAGGT GCAACAATGG 660
 10 GTGCAATCAT GAGTACATTA TTTGCAGTTG CATTATTAGC GTCAGGTCAA AATTCAACGA 720
 TTAGTGGTAC TTTAGCAGGA CAAATTGTAA TGGAAGGATT TTTAAGATTA CACATACCAA 780
 ATTGGTTAAG ACGTTTAATT ACACGTTCTC TTGCTGTCAAT TCCTGTTATC GTATGCTTAA 840
 15 TCAT 844

(2) INFORMATION FOR SEQ ID NO: 620:

(i) SEQUENCE CHARACTERISTICS:
 20 (A) LENGTH: 574 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 620:

TGTATTGCAG TCATGCCATA AATTTTTTTGA ATAATAAAGG GAGACGCAGA AATATAAGTA 60
 30 AATAATATTA CAAATGTCAT ACCTTGGAAT GAGCATTGGT AATACAAAAC GTGGCGTCTT 120
 CAhTAATATT TTGAAGTTTT TAAACATTGT CTTTaaTCCA CTACTTGACT CACGATTTGT 180
 CACTGTTAAT GATTcAGGTA CTTTTAATAA AGAACCTATG ACCATGACGA AGCCAAAGAT 240
 35 AGTCAGAATG ACAAAGACCA TACGCCAGAC AGAATAATTT AAAATTATGC CCCCTATTGT 300
 TGGCGCAACA ACTGGTGCAA TACCATTAAc AAGCATCAAT AATGCCATAA ACTTAGTTAG 360
 TTCATTACCA CTATACATAT CACTTGCTAT AGCTCTTGAA ATAActGCTG tGCGCCACCT 420
 40 GTCACTCCTT GAAGaAATCT TAATGCAACC ATCAGCCAAA TATTATGTAC AAAAACAATA 480
 CCTAAACTTG CTAATGTAAA AATAATCATA GCTATAATAA GCGGCTTTCT GCGCCcAGTT 540
 45 GAATCTGAAA TTGGACnAGC AACCAAATTA CCAA 574

(2) INFORMATION FOR SEQ ID NO: 621:

(i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 523 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

AAGCAAGAGA GAGTACAACA TTTATATGAT ATTAAAGACT TACATCGATA CTACTCATCA 60
 5 GAAAGTTTTG AATTCAGTAA TATTAGTGGT AAGGTTGAAA ATTATAACGG TTCTAACGTT 120
 GTACGCTTTA ACCAAGAAAA TCAAATCAC CAATTATTCT TATTAGGTAA AGATAAAGAG 180
 AAATATAAAG AAGGCATTGA AGGCAAAGAT GTCTTTGTGG TAAAAGAATT AATTGATCCA 240
 10 AACGGTAGAT TATCTACTGT TGGTGGTGTG ACTAAGAAAA ATAACAAATC TTCTGAAACT 300
 AATACACATT TATTTGTTAA TAAAGTGAT GCGGAAATT TAGATGCATC AATTGACTCA 360
 TTTTCAATTA ATAAAGAAGA AGTTTCACTG AAAGAACTTG aTTTCAAAAT TAGACAACAT 420
 15 TTAGTTAAAA ATTATGTTTT ATATAAAGGT ACGACTAAAT ACGGTAAGat CACTATCAAT 480
 TTGAAAGATG GAGAAAAGCA AGAAATTGAT TTAGGTGATA AAT 523

(2) INFORMATION FOR SEQ ID NO: 622:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1871 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 622:

30 AGAAGTGTGA nAAAAATTTA AnAGAGATAT GCACATAGAT GACGCATTGC TATATCCAAG 60
 CAATTGAGAA AGCTGCTGAT GCTCCAAATC aCGGAATGAG GGAACCATGG AGAGTTGTGC 120
 ATGTTCCGAA AGACAGATTA GGAGATATGA GTAAGGATAT TTCTAAATTT GCATTTCTTA 180
 35 ATGAATTAGA TAAGCAACAA TGTCAATTATG ATGCAGTTAC GAAACTAGGT GGCATGTTAT 240
 TGCTTATTTT AAAAACAGAT CCAAGACAAC GTCAAAATGA TGAAACTAC TTTGCATTTG 300
 tGCATATGCA CAAAATCTTA TGTGTTACT TTATGAAGCG GGAATAGGTA CATGTTGGAA 360
 40 aTCGCCATTA TATATCTATG ATCCTAAAGT AAGnAAACAC TTGGTATAAA GnAAGATGAA 420
 GTTCTTGCTG GATTCTTATA TTAAACGGAT TTAGAAGnAG ATATGCCTAA AGCACCACGT 480
 45 AAAAATAGAA ACTTAATTAC ATTATATTAA TATGTATAAT TATAGAAACA TTAATAAAAG 540
 CTGAGTCATG AATTGATGGA CATCTATCGA GTTAGAGATT TAATCTAACT TACTAGAGTC 600
 GGTACAATAA CAGTCTCAGC TTTTATTGT GCAGTATATA CACATTTTTA TTTTAGTATT 660
 50 TATTtAAAAG TTtCTGCTaA AAATGATTCA ACTTGTTTCA GTGACTTAGC ATTTGCTGAA 720
 TGAAGGTGTG CAATTTTATC GCCGTTTTTA AATACTAGCA AGCTAGGGAT ACCCATAACT 780
 55 TCATTTTCAA CAACTACATC TTCTAATTCA TCACGATTAA CAGTATACCA TTGGTAATCA 840

CCTGCCTCAA ATTAAACAAT TACAGGTGTA TCGCTATTAA TTACAGATTT AAATGATTCA 960
 TTACTTTTGA TTGATTGCAT TGTAACAACT CCTCTAGATA GTTAAATAAT TTTTATTATA 1020
 5 GCTAAATTTA TATCATAATA AAAAATTTTA GCTTCAAAAT GAAAGCCTTT TGCTTTGGAA 1080
 AATGATATAT TTATTTTAAA TACATAAAGG AGGTTGCAGT CGTATGATTA AATTTTACCA 1140
 10 ATATAAGAAT TGTACAACTT GTAAAAAGGC AGCAAAGTTT TTAGATGAAT aTGGCGTAAG 1200
 TTATGAACCA ATTGATATCG TTCAACATAC ACCTACAATA AATGAATTTA AAACAATAAT 1260
 TGCAAATACA GCGGTAGAAA TTAATAAATT GTTTAATACA CACGGCGCGA AATATCGTGA 1320
 15 GCTTGATTTG AAAAATAAAT TACAACTTT ATCAGATGAT GAAAAGTTAG AGTTGTTATC 1380
 ATCTGATGGT ATGTTAGTAA AGCGTCCTCT AGCAGTAATG GCGGATAAGA TAACATTAGG 1440
 ATTTAAAGAA GATCAATATA AAGAGACTTG GTTAGCGTAA GTGaAATGTA AGCGTTTACT 1500
 20 AAATATCTCG ATATTTAGAT TCATTACATG TAAATGAAA TAAGCTATAC AATTGTTAAT 1560
 TTTTATAAAT ATAGTTGAAT AGCATCTAGC CTTATGGCAT CATTAAATGAT GTAAAGATTA 1620
 ATTAGGAGGG GATTCTCTTG GCAGTACCAA ATGAAYTGAA ATATTCAAAA GAGCATGAAT 1680
 25 GGGTTAAAGT TGAAGGTAAT GTAGCAATAA TTGGAATCAC AGAATACGCA CAAAGCGAGT 1740
 TAGGTGATAT TGTTTTCGTT GAATTACCAG AAACAGATGA TGAAATTAAT GAAGGGGATA 1800
 30 CGTTTGGTAG CGTAGAATCA GTTAAACTG TATCAGAATT ATATGCACCA ATCTCTGGTA 1860
 AAGTAGTTnA A 1871

(2) INFORMATION FOR SEQ ID NO: 623:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 451 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 623:

45 TTGCTGTAGA AGATAAAGAG TCCCATCAAT GGATAGGCTT TATAGGTTTG AATTATATTC 60
 CAGAAACAAG CGATTATCCA TTAAAGAAT TACCGCTTTA TGAAATAGGT TGGCGCTTGT 120
 TGCCAGAATT TTGGGGAAAA GGATTAGCAA CTGAAGGCGC AAAGGCAACA TTGAAGTTAG 180
 50 CAGAAGAACA TCAATATAC GATGTCTATA GTTTTACAGC AGAAGCAAAT AAAGCTTCAC 240
 AACGTGTAAT GGAAAAAATT GGCATGACAG TGTATGATCA TTTCGAATTA CCCAATCTAA 300
 55 GTAAGTATCA TTTATTaAAA AGGcAAGTGC GCTATTACAT TAATCTTCcG AAAGTGGA 360

ATTAGGGGGT GtTtnGTkTn ATTTTTTTAA n

451

(2) INFORMATION FOR SEQ ID NO: 624:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 665 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 624:

CGATAGTTAT CTAAAGCCAT TTTnGATTGT GTTATGAAAT CTAATGATGC GTGATAATTT	60
AATGCTACAT AACGtTaATA TAAAATATCA ATAGTGAACA TTTGAGCAA TAATGAAGTT	120
GTTGCTCCCA TATGCATCTC ATTTTCATCA GTTTTCCCAT AAGTTAAAAC AATATTTGAT	180
GCCTGTGCTA CGGGATTATC CCTTGTTACTA GTAATTGTAA TTATAGGTAT ATGGTAGTCA	240
TCAATAACTT TAACCATTGA TTGCATTTC CTTTGCCTAC CATTGTTAGT AATAAGAATA	300
aCACTGTCGT TCGAATTGTG AGTTGCTAAT AATGTAGCAA AAATATGTGT TTCTTGAACA	360
AGTTGAATAT TAAGACCTAT TCTTGATAAC TTTTGGTATA AGTCGGTAGC AACTACAAAA	420
GATGCGCCAA AACCATATAT AAAAATCGTC TCAGAACGTT TTAAACAATG ACATATTTGA	480
TCAATAGTTT TATCATTTAA TTCGTTATTT GCATGATTAA GTGCGCGTGT AGTACGTGTA	540
TGGAGTTTAG TTCTTAAAGA TTCTGTGCTT TCGTTATTCA TTAATTCAAC ATTGTAAATT	600
GATGATGCTT TAGGAACATA TTTAGATATA TTTATTTTCA AGTCGTGAAA ACCGCCATCA	660
GTAAT	665

(2) INFORMATION FOR SEQ ID NO: 625:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2549 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 625:

GATGCCAATT AACCGTGCAT ATAATGTTGA GAAGTTAATC GAAGCAATTC AATATTATCA	60
AGAAAAACA AATCGTCGTG TTACTTTTGA ATATGGTCTG TTTGGTGGTG TGAATGACCA	120
ACTAGAACAT GCAAGAGAAT TAGCACATTT AwTAAAAGGC TTAAACTGCC ATGTTAACCT	180
AATTCCTGTC AACCATGTTT CAGAAAGAAA TTATGTGAAA ACGGCTAAAA ATGATATCTT	240

	TTCCGATATT GACGCAGCTT GTGGTCAATT AAGAGCAAAG GAACGACAAG TAGAAACGAG	360
	GTAAAGACAA ATGCTAGAGG CACAATTTTT TACTGATACT GGACAACATA GAGATAAGAA	420
5	TGAAGATGCG GGTGGTATTT TTTATAATCA AACTAATCAA CAACTTTTAG TTCTGTGTGA	480
	TGGTATGGGT GGCCATAAAG CAGGAGAAGT TGCAAGTAAA TTTGTTACAG ATGAGTTGAA	540
	ATCyCGTTTT GAAGCGGAAA ATCTTATAGA ACAACATCAA GCTGAAAATT GGTTGCGTAA	600
10	TAATATAAAA GATATAAATT TTCAGTTATA TCACTATGCA CAAGAAAATG CAGAATATAA	660
	AGGTATGGGT ACAACATGTG TTTGTGCACT TGTTTTTGAA AAATCAGTTG TGATAGCAAA	720
	TGTCGGTGAT TCTAGAGCCT ATGTTATTAA TAGTCGACAA ATTGAACAAA TTAGTAGTGA	780
15	TCACTCATTG GTTAATCATC TTGTTTTAAC GGGTCAAATT ACGCCGGAAG AAGCATTTC	840
	ACATCCACAA CGTAATATTA TTACGAAGGT GATGGGCACA GATAAACGTG TGAGTCCAGA	900
20	TTTGTATTAT AAGCGATTAA ATTTTTATGA TTATTTATTA TTAAATTCAG ATGGATTAA	960
	TGATTATGTT AAAGACAATG AAATTAAGCG TTTGTTAGTA AAAGAAGTA CAATAGAAGA	1020
	TCATGGTGAT CAATTAATGC AATTGGCATT AGATAACCAT TCGAAAGATA ACGTTACTTT	1080
25	CATACTCGCG GCTATTGAAG GTGATAAAGT ATGATAGGTA AAATAATAAA TGAACGATAT	1140
	AAAATTGTAG ATAAGCTTGG CGGCGGTGGC ATGAGTACCG TTTATCTTGC TGAAGATACG	1200
	ATACTTAACA TTAAAGTTGC AATTAAGGCG ATTTTTATAC CACCTAGAGA AAAAGAAGAA	1260
30	ACATTAAAAC GTTTTGAACG AGAAGTACAT AACTCATCAC AGCTATCACA TCAAAATATA	1320
	GTAAGTATGA TCGATGTTGA TGAAGAAGAT GACTGTTACT ACTTAGTAAT GGAATATATC	1380
35	GAAGGTCCGA CTTTGTCTGA GTATATTGAA AGTCATGGGC CATTAGTGTG TGACACAGCG	1440
	ATTAATTTTA CGAATCAAAT ATTGGATGGC ATTAAACATG CGCATGATAT GCGTATTGTA	1500
	CATAGAGATA TTAAGCCACA AAATATATTA ATTGACAGCA ATAAAACGTT GAAAATATTT	1560
40	GATTTTGGAA TTGCTAAAGC TTTAAGTGAG ACGTCTTTAA CTCAGACTAA TCATGTGTTA	1620
	GGTACTGTGC AGTACTTTTC GCCAGAACAA GCAAAAGGTG AGGCAACGGA TGAATGTACA	1680
	GATATTTATT CTATAGGTAT kGTGTTATAT GAAaTGCTTG TTGGTGAACC ACCCTTTAAT	1740
45	GGAGAAACTG CAGTTAGCAT TGCGATTAAA CATATTCAGG ATTCTGTGCC AAATGTGACA	1800
	ACAGATGTAC GTAAGGATAT TCCGCAATCT TTAAGTAATG TCATTTTACG CGCTACAGAA	1860
50	AAAGACnAAG CGAATCGTTA CAAAACAATT CAAGAAATGA AAGATGATTT GAGTAGTGTT	1920
	TTACATGAAA ATCGAGCGAA TGAAGATGTC TATGAACCTG ATAAAATGAA AACGATAGCG	1980
	GTACCTTTGA AAAAAGAAGA TCTAGCAAAG CATATTAGTG AACATAAGTC GAATCAACCT	2040

55

AGCCAGAAGG TACGGTGTAC GAACCAAAAC CTAAAAAGAA ATCAACACGA AAGATTGTGC 2160
TCTTATCACT AATCTTTTCG TTGTTAATGA TTGCACCTGT TTCTTTTGTG GCAATGGCAA 2220
5 TGTTCGGTAA TAAATACGAA GAGaCACCTG ATGTAATCGG GAAATCTGTA AAAGAAGCAG 2280
AGCAAATATT CAATAAAAC AACCTGAAAT TGGGTAAAAT TTCTAGAAGT TATAGTGATA 2340
AATATCCTGA AAATGAAATT ATTAAGACAA CTCCTAATAC TGGTGAACGT GTTGAACGTG 2400
10 GTGACAGTGT TGATGTTGTT ATATCAAAGG GCCCTGAAAA GGTTAAAATG CCAAATGTCA 2460
TTGGTTTACC TAAGGAGGAA GCCTTGCAGA AATTAAAATC GTTAGGTCTT AAAGATGTTA 2520
CGATTGAAAA AGTATATAAT AATCAAGCG 2549

(2) INFORMATION FOR SEQ ID NO: 626:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2286 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 626:

TGCTTACTTC GCCTTCAATA CGTACTAATT CATGTCCACA ACTTGGACAA TGGGTTGGCA 60
TATGATATGT GACAGCATCC TCAGGTCTAC GTTCTGGAAT ACTACGTACA ACTTCAGGTA 120
30 TGATGTCACC TGCTTTTTTC ACTACAACAC TATCACCAAT TCGAATATCT CTGTCATGAA 180
TTAAATCCTC ATTGTGCAAA GATGCTCTTG aTACAGTTGT ACCAGCTACT TTTACTGGTT 240
35 CTAAAATAGC AGTAGGTGTG ACTACACCTG TTCGTCCAAT ACTTAATTCA ATATCTAATA 300
ATTTAGTTAC TACTTCCTCA GCTGGAAATT TATAAGCAAT GGCCCATCTA GGAGATTTTT 360
GTGTGAATCC CATCTCATCC TGTGATCTA AATCATTAACT CTTAATAACA ATCCCATCAA 420
40 TATCATAAGG TAATGACTCT CTTTGGCTTG TCCATTTTTT AATATACTCT AAAACACCAT 480
CGATATTATT TACACGCGCT CTATTTTAT TCGTTGTAAA ACCTAATTTA TCTAACTCAT 540
CTAATGCTTC ACTTTGCGAA CGCGCATTGA AATCAGTGAA ATCATTGACA CTATATATAA 600
45 ATACGCTTAG CTTTCGTTTT GCCGTTAATT TAGAATCTAA CTGTCTTAAT GATCCCGCAG 660
CAGCGTTTCT TGGATTnGCA AATAACTGCT CATCATTTTT TTCTTTTTCT TCATTTAATC 720
50 GTAAAAATGA ACGTCTCGGC ATATATGCTT CACCACGAAC TTCTACATTT AATGGtTCIT 780
TCATTTTCAA AGGTATCGCA TGAATTGTTT TTaAATTTTC GGTAATATCT TCACCTGTTG 840
TTCCATCACC ACGTGTTAAA CCTTGaACGA AGTATCCATC AACATATTTT AATGATACTG 900

GTTGGTTCGAA TTTTCTCAAA TCATCCTCAT TAAATGCATT CCCTAAACTT AACATTGGCG 1020
 TGTTCATGGTT GACTTTTATTG AAAGAGGCTT GGGCTTCACC GCCAACTCTA ACTGTTGGAG 1080
 5 AATCTACAGT CTTATACTCA GGATGCTCCT CTTCTATTTT AATCAGTTCA TGAAGTAATT 1140
 TGTTCATATTC ACTATCTGGT ACAGATGGAT TATCCTCTAC ATAGTATTCA TAACTGTATT 1200
 GATTTAATAA ATCATGTAAC TCGTTCACAC GAGACGATAA ATCAGCCATC CCTTAATCCT 1260
 10 CCTTTTTTtC AATTGGTGCA AATTGCGCTA ACAAACGTTT TGGCCCTTGT GATTTAAAGA 1320
 TAATATCTAG TTCGATTGAG CCATTTTTTCT CGTTTACATT ACTCACCATG CCTTCTCCCC 1380
 15 AGGCTTTATG CATCACTTTG TCACCTACAT TCCAATCAGA TGACAATACT TGTTTTTTCG 1440
 TTGACGTTGT TCGTTGACTA AATCCGCGTT TAGCAAAAGG TTTTGCCTTA GGTGTATCG 1500
 TTTGTCGTTT GCCACTTGAA TGATTTTCTA ATAGTGATTC TGAATTTC TTTAAAAATC 1560
 20 TGGATGGCAT ATTTGACTGA GGGCGACCAA ATAACATTCT TGATGTCGCA TGAGTGATAT 1620
 ATAACACCTC TTCAGCCCTT GTAATTGCTA CATAACAAAT ACGACGTTCT TCTTGCATTT 1680
 CATGATCATC TTCACTCTTA ATCGCTCTAA TATGTGGGAA TAAAGATTCT TCCATCCCCA 1740
 25 TTATAAAGAC AATTGGAAAT TCAAGGCCCT TAGCCGAGTG CATCGTCATT AGTGTTACGC 1800
 CATTTTCAGT ATCTGCCTCA TCAATATCAG CTAATAACGA TAAATCCGTT AAAAAGTTAA 1860
 TTAATGACTG TTCTTCTAAT GGGGTATTTT CCTCATAGTC TTTTGGTACT GACATAAATT 1920
 30 CATCGATGTT TTCTAATCTA CTTGAGATT CTAATGTATT TTCACGTTCA AGCATTTCTC 1980
 GATAGCCAGA CTTTTGTAAT ACTTCATCAA CTATTTTCATG AATTTCTAAA AATTCTTGTT 2040
 35 CTTTTATCAA GCTTTGGATT AACTCGTAAA AATTAAGACA CTCTTGTC ACCTTTTTTTG 2100
 aCAATCCGAT AAAATCAGCT TCTCCAAGTG CATCAAACAT ACTGATATTG TTTTGAAGTG 2160
 CATAGTTTTG AACTTTTTCA ACAGATGAAG GACCTACACC TCTTTTGGGA ACATTaATAA 2220
 40 TACGTTGCAA ACTAATGTCA TCATTACTAT TGGCAATTAT ACGCAAATAA CTTAATAAAT 2280
 CTTTGA 2286

(2) INFORMATION FOR SEQ ID NO: 627:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 627:

TGACATTAAA AGTTGGGTGA AnAAACAAGG CATACTTTAG TTAAGCTTGA TGAAAATAAC 120
 AATGGAATTA ATGCGATTAT TCAAAAAGAA AAAGCAAAAG ATTTAGATAT AAATTATTCT 180
 5 GCTAAAGGTA CTACCAATTG TATTATTTAG TGGAGAATTA GACAAGCTGT AGCAGCGTTG 240
 ATTATTGCCA ATGGTGCTAG AGCTGCTGGA AAAGATGTAA CTACCTCCTT ACTTTTTTGGG 300
 GGCTTATGCC nTTAAAAAAG TGCCACCGTT AATGTTAAAA GCAGTTGCCA AAATGTTTGA 360
 10 TTATGTTGC CCCAAAAGAT TTACGATGCC CCTTCCCAA 400

(2) INFORMATION FOR SEQ ID NO: 628:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 453 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 628:

AATAATTTGG GCGCTnTTTG CGTCGGGATA TTATACCGCT TCCTTAATTG TTCAACATTG 60
 25 TAATCACTGT TTTTCAATTG ATATTTTGCA GAGTAAATTG GTACTTCTGG GTTATATGAC 120
 ACTTCGTCCT CTTTATAGTT TTCCAATTCT TTGAAATTCC CGTATTGTAC AAAGAAGTTA 180
 AATTCTTCGA TTTCTTTTTT TACTTTTTCG TCATCGATTG GTTTTAATGG AATCATTTTA 240
 30 TTAKTTTCCA TTTTCACAGG ATATCTTTTT GTATGATTGT GTGTCAATCC ATCGCTATCT 300
 TCAACAACTT CTCTAACAAT ATAATGCCCT TTAGCCGTTT TAGTATTCTT GTTAATTTCT 360
 AAAACTGCTC CTCTrGATTG CAGATTTTCT CCTTTTAATT GGATTTTCAT TTCAGATCTA 420
 35 ATTAGcCAAG TACCTTTATC ATCTtTTTTA AAT 453

(2) INFORMATION FOR SEQ ID NO: 629:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1221 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 629:

TCATATGCAT TTGCAAAATA AACGCCAGAA GCAAGGTTTA GAATTGGGCC GTCCGTTTTG 60
 50 CTCaATTAC TTGcATTCAA TAATTATGCT TGATCATGAT CAATTGCTT ATCTAATTCT 120
 GCaATTTTCT TCATTGCTT ATCTGATTG TTTTCTTTCT CCATCATTG ATCAGACGA 180

5 GCTAATGGTA CTAATACTTT ATCTCCATAT GTGTCAATAA AGTTATAAAA ATAATCATCT 300
 GTTTTTGATA CAAATCCAGC ACGCTCTTCA GTTTCACGAT ATAAATCTAA GAAAAGATTG 360
 10 AACTCATCAC GTTCAAGGAA TCTGACTTTA ACACCATAGT TTATCGCTTT ATTAATATTA 420
 CGTTTACGTT GACTATCAAA TGTCTTTTTC AATGTTTCGG GTGTTTTIACC TTCAAGGTTT 480
 AATACGCCCA TCCATCGTAC TTGGCTCGAT GTATCATACT CAGTTGTAAA GCCATGATGC 540
 TCGTAACCAT GTGATTTAAA CAAGTTTACT AGGGCATCAT TTTTCTCGCG ACCTTCAAAT 600
 GGCACGATAT CTTTATCATA TAGATGATAT AACCAATACG GATCTAATTT AACATATAAA 660
 15 CATTGATGTT GCTGTAAATA TTTATCTAAC TCTTTTAAAT AATAATCAAC TAATCCTAAA 720
 TCTGAAAAAT CCATTACTGG ACCACGATTG GAATAGTAAA CATAACTTCC CATAGTAGGA 780
 ATTTTAGAGA AAAGGCTTGC TGCAATTACT TTGTTATTGT CGTCTTTAAT ACCTAATAAA 840
 20 ACTACTTCAA AGCCATCATT CTCACGGGTA ACTATATTTT CTTTACTTG GAAATAATGA 900
 CTTTCCAATG ATGGATTTTG TACAAAGTTG TCAAATTCGG TAACAGTTAA CTCTGTAAAT 960
 TTCATGTTTT GATAATTCCT TCCTAAAAAA TTCTGTCTTT AACTTTTTTA AGTGCGGTAT 1020
 25 ATGCTGCGTA AACAGGTTTA TTAATTGGTT TAATAAAGTC ACCAACATAT TCmATAATTT 1080
 CAGCATTGTA ACCTTTTTTG AATTAACTA CACCAGCATC TTCAGCATCT TCTGTAAATT 1140
 TrCCACTAAC ACCATAGAAA TTATAACGGT CAATGCATGA TTTAATGCAT AATTAATCAT 1200
 30 TTCCCATGTC ACTGCATAAC T 1221

(2) INFORMATION FOR SEQ ID NO: 630:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1121 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 630:

45 TGGCCCAAT AACACAAGTA ATTGCTGCTA ATGGTAACAT CACAAAAAAT GAAATCGTAA 60
 CTACAAATGT TAAACCTTGG AATACACCAA CCATTTCTGG TAATCGTTTA CTATAGTATC 120
 TATTGTGAAT CCAAGTAATA ATAGCTGAAA TAATAATACC ACCTAGAATA TTCGTATCCA 180
 50 ATGTGGCAAT ACCTGCAATT GATTTTAAAC CAGGTACATT TTCAACGCCT TTTTCTAAAT 240
 TAGCGCCAAA CGTATGTGGC CATTGTGTTA AAATGGCATT TATAAATGTA TTAAACATTA 300
 AGTAACCCAT CAATGCTGCA AGTGCTGCAT GACCTGGTGC TTTTITAGCT AAAGAAAGTG 360

55

TTA CTGACCA AAATTTAAAC CAAAACGTAT GTTGATCTGC TAAACTCCCC ATGATTGTAG 480
 GATTTTTTAAA TAATGTCGCA AAGCCAAGCA CTATCCCAAA GAAAGCGAAC ATTAATACCG 540
 5 GTACAATCAT TGCCTACCG AAACGCTTTA TCGCATTCTAT CTTCTATTCC CTCCATATCA 600
 TCTTTCTTAA CAATACATCT AATTAGATTC ATTTATAAAT AGATGTCTTA CTATTTAAAT 660
 ATAATATATA GTAAACGCTT ACACACCTAC AAGCAGATTG ACGTATTTTG AAAGTATTTT 720
 10 GTATAATCAG ATTATCTTTT CATATAGTGA AAATTTTTTC ACGACCTTAT ATATGACATC 780
 GTTGATTTTG TAATACATTC GTTTTAAACG CATAATCAAA CCTATATCAA TACACAAATA 840
 TATATAATGA CATAAAGAT TTTAATGTAA TAACGATCTA TTACACATTT ATTTTCAAGG 900
 15 AGGTTGAATA TGTTTTTAGA TGAACACATT AATCGAACT TTGATAAACT TAATGATAAT 960
 GATTTACATA TCGCTCACTT TATCAATACA CATATAGATG AATGTAAAAA TATGAAAATA 1020
 20 CAAGATTTAG CGCAATTCAC ACATGCCTCG AATGCAACCA TTCACAGATT TACACGCAAA 1080
 TnAGGTTTTG ACGGTnATAG TGGATTTAAA TCGTACCTTA A 1121

(2) INFORMATION FOR SEQ ID NO: 631:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4005 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 631:

35 AACCTTCCAT TTTACTTGAT CGATAACATC AGTTCTGCTT TACAAATCTC TTCATTAATA 60
 TCGCTCTTAA AACCATGAAA TTAAACATGT TCCGATAAAT GATAATCTTC TACAAGTTGT 120
 CGATATTCTG ACAAACCATT TCCATGTCCA TAAATATTCA ATTGAATATT GGGATGTTTT 180
 40 GTTACTAATT GCTTGATTAC TTCAATTGTA TGTTTAATTT GTTTATTTTC AACGAGGCGA 240
 GCAATTGATA TGATATGATT TTTCTCCTTT TGATTGATGT CAAATTGATA CTTTAAATTT 300
 GCCACGTAGC CAACCGGAAT ATTGATAACT GGTATTTTAT TTTCAATATA TTGTGAAATA 360
 45 TCTTGGCATT GCTTTTCTGT TGATACAACA ATCGCTTTAT AACGTGTTAA ATTATTAAAC 420
 ACTGTTTTAT AAAAATTTT TATACCATT ACGGCACCGG ATAAATGTGT ACTATGGAGC 480
 ACAACAATAA CTGGAATACT TTGATTTAAT CCCGCTATAA CATTTCTTAA TTCATGAGGA 540
 50 CGATCTAATA TGATTTGATC ATTATTTTAA CATAATTGAT GGAGAAAATA TTGAATAAT 600
 TCATCTTCTG TATCAAAAAA TTGTTGATGC TGGTCTTCAT TTAAGATAAC CTTTGTGAGC 660

55

	TAGTAATTTT CGAGTACAAT CCGTTGTCCT TCACCTAAAA TTCGAGAACA ACTTAAAAAG	780
	CCTCTTCCAT CATACAATTC GCGTTTTACT TTTCTTCTTT TATGATCAAA ATAATTCACA	840
5	TAATTTAATT GATGATACTG TTTATCTAAA AAATGAGCAT ACATTACAAA TTGCTCTTCA	900
	TCATATATTC TGACATCATT TGAATTTTCC ACAAATTTCA ATGTGTACCT ACATGACTTT	960
10	TCCCAATACT GTATCCAGTT AACTTGCTTT GTCTTTTTAT AATTGATTGC TTTTGA AAAA	1020
	TAGTCATACA TTGTAAATAC ATCATTTTCA ATCTGATGTT GCTTCGCATA TGTGTATGAA	1080
	TAAGGATTCC ATTTAACATA TACACATTTT GAAGATATGC CGTGTGTTTT GAACAACTTC	1140
15	AATCTAATTA TTTGCGCTTT TTCTACACCT GTAATTTTAC TTTCTAAAAT TGTTCTTAAA	1200
	ATGTAATTCA TATTATCGCC TCATATAAGT TTTATTCCGT ATCTTTATTG TTTATTTTAT	1260
	ATGAAAAATA CATCTATTGC ATGTGTAATT ATAAAAAAC CAGGCCACAA GGACCTGGGT	1320
20	CATATTGTAT TATTTGTTTT GTTTTTTGCG ACGACCGAAT AACAATAATG AACCTAATGC	1380
	TGCAATAAT CCACCAAATA ACGTTGCGTT ATTTGAGCCG TTATTTTCAC TACCTGTTTC	1440
	TGGTAATGCT TTTGCTTTAT TGTGATGGTC TTTAGTAGTA CTCATTGGTT TAACAGGTGT	1500
25	ATGTTTTCTT GCATCCGAGT CTGAATCGCT GTCTGAATCA CTGTCTGAGT CTGAGTCGCT	1560
	ATCAGAGTCT GAGTCGCTGT CCGAATCTGA GTCGCTATCT GAGTCTGAGT CGCTGTCTGA	1620
30	ATCTGAATCA CTGTCTGAGT CTGAGTCGCT ATCTGAGTCT GAATCGCTGT CTGAATCTGA	1680
	GTCGCTATCT GAGTCTGAAT CGCTGTCTGA ATCTGAGTCG CTATCTGAGT CCGAATCGCT	1740
	ATCTGAATCT GAGTCGCTGT CTGAGTCTGA GTCGCTATCT GAGTCTGAAT CGCTGTCTGA	1800
35	GTCGGAATCG CTATCTGAAT CTGAGTCGCT GTCTGAGTCT GAATCGCTAT CTGAATCTGA	1860
	GTCGCTATCT GAGTCTGAAT CGCTGTCCGA ATCTGAGTCG CTATCTGAAT CTGAGTCGCT	1920
	GTCTGAATCT GAATCACTGT CTGAGTCTGA GTCGCTGTCT GAGTCTGAGT CGCTGTCTGA	1980
40	GTCACATCTT GAGTCTGAAT CGCTGTCTGA TGTATCTTCT TCGAAGTATC CGTTATCAAG	2040
	TGTGAAATCA TCATGATCCG TAATTGTTAC GTCAACTTCG CCACCATCTG CATCTTTATC	2100
	ATCTTCAGTT GTATTGTAA CTGTTGTGT TAAGCCAGCA GGCTTTTCAA AAATAACTTT	2160
45	GTATTTACCG CTATCTAAAT TATCAAAGCA GTATTTACCA TTTTCATCTG TTTTAGTTGT	2220
	TCCAATTACT TCGCCTTTTT CATTTAATAA AGTAACTTTA ACATCTTTGA TACCTTTTTTC	2280
50	AGTTGAATCT TGTTTGCCGT CTTTATTACT GTCGTACCAA ACATAATCAC CTAAACTATA	2340
	TTTTGGTGTT TTATAGAAAC CACTGTCTAA TGTGATGTTA TCTGCATCTT TAATGACACC	2400
55	TGTTGTTGTT AAACCATTAG AATCTTTTTT AGTATCATTT CCAGAAGTTA CTGAAGTTGG	2460

	TTGATATTTA CCATTTTCAT CTGTTGTAAC TGTITTTTAAA ACTTTGTCGT TTTCATCTTT	2580
	TAACGTAAGT GTTACACCTG AAATGCCCTT TTCATCTTTA TCTTGAACAC CGTTTTTATT	2640
5	TGTATCTTCC CATACATAGT CACCTAAGTT GTAAGTCGGT TTGTAGAAAC CAGAGTCAAT	2700
	AGTATCGTTA TCTTTATCTT TAATGACACC TGTGTGTGAT GTACCATTTG AATCTATACC	2760
10	TTCATCAGTT CCTGAACCTA CTTGTGTTGG TGTGTAACCT GATGGTGTIT CGAATTCAAC	2820
	TTTATAAGTT CCATTTTCTA ATCCAGTAAA TTGATATTTA CCATCTTTAT CTGTTTTAGT	2880
	TGTTTGTAAG ACTTCACCGT TTTCAATTTT CAATGTAAGT GTTACGCCTG AAATACCTTT	2940
15	TTCAGTTGAA TCCTGCTTAC CATCTTTATT TGTATCTTCC CATAcataat TACCTAAATT	3000
	ATATTTTGGT GTTTTGTAGA ATCCACTATC TAATGTCATG TTATCAGCAC CATTAAATAAC	3060
	ACCTGTTGTT GTTAAACCAT TAGAGTCTTT TTCAATGTCG CTACCAGATG TTAGTGTAGT	3120
20	CGGTGTATAG CCTTCTGGTG TAGTAAATTC AACTTTTATAA TTACCATTAT CTAAATCAGT	3180
	AAATTTATAT TTGCCATCAG CGTCTGTTGT AACTGTTTTT AaCagTTACC GTTTTCATCT	3240
	TTTAATGTTA CCGTTACGCC AGATATACCT TTTTCATCTT GGTCTTGGAT ACCATTTTAA	3300
25	TTTGTATCTT CCCAGACATA GTCACCTAAG TTGTATTTAG GTTTGTAAAT ACCTAAGTCT	3360
	GCAGATAAGT TATCTTTGCC ATTAAGTGTATAAATAAGATAAGATAAGCCGTT TGAATCTAAT	3420
	TCTTCGTTAT TACCTTGTTT TGAAGGGGTT ACTTCATAAC CTTTGGTAA GTTTGAAAAT	3480
30	TCTACACGGT AATCTCCATT AGGTAAGTTT GGAATCAAGT ATGACCCATC TTCTTTAGTA	3540
	ACTGCTTCTC CTACTTTTGT ATTTGTATTA TTATCAAATA CAGTTACAGT TACATTGCCA	3600
35	ACGCCTTTTT CTCCTAATTC TTGAACACCG TTTTATTAG TATCTTCCCA TACGTAGTTA	3660
	CCAATTTTAT ATACTTCTTG ACCAGctCCG CCACTTTGGT TATTAGTAAA TCCTAAAGCA	3720
	TTGCCAGTAG AAACGGATTT ATTACCTGTT GAAGATAAAG TAGCCATTTG AACAAGTGTT	3780
40	GGGCTTTTCG TATTTGTATA TTGGAATTTT GTATTAACCA TTACAACATA AGCAGAATCT	3840
	GCATTTCCAA AATCAATAAC AGCGCTATTG TtGTGCGCCAT ATGTAATTTT CTGCAAGTAT	3900
	TGATTTGTTA CATCTGTAAG CTCTTTAGTA TTCACATCGT ATCCTTTATT TAATGTATAA	3960
45	CCTTTAGGAA CTGTATATAT TTTTATATCT GTTACATCTT TATTT	4005

(2) INFORMATION FOR SEQ ID NO: 632:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1440 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 632:

5	TATTAGGTTA CTCTAGTTTC CAAAGCGGGA ATTTTAATGT TATTAACAGC AAGGACAGCA	60
	AAAGCAATAT CGGCGCATTG ATTGAAAATC CAGGAATATA TCCTTTTATG TCTGGATATG	120
	AAAACCTTGAA GTTATTGAAT GAATCAAAAA ACACTCAAGA TATCGATAAA ATTGTCTCAC	180
10	AACCTTCATAT GGATGAATAC ATTCATAAAA AAGCTAAAAC GTATTCTCTT GGTATGAAAC	240
	AAAAATTAGG AATTGCTATA GCATTTTTTAA ATAAACCTCA ATTCATTATC TTAGATGAAC	300
	CAATGAATGG CTTAGATCCA AAAGCTGTGC GAGATGTACG TGAATTGATT GTCCAAAAAG	360
15	CGCAAGAAGG TGTTACTTTC TTAATTTCTGA GTCATATTTT AAGTGAATTA GTTAAAATCA	420
	CAAACCTCTAT CCTTATTATT AACAAAGGTA AAATTGTTAC AGAAACATCG GAAGAAGAAC	480
	TTAAACAATT TAAAGATAAT GATTTAGAAA ATGTATTACT AGAAATCATA GAAAGGGAGG	540
20	ACCAAGCATA AAATGGGAAC TTTAATTAAA CAAGAATGTT TCAAATTATT TAAAAAGAAA	600
	TCAACTTTTA TCGCACCTAT TGTCTTTAAT CTACTAATGG TTGCTCAAGG TTATATTGCT	660
	ACAAAATACA ATGAAATTTT TACGCCACAG GAATCTTTCA CATCTGCTTA TAATGGTTTT	720
25	TCATGGTTTG CATTTTTTATT AATTATTCAA GCAAGTACAA TCATTTCAT GGAATTTTCA	780
	TACGGTACGA TTAAAAATTT ACTCTATCGT GAATATTCAA GAACAACTAT GATTGTTAGC	840
30	AAAATCATCA CATTATTTAT TATTTCTTTA ATTTATTTTG TTATTACAAT TATTGCTTCA	900
	ATTGTTATTG GGTCTTTATT CTTAATGAT TTAAATATAT TTGAAAGTAG CGGTAATCAA	960
	TTATCTTTAT TGAATCAATT ATTATTAGTT AGTTTAGGCA CATTGTGTTG CGTTTGGTTA	1020
35	GTTTTAAGCT TAACGTTGCT ATTATCATCT GCAACAAATT CAACGGGAGT AGCCATTGCT	1080
	GTAGGTATTG TTTTTTATTT TGCAAGTTCT ATTTTAGCAG TTATTCAAAC GGCACITTTA	1140
	GAAAAAATAG ACTGGCTAAA GTGGAATCCT ATTAATATGA TGAATATTAT GCTTCAAACA	1200
40	GTTGAAAAAG GCTTTAGTAA GTCGACAAAA TTAGAACTTC ATGAATTGTT TATTGGTAAT	1260
	ATTGCTTATA TTTCTATTTT CTTAATACTT GTAGTATTTA TTTTCAAGAA GAAAAATATT	1320
	TAGTAACTTA AAGTATTAAA TGTCTAAATA CACACATATT CCATCGTAAT TCAAAATCAT	1380
45	TTTCAAATCC CTTCACCCAA ATAATGGTGC GGGGATTTTT TCATCCAAAT TTTGGAAATT	1440

(2) INFORMATION FOR SEQ ID NO: 633:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1323 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 633:

	GCTGACATAA TTGCATCAAA TTTACATCC CCATAAAAT CGCTACCACA TAACCTACGA	60
5	TAATACCTAC AAGAACTGGA ATTAAAGATA GGAATCCTTT AAAAAATCCT TGAACGACTA	120
	TTGTTACAAG CAAGGTTATC ATTGCAACAA TTAAGAACT GATATTGTAA CCTTTCATAT	180
	CTCCAGGATT TTCATACATT GCCATATTGA CTGCAGTAGG CGCTAAGCTT AAACCAATTA	240
10	CCATGATGAC TGGTCCAACA ACAACTGGTG GTAATAATTT CATTAAACCAT GCTGTCCCAC	300
	TTAATTTGAT TAGAATCCCG ATGATGACGT ACATAACACC ACTCATGAT AATGCTACAA	360
	GCATGTCTCC TAAGCTATGC GTACTTAATC CCGTGATAAT TGGCGTGATA AATGCAAAGC	420
15	TAGATCCCAA GTATGCTGGT ATTTGCGCCT TCGTTATTAA GATATAAAGT AATGTACCGA	480
	TTCCCGAAGC TAGTAACGCT GCTGATATTG GTAGTCCTGT TAAGAATGGT ACTAGTACTG	540
20	TTGCGCCAAA CATCGCAAAT AAATGTTGTA AGCTTAAAAA TGCCCATTCG GCTGGTTGTG	600
	GTTTTTCATT TACATCTAGT ACGGGTTTTA CTGTTCTGTC AACATTTCA TCATTTTGCA	660
	TAATATTCAT TTCTCCGAT AATAAAAAAA TCTCTTTACA TCAGTATATG TAAAGAGACA	720
25	AAAAGTGTGA CAAGTTGCTA CAAGTCATTT TCGTCCATAG AAATTGACTT ATAGTTGTGCG	780
	AACATGAGGG TATTATTAGA TAAACAAGCA TATGAAAAT TATTTATCAT TCAACTCCCC	840
	CACCTTTTTTC AGTCTCTCGT ACTGAATTAA AAGGGGtATT ATTTAATTAT AACTGCATTT	900
30	CTTTGATCCA TTtCTTCyAA ATAGACACTT ACCGTTTCCT CTTTAGAAGT AGGTAWATTT	960
	TTACCAACAA AATCTGCTCG AATTGGTAAC TCACGATGTC CTCGATCAAC CAAAGCAGCT	1020
	AAACCAATTT TAATAGGTCT AGCATTTAGC AAAATAGCAT CAAGTGAAGC ACGAACCGTT	1080
35	CGACCAGTAT ACAGCACATC GTCAATAATG ATGACTACTT TATCTGTAAT ATCTGTGTCG	1140
	ATGTCTATTG CGTCTTTTGT CGTAAGTGAT GACATGTGcT CTATATCATC TCTAAAGTAT	1200
40	GTAATATCAA TTGTTCCAGT AGGTATACGT TGTTGCTCAA TTTGATGAAT TTtATCTTGT	1260
	ATACGATTCTG CTAAATATTc ACCTCTTGTh TTGATACCTA AAAGATTAAA TTATCAGTAC	1320
	CTT	1323

(2) INFORMATION FOR SEQ ID NO: 634:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 761 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

ACCCATCTCA TCGTATTTTG AATTTAATAG ACGAGATCGA TGTATATCTG AATTCATCCC 60
 AACTATGGAT TAATGTTGGT ACATCATtAA cGCATAACCA ACATTTTGAG CAGTTGTTTT 120
 5 ATAAGTAACG TGATTTTTAT CTAATTGCCC TCTTAATGCG TCCTCTGTAA ATTCAACACT 180
 ATCAGAACCA TTAGAGGTCG CTTCATATAA GTTATTAGAT GCAATATGTG CTAAATCGCT 240
 ATTGATTTTC AATGGTTTTA ATCCTTTTAA TTTTCTCATT TCATTCGTTA CTTCATAAAG 300
 10 AGAAATTAAT TGATTGGAT TTTGCTCAAC TGGACGCTTA TTATGCTCTT CTGACGTAGA 360
 ATTAGAATTT AATTGATAAG GTTCAATATC TGCTAACATT TCTTTTGTTA AAAATCGTAC 420
 ACTTAGCACC TTTTTCGATT GTTGATCAGA ATACACTTGT GCATATATGT CGCCATATTT 480
 15 AATCaGTGTT TGTGTTTTTA AATCTTCATC TGAAAGTTCA AATTCATATT TTTTACCATC 540
 AACTTTAAAG GACGGTCTG GATTAATACT TGTATGATTA AAAATTTCTG CAGAATGTTG 600
 20 TCCTATTTTT AACGGACTAA CATTGACTTT CTCACCTGTA GCATACACTG AAACGATTTT 660
 TTCACGTTTA GTTGAAACAA TGTAATAACT GTTTTTGTCT TTAAACACAT AATTTTTGTA 720
 rCCATCTCTA AAAGGGTAGA CrCGATCTGC TTGTCCAAAT T 761

25 (2) INFORMATION FOR SEQ ID NO: 635:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 827 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 635:

35 nTAAATATAT TTATATATTA TAGAATAGAA AGACCTGAAG ATTGAATATC TTTGCGAAAG 60
 CCTTTAACTG TATCTACTGA TAATTCGTTA ATATCGCGAC CTAAGTTTGT ATTCACTTTT 120
 40 TTCACAACAT CTGCTGGGCA TGTAATAATA TCTGCACCAA TTTCATCAGC TTGAATCACA 180
 TTGAATAATT CGCGGCAACT TGCCCATAAAT AATTTAACGC CGTCTTTACT ATGCGTAACT 240
 TThACAGCCT CTkTCATTAA TGGTaATGGA TCTACGCCTG TAtCTGCAAT ACGTCCTGCA 300
 45 AATACTGAAA CATATGTTGG CACACCTTCA GTTACTGCTT CAGTTATTTT TTTAACTTGT 360
 TCAATTGTGT AAACAGCCGT AACGTTTAAAT CTCACATTGT CAGCTGAAAG TTTTTTAATT 420
 AAAGGAATCG TTGATTCAAC TTTTGTATTT ACAATAGGAA TTTTAACAAA TACATTTTCG 480
 50 CCATATTGTT TTAAAATTGC TGCTTCTTTk TCCATAGTTT CTAAATCGTC TGCAAATACT 540
 TCAAATGArA TTGAAGCATC TGAATTTCT TTCACAGCTT CTTCAGCAAA AGCTTTGTAA 600

55

TTTTTATAAG CTGCTTTTCAT TnCTTCAATA TCTGCACCGk CCGCAAATAC TTCTACATTT 720
 AGTTTAGCCA TATAAyATAG CCTCCTTGAT TCTTATTAAA ATTTTAACAA CATCTGCATG 780
 5 KCTTTTCTT ACAACCATT GTAAAAATG ATTTTATTT CTTtGTT 827

(2) INFORMATION FOR SEQ ID NO: 636:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1478 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 636:

TTAGTCGGTA TAACCATCGG CTAGGTGGTT TTTGTATTAA AAAAGTGGAT aCCAAAATTT 60
 20 ATTAATAATT ATTTTAATGT TAGAAAAAAA CTAAATAAAA ACTCGCTAAT GATATCCAAT 120
 AATATGTATA CAAAACGAGA CATATATTGC ATATGATTAA CGAGATACTG AAAATATTTT 180
 ATCACCCTTA AAATGATTAT TCATTTTCAG CGGTAATTCG ACCTAAAGTC AAACCTACAA 240
 25 TAAACCGAT GATAAATACT ACTAATGAAA CGAACCACAT CACGATATTA GTTGGTAAAC 300
 CTGGAAATAC TGCAAAGAGG GAGCCAACAA CAAAACCAAT GATTAATGCA AAAGTCATTA 360
 GTTTATGATG TGTTAGGAAA TACTGGATAA TTTTGCTTGA AATAATGAAT CCAGCAAGCA 420
 30 CGCCAAATCC GACTGCAAGT AATATAGGAA GACCTGCAAA GTTAAGTTTA ACAACTTCAG 480
 ATATTGCTAG CATGACCGTA CCATAGACGC CAAATACTAA TAACATAAAT GACCCTGAAA 540
 TACCTGGGAG TAACATAGCA CTAGATGCAC ACATACCTGC AATAAAATAT TTAATAATAA 600
 35 GACTAGTTGA TAGAGTAAGT GTTCTCCAG CATGTTTATC ACCATTATTC ATTAATGTAA 660
 TAACAATTAA GATAGCGATA CCAGCTATAA CCATCATGTA ATGTYTAGTT GTAAATGACG 720
 40 TTTTATAGTT AGAAATTTTC AATAAATATG GAACGATACC AATGATTAAT CCACCAAAGA 780
 AAAACATAGT TGGAAATATGG TGTGGCTTA ATAAATAATT AAAAAGATTA CTTAGTGATC 840
 CCATTGCCAG TAACATTCCA ATTATAATGG GGATTAAAAA TGTAAACTT GGCCAAAAAC 900
 45 GTCGTGAGAA TATGCCGCTA ATTGAAGCGA TAAATTGATT GTAAATACCT AACAATAATG 960
 CGATAGTCCC ACCGCTAACA CCAGGTACCA AGTCACTCGT TCCCATAGCA AAACCTTTTA 1020
 GAATATTAAT CCATTAAAT TGTGTCATGA ATAACCTCTT TCAAACGATT GGAATAAAAT 1080
 50 CATAAATAGC ATCATACCAT ATTACAAATG TCCTAGTGAA ATGATAACAT ATTTTAAATT 1140
 CATAAATCC ATTGAGAAAT TATGTGCACT TATTATCATT TATATTTTAA AAGAGAGCGG 1200

AGGTATAAGT AAGTTATAAT TAACTGAACG CATTATTACA AAGTCTTTTT GACTACAAAT 1320
TAAAATTATT ATAACTAGT TAAGAAACT TTATATTTTA CGGAGGGAAT ATAAATGGC 1380
5 ATCAACATTA GAAATyAAAG ACCTACATGT GTCTATTGAG GATAAAGAAA TCTTAAAAGG 1440
TGTTAACTTG ACAATTAACA CTGATGAAAT ACATGCCA 1478

(2) INFORMATION FOR SEQ ID NO: 637:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1995 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 637:

20 ATTACAGCAT CTTCTCTAGG TAGATTATTA AAAGATAGAG GTCTAAATGT AACAATTCAA 60
AAATTCGATC CATACTTAAA TGTGACCCA GGTACAATGA GTCCTTATCA ACATGGTGAA 120
GTATTCGTAA nGGATGATGG TGCAGAACT GACCTAGACT TAGGACATTA CGAAAGATTT 180
25 ATTGATATTA ATTTAAACAA GTTTTCAAAT GTGACAGCCG GTAAAGTGTA TTCACACGTA 240
TTGAAAAAAG AACGTCGTGG TGATTACTTA GCGGGAACAG TTCAAGTTAT TCCGCATATT 300
ACAAATGAAA TTAAAGAACG TTTATTACTT GCAGGGGAAA GTACGAATGC AGACGTTGTT 360
30 ATCACTGAAA TTGGCGGTAC AACAGGTGAT ATTGAGTCAT TACCGTTTAT TGAAGCGATT 420
CGTCAAATTC GTAGCGATTT AGGTAGAGAA AATGTTATGT ATGTTCACTG TACATTACTG 480
CCTTATATTA AAGCTGCTGG AGAAATGAAA ACGAAGCCAA CACAACATAG TGTTAAAGAA 540
35 TTACGAGGCT TAGGTATTCA ACCAGACTTA ATCGTTGTAA GAACTGAATA TGAAATGACA 600
CAAGATTTAA AAGATAAAAT TGCATTATTC TGTGACATTA ATAAAGAAAG TGTTATTGAA 660
40 TGTCGTGATG CAGACTCTTT ATACGAAATT CCATTACAAT TAAGCCAACA AAATATGGAT 720
GATATCGTTA TTAAACGTTT ACAATTAAAC GCGAAATATG AAACACAGCT TGATGAATGG 780
AAACAGTTGT TAGATATCGT TAATAATTTA GATGGTAAAA TTACAATTGG TTTAGTAGGT 840
45 AAATATGTTA GCTTACAAGA TGCATATTTA TCAGTTGTTG AATCATTGAA ACATGCTGGA 900
TATCCTTTTG CCAAAGATAT TGACATTAGA TGGATTGATT CAAGTGAAGT AACAGATGAA 960
AATGCAGCCG AATACCTTGC AGATGTCGAC GGTATTTTAG TACCAGGTGG ATTTGGTTTC 1020
50 CGTGCAAGTG AAGGTAAAAT TAGTGCAATT AAGTATGCTA GAGAAAACAA TGTACCATTG 1080
TTTGGTATTT GTTTAGGAAT GCAACTTGCA ACAGTTGAAT TTTCAAGAAA CGTATTAGGC 1140

TTACCAGAAC AAAAAGATAT CGAAGATTTA GGTGGTACAT TACGCTTAGG CTTATATTCA 1260
 TGTTC AATTA AAGAAGGCAC ATTGGCACAA GATGTTTATG GTAAAGCGGA AATTGAAGAA 1320
 5 AGACATCGTC ATCGTTATGA ATTTAATAAT GACTATAGAG AACAATTAGA AGCAAATGGT 1380
 ATGGTGATTT CTGGTACAAG tCCAGATGGA CGTTTAGTAG AAATGGTAGA GATTCCGACA 1440
 AATGtTTCTT TATTGCTTGT CAATTCACC CAGAATTCTT ATCTAGACCA AATCGTCCGC 1500
 10 ACCCGATTTT TAAATCATTT ATTGAAGCTT CATTAAAATA TCAACAAAAT AAATAAATTT 1560
 GCTAATAAAA CCGGTACTTT CATTGTTAAA CATTGAAAGT ACCGGTTTnT CGTATAATTT 1620
 15 TAATATTATG TTAGTGACAA GGTATGAAAT AACAATAGTG ACTTTTATAA TTCTAAGTCT 1680
 CTGTGCATTT CAATCATTTG TGTATAAATG TCATAGTATA CATAATTCAA TGCCATCGCA 1740
 TGTGGTyGGA CAATCTTATC GTAATCTTCA GTGTAGACTA TAGGTcTTGG TGTAGATAAA 1800
 20 TCGATAAAAT GTACGAGATG ATCAGGGAAA TCATCTGTTT TAGGTTTGTT GCTTATTAAG 1860
 ACCACATCGA TATCTAAGTC GATAAGTTTT TGAATATCTA ATGCAACTTG ATyATTATAA 1920
 AATGGTGCGA ATAATAATAC ACGATCAGTT GAGTCAATTT CTTTAAwkTC TTTAATAGCG 1980
 25 TaAGTTTnCG GCTAG 1995

(2) INFORMATION FOR SEQ ID NO: 638:

(i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 1107 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 638:

ATTAGTGATG AAAGTCAGAT TGAAGCTTTA TTAACAGCTG AAAAATATTC AGAAATGATT 60
 40 GGTGAATAAT CACCGTGTA CTCCTTAATC TAAGATTGAG GAGTTTATTT TTAGTCTGAG 120
 AAAATAAATG ATATGAAAGA AAAATTATTA GGTACTATTA TTTGGAGTAT TGCTACATTT 180
 TATTATTCAA GAATGATGGA AATAATGAAT TTAGCTATTT TAAAAATAAA AATTGGGGGA 240
 45 AGTTAATATG CTAAACATTC AAGACGTTAn CATnCTTTCT AAAAAGGAGC AAAAAGCATA 300
 TAACCGTTTC GTAGAATCTG TAGAAAACGG TAATTTACCA GTACTACCAT GTATTGAAAT 360
 50 GGATCTAAAA GAGATGCAAG AAGAAACATT AAACCAGAGT AAGATTGGTG GAATGCCATT 420
 TTTAAAATCT TTTAAAGATA TACCATTAGA TGAAAATAAT GTACCAATGG TATTGCTAGC 480
 ACAGATTAAT TTGGATGATC TTCCAGAACA ACAAGAATTA TTTCTGTAA AAGAAGGGAT 540

55

AAACAATATA AACTCAAGGC TTGTTTATAT AAAAGAGCCA ATTACAGATT TATCACTCGA 660
 AAATATTCAA GCGCATTGGA AGTCATTAGA TGCTGATAAT GAGGATATCC CGTTCAGTGG 720
 5 AGCATTCTTCT ATAGAATTTA GATTGTGCGAA ACAAACCTATT ACATGTACTG ATTATAAGTA 780
 CGATGAGGAC GTGCTTGCAAT TGTGGAATAA AGTCAATCCA TCCTTCGCGC TAAAATCAAT 840
 GTTTGGTGGT TATGATGAAT TGATGGAACC TGTGTGTAmC AywTTTACTG CTAAGgAACC 900
 10 ATTTAATCAA CTTGGTGGTT ATCCATATTT TGACCAAATA GATCCAAGAA CGAACGATCa 960
 AGAACTGAAA ATGTATGATA GAGTCTTACT GCAAATTGaT TCTACAAGAG ATGGTAATTC 1020
 15 TTCGATTATa TGGGGTGaTT TAGGTATTGc CAATATCtTA GTGaAATCTA CTGrACCTTG 1080
 aGGcTatGaa GTTTTGAATG ATTACCT 1107

(2) INFORMATION FOR SEQ ID NO: 639:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 904 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 639:

30 ATTCATATTA TTATAAATTA TTTCTACACC ATCCCAATTG AGTTGTTTTT CATAATTTAA 60
 ATGTAATTCC ACTAACTCCC TACCAATTTT AACAAATCCA TATACATCCT TTAATATCGG 120
 TATTCGCGGA AAACCTTTAC TCAATCACT TGAATATTTG TTCACATAAT ATTTATGATG 180
 35 CAAAATTGCA TATATATAAT ATACTATCTC TTCTGAATTA AGATTTATTT TCTTTTTTAA 240
 AGAATTAGga AATATTATCT ACAAGCCTCA AACTATCTTT ACCTTTGcAT GTAGCAAAGC 300
 CTTkGCCATT ACCAATAAAc TGGAAATTAG GTAATATGTC CGTGATCATA GCCGAGAATT 360
 40 CTTTATTCAT TCCCTGTCCT TGTATATAAA TCACCTGTCC AGTATTCTCC ATTATATTAT 420
 AATATCTACT TGGCATTTC ATAATATTTT TGTCGTACAC TATCCATTTT TTTGTAAATG 480
 GTCTATGCAT AAATTTAACA ATTCTCTCTG GATTAATTGA AATATTTTTT CCTTTAGAAA 540
 45 ATTTTGGGT AAGTCCTCGT GTCCAACtaA TaAATGTTTC ATCTTTGTTC ACTAAATTTA 600
 TACGTTCTCT TGAATCTAAG ATATCAATTA ATCTATCTAT TTCAGAATTA TAGTTATCTA 660
 50 CAAGTAATTT TGcATTTACT AATGCTTTTT CATTCGAAAA ATTTGTTACC CAATTATCTC 720
 TTGCTGaATT TACTCCATTA AATkGAKCTA AATATATAGA ATTTTCAATA TCCTTGGAAT 780
 CATACATTGG TAAATAATTC CCCATAGTTT ATGTCTCGGT GATTAATCCA ATCATTGGGG 840

55

AAAG

904

(2) INFORMATION FOR SEQ ID NO: 640:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 640:

```

15 CGATGTCTTT ACTATTAGAC TTAGCCATTG GTTTCACCTC TCCAAAAATT GTAAATGTGT      60
   ATCATCAATA TGAAAGTTAC ATAAACTGA CATATTTCTT TAAATATCA ACGCCATTGA      120
   TAACTTCCTG TTTTAATTGA TACGCTGTAA CAAAATACTA TAGTTAGTGC TTACATGTAT      180
20 ATGTTAAAGC AAGCAGTGGT AAATGTAAAT TATAATTATT CATTAACTTT GCAATATATT      240
   AAATCTTTTA TTCATAGAAG ATAAATATCA AATCAATCAT AATTATTTGA CAACAAATAG      300
   CTAACGATTG TTTTAATCTA CATTTGGCTT ATAGCATTTT AAACCTATAC TCTATTTTGA      360
25 TACAATATAA GTGTAAAATC AATCATAAAA AGGATATTCA ATATCTGCAT CCAAGAAAAA      420
   CATTACAATT ACCTTT                                     436

```

(2) INFORMATION FOR SEQ ID NO: 641:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 442 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 641:

```

40 GTTATTAAAT TCAGAGTGGT AGCAAATTAA AGTTANTCAA GAGTTAAGAT GAATTTAATT      60
   CATGAACACG TCTATTATTT TTATAATTGT AGCAAATAAA GCTTTACATC AAGGAGGTAA      120
   TTAAATATGT TCAAAAAATA TGA CTCAAAA AATTCAATCG TATTAAAATC TATTCTATCG      180
45 CTAGGTATCA TCTATGGGGG AACATTTGGA ATATATCCAA AAGCAGACGC GTCAACACAA      240
   AATTCCTCAA GTGTACAAGA TAAACAATTA CAAAAGTTG AAGAAGTACC AAATAATTCA      300
50 GAAAAAGCTT TGGTTAAAAA ACTTTACGAT AGATACAGCA AGGATACAAT AAATGGAAAA      360
   TCTAATAAAT CTAGGAATTG GGTATTATCA GAGAGACCTT TAAATGAAAA CCAAGTTCGT      420
   ATACATTTAG AAGGAACATA CA                                     442

```

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2472 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 642:

10 CCAATTTTGG TATGAATTAT ACAGATAATT CnGCGCCCCG AGGATCATT T GCTTATTTAA 60
 ATCAATTCGG TGTGGATAAA TGGATGAATG AAGGGTATAT GGCATAAGGA GAACATTTTA 120
 15 ACTACTGCCA ATAACGGAAG ATATATTTAT CAAGCTGGAA CTTCATTAGC CACACCTAAA 180
 GTTTCGGGAG CACTAGCTTT AATCATTGAT AAATATCATC TTGAAAAACA TCCAGATAAA 240
 GCGATTGAAT TGTATATCA GCATGGGACA TCTAAGAATA ATAAACCATT TAGTAGATAT 300
 20 GGGCATGGTG AGCTTGATGT GTATAAGCA TTAAATGTAG CAAATCAAAA AGCAAGTTAA 360
 TAAATCAAAG GAGTTTTTGA TTATGGCAAA ATTAGTTACT GAAAACATTT CGAAGCGGTT 420
 TAAAAATCAA GATGTATTAA AGCATATTAA TATCACTTTA GAAAATAACG AAGTTTATGG 480
 25 ATTACTTGGT ATTAATGGAG CCGGTAAAAC GACACTTATG AAAATTATAT GTGGCATACT 540
 TCAACAAGAT TCAGGGGAAA TTAAATTAGA TAATAGACCA ATGACACGAA ATGATTTGCA 600
 CAAAGTTGGT TCGCTTATTG AAACACCTGC GACATATAAT CATTTAAGTG CACAAGATAA 660
 30 TTTGAAAATT GTGTGTTTAA ATGAAAGCGT TGATTTCaGC GAAATTAATA GTGTTTTAAG 720
 CTTAGTCAAT TTAAATGTCG ATAAAAAGAA AAAGGTTAAG GACTTTTCTT TAGGTATGAA 780
 35 ACAAAGACTT GGAATTGCAA TGGCaTTAAT TAAAAaGCCA GAAATTTTAG TATTAGACGA 840
 ACCATCTAAT GGTTTAGACC CATATGGAAT CCAAGAACTT AGAGAACTTC TAAAATTATT 900
 AACAGAACAA GGTACTAGTA TTATTATTC AAGTCACATT TTATCTGAAA TCCAAGTTTT 960
 40 AGCAGATCAT ATCGGTATTA TTCATGAGGG TGAGCTAAAA TATCAGCAA GAAATAACAA 1020
 AGATGAAAAC TTAGAAGAGA TATTCTTCAA AATAACGAAA GGTGATTACA AATGATACAT 1080
 TTAAAGATaG AAGGTATCAA ATTTAAAAaT TCTTTCAGTA TGTATGTTTT ATTAATAAGT 1140
 45 CcGcTGGTAT TTCTTTGTTT TGCTATTTTC ACAGTCTTAT TCGCCAAAAG TAATACGGGA 1200
 ACAGCGAATA GTGTGTCACC ATATATAACT TTACTATTTA ATATTTGGCC AATTGCTTTC 1260
 ATCCCGATTG TATTATGTAT GGCTTGTAAT TCGTTATTTA AAATTGAAAT GAGAAATAAA 1320
 50 TCATTTAATT ATTACTTAAG TAATAATTGG TCGATTACAA AAGAAATAAG AGCAAAGATT 1380
 TTCATTTTAT CAATAGCATT TTTGGTACAT TGCTTTTTAG TATTTATTAT TGCTTATATA 1440

TTGATGTATG TAGTATCTCT ACCATTGATA CCGCTCAACT TTTTATTAAC TCGATACTTT 1560
 GGTGTGTTTCG TATCAATATT AATAAACTTA GTATTATCAG TCATTTGTGT CTTGTTTTTA 1620
 5 ACATTGAAGA GTTTATTTTG GGTGTTGCCG TGGGGGATAA TGCAGAGAAT CCCGCTTATT 1680
 ACGCTTGGTA TACTACCTAA TGGcTTAGTT GTAAACCATA ATTCAAAATA CTTTAATGAT 1740
 10 CTCAATGCCT TATATATTTC GATTATTGTT AGCATCATT TTTTCGCGAT AGTAACATT 1800
 TTAAATAATA AGAAAAGTTG GCGATTAAAA TGATAATTAA CGAATTAAAG TCATGTAAGT 1860
 TGAAATTTTC TAAGCAAGCG CTCACATTG TACCCATTAT TGTAACCATA TTGTTTATAT 1920
 15 TATTTATAAA TTGGTATTTA AACGTAAATT TATGGAATGG TCGACAAATm AGTTTGTITA 1980
 CAGCGAGTTT TAATGCAATT ACATCGCTAT TAATTTCTAT AAACGTCTAT CAAGTTATCA 2040
 ATTTTGAAGA AAATATTGGT CACTTTAATC ATATTTTAGG AAAAGCTAAT AGGCTAAATT 2100
 20 GGTAAATGC ATCAATGATT TTTACTTATA CTATTACAGC CATATGTATT CTATTAGCAT 2160
 CAATTAATTT ATTGTGGCAT TCACATGATA TGAAAATAAC ACTTATGTTT ATAGGcGTaT 2220
 CATTGTTTTT CAATGTaATT ATATTACTGC TACTTTTTAT TTTTAGTATT TTCATTAAAG 2280
 25 ATGTAATGGC TATTGTTGTC GGAGTTTTAA TGTTTATTTT TAACGTTTAT TTTGGATTAG 2340
 AAGTGCTTGG AGATCATTCTG TGGTTCTATT TACCAATCAC ATATGCTACA CGTTACGTTT 2400
 ATATGTTTAn CGAAGGGGAG TATACCAGTT ACATTACAT TGGGCAATCT ATATnATTAT 2460
 30 CACGnTGCCG AT 2472

(2) INFORMATION FOR SEQ ID NO: 643:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 646 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 643:

45 ATAATATAGG AATTATTTTCG ATnACAGATT TTACGAATCG TGCTACGATG CAAAATGAAn 60
 ATAAAGATCC ATATGGCGAA AAGTTAGCTT ATGGAATTGC TTTTAATGGC AGTGTGGATA 120
 TGCAAGGGGA TAAACAAGTC ACAATTCCAA AATATAGTGT AGTTACAATT ACTGGCGAAA 180
 50 ATAGTAAAAA TTATCGTGTT ACCGCCGATA ATAAGACTTA CTATGTTAGT AAAGATAAAT 240
 TAGAATATTT TAACCCGGCA GGTTTATATC AAACGCATAG TTTTAAAAAA TTAGCACCAT 300
 ATATGAAATC AAATTATAGT AATTACTATG CATACTTTAA TAGTCAATTA CATAAAAAGC 360

55

CACAACAACC GATACAATTA CTTTTCAATG ATAATAATCA GTTATACGGT TTTGTTTATC 480
 CAATTGTAGA TAAAAAGAA TTAAAGATA AGTTTAATAT TAACAATAAC ATTTGGATTA 540
 5 CTAAAGTTGG GAATGGATAT TGTATTGCCA ATTTGAAAGA AGACAAATGG ATTTATATTG 600
 AATTGTAGGT GTAAAGATGC TAGATAATAT TATTTTATAT TTAAA 646

(2) INFORMATION FOR SEQ ID NO: 644:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 426 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 644:

20 TAATATCGGA ATTTGATAAT GAAGATATCT AATTTTTTAA TATTCGTAGC TTTTATTTTT 60
 CTACTTATTT TATGTTTATT TTTAATCTTA CAAATGACAA ACCATTAAAA GTAGCATCCC 120
 AACATCAAAC AAAAAACAA TTCATCAAAT AAAAATCGCT ACAAACCAA GTCATTAAAC 180
 25 ACGCAATAAT TAAAATTTTC CACTCATTAT AATTCTGAAT TCCAAATGTC GAATTCCGAA 240
 AACCAAACCT CAAATTCCAA AAACGCAACT CCAAATTA AAGCATTTC CTACCATTCC 300
 GGAAATGCTT TTTACATACT GgATTACTCT GTCATTAATG ATTTTACAAC GGgAAACCAT 360
 30 GTCGtCATGT ATGACCaaAG TAGCGTCGCT AtCaTaAgGt GGTTCGGATC TTTATTGGAT 420
 AAThAT 426

(2) INFORMATION FOR SEQ ID NO: 645:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3241 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 645:

45 nTTATTTTACC CACAACATGT TGCACATTA GGTAAATGGG TACCTTATTT ACTTGGTATT 60
 GTTATGTTAG GTATGGGATT AACAAATACA CCTAATGATT TCAAAATGGT CTTTAAAGCA 120
 CCTAGAGCAG TAATTATTGG TGTCTGTCTA CAATTCAGTA TTATGCCCAC ATTAGCATTT 180
 50 ATAATTGCAA AGTCTTTTCA TTTACCACCT GATATTGCTG TTGGCGTAAT ATTAGTTGGA 240
 TGTGTCCGG GTGGGACATC AAGTAATGTA ATGAGTTATT TAGCCAAAGC TAACGTAGCA 300

	ATATATCTAT TTGCAAATGA ATGGTTGGAA GTATCTTTTCG TGAGTATGTT GTGGTCAGTT	420
	GTTCAAGTTG TATTAATTCC AATTGCTTTA GGTATTGTTT TGCAAATTAT TAATCGTAAA	480
5	ATTGCTGAAA AAGCTTCTAC AGCTTTGCCA ATTATATCAG TTGTTGCTAT TTCATTAATT	540
	TTAGCAATAG TTGTAGGTGG CAGTAAGCAC CAAATCTTAA CTACAGGATT ATTAATATTT	600
10	TTAGTAGTTA TTTTACATAA CGTATTAGGG TATACGATTG GATATTGGTT AGCTCGTCTT	660
	TTAAATTAG ATCGACAAGA TCAAAAAGCA GTCAGTATTG AAGTTGGAAT GCAGAACTCT	720
	GGTTTAGCTG TGTCATTAGC aGCATTGCAT TTTAATCCAA TTGCAGCAGT ACCAGGCGCA	780
15	GTGTTTAGTT TCATTACATAA TATAACAGGG CCTATTTTAG CAAAGTATTG GTCAAAAAAG	840
	TTATAATTGC ACTAATAGAA TGAAGTGGTC ATCGGACTAT GTTAAGCTTT GATAAAGAGA	900
	AAAAATAGAG GAGTAAATAT ATGTATAGAG CAGTTATATT TGATTTCGAT GGAACAATAA	960
20	TAGATACGGA ACAACATTTA TTTAATGTTA TTAATAAACA TTTAGAGATG CATAATGCCG	1020
	ATCCTATAAG CATTGATTTT TATCGTTCTT CTATTGGAGG AGCAGCTACA GATTGTCATG	1080
	ACCATTTAAT TAAAGCGATT GGTTCGGAAT ATAAAGATAA ACTTTATGAA GAACATCATC	1140
25	TTACTAGTAC AACATTACCG ATGATTGATA CGATTanATC ATTGATGGCA TTTTTAAAGC	1200
	AACGTCACAT TCCTATGGCA ATTGCCACAA GTAGTGTGAA AGCGGAAATA ATGCCCACCT	1260
30	TTAAAGCATT AGGTCTAGAC GATTATATAG AGGTAGTTGT TGGTAGAGAm GATGTTGAAC	1320
	AAGTTAAACC TGACCCTGAA TTATATTTAT CTGCAGTACA ACAATTAAAT TATATGCCGA	1380
	CACAAATGTTT GGCTATTGAA GATTCTGTAA ATGGTGCAAC AGCCGCGATT GCAGCTGGAT	1440
35	TAGATGTTAT TGTTAATACG AATAAAATGA CAAGCGCACA GGACTTTTCT AATGTAGATT	1500
	ATGTAGCAAA AGATATTGAT TACGATCAAA TTGTAGCGCG TTTCTTTACG AAATAGGAGG	1560
	CGTATCATGA TGGGTTACAT TATATTGTTT TTTCTAGCTG GTCCAGTAAT TTTAGGCGTT	1620
40	GGAAATTTGG TGATTGGTCC TATATTTAAC AAACAGACAC CATTTGCGGT GCAAGTAAGA	1680
	TCTTTTGTGG kTGGkTCmAT GrTTTACTTA ATACTCGCAA CAATTGGCTA TTTTTTACTA	1740
	TTACAAGGTA AACTTTAACG AGAAAACCAC CTTACCTCAT TAAATGGACG ACCATATGTA	1800
45	TGTGAAATGG TAGAACGTTT ATGTTTATGT ATGAGATAGG GTGGTTTAAA TAGTTACATA	1860
	TATTTTAATA ATAACGTCAC GATGATAAGT ACAATTAAGA TAATATCTAT GCCTACCATA	1920
50	ATTGfAGCTC TTGTTGCAIT ACTTCCTTGT TCTTTTGCTG ATTTCATAGC ACGGTAGTTT	1980
	GGCACAAAGC TAATAATTAG TAAGATTAAT ACAATTACAC CAATTAATGC TGTGTCATG	2040
55	ATGAACGACC TCCTTTATTT TTTTCAATCA ATTCCCAAAT AAACGTAGCA ATCACACCGA	2100

CAATAATTAA TGCAATCGGT AAAGTCGTAC CGAGTTTAAT CTTGCGCTCT GGAGAATTAA 2220
 TAATAGTAAA TACTGTAAAG CAAATGAGTA TGAAAGCAAG TGTGCAATA ATAGTTCTTC 2280
 5 CAACTAAATA TAGGATGTCA GGTTTTTCCA TACCGATATA ATTTATGATG AAAAATGCTA 2340
 CAGCAAAGAG TACCGATATT TTTGTAGCAC GTAGCAGTAT TTGTTTTAAC ATTGATATAC 2400
 TCCTTTTTAA TATTATTAAA ATTATATCAT AATTACCAAG AATAGCTGAA GTTGATGTG 2460
 10 ACTCAACGGT ACTTGAGCAA CTTTTTTAAT TTTTGTAGAAA AATCACAAAA TAATTGTTTG 2520
 CAAAGTTGCA AAAGCCTGCT ATAGTAGTTC TGTAACGAT TGCATGGTAT GCAAATATTA 2580
 ATGTACCAAA ATCGATAATT TATAGTATAA TTACGGCAAT AAGTTTTTTT ATGGATTTAT 2640
 15 TTAGTATCAA TCAGAGATGG GGTAAGAAGT TATGGAGAAC AATGAACACT AAAGGGGATT 2700
 GAGTGCCCGT CAAATTCAAA TGATTGCACT TGGTGGTACG ATTGGCGTGG GGCTTTTCAT 2760
 GGGTGCGACA AGTACAATTA AATGGACAGG CCCATCAGTT ATCCTTGCAT ATTTAATTGC 2820
 20 GGGTATCTTT TTATTTTTAA TCATGAGAGC AATGGGGGAA ATGATTTATT TAAACCCTAC 2880
 AACAGGATCA TTTGCAACAT TTGCAAGTGA TTATATACAT CCTGCAGCAG GTTATATGAC 2940
 25 AGCATGGAGT AATATATTCC AATGGATTGT AGTTGGTATG AGTGAGGTCA TCGCAGTAGG 3000
 AGAATATATG AAGTTTTGGT TCCCGGAATT GCCAACTTGG ATTCCTGGTG TTATTGCTAT 3060
 TTTATTATTA ATGGCAGCGA ATTTATTCTC GGTAAGCG TTTGGAGAAT TTGAATTTTG 3120
 30 GTTTGCTTTA ATTAAAGTTG TAACAATTAT TTTAATGATT ATTGCTGGTT TTGGTCTTAT 3180
 nntCTnTGGT TTTGGAAATG GTGGCCATGC GGTAGGTATT TCTAATCTAT GGACAAATGG 3240
 C 3241

35 (2) INFORMATION FOR SEQ ID NO: 646:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1311 base pairs
 (B) TYPE: nucleic acid
 40 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 646:

AGGCGTCAAC TCAGATGGTT TAATAATTGC CGTATTACCT GCTGCAATAG CACCGATTAA 60
 AGGTtCGAaC ACTAGTtGAA AAGGATAGTT AAATGGTGCA ATGATCAAAA CTGTCCATA 120
 50 AGGTTCTTTT TTGATATAGC TTTTGTGTGG AAATAAATAT AAAGGTGTGT CTACATTkTT 180
 TGTTTTAGTC CAGTTTTTAA GTTCcTTACG GGCAATTTTG ATACTTyTCA AAGTTATGCC 240

55

AATATCGCTC TCGTATGATT TAATAGCTTT GCTTAACTTC TTTAATTGCT CTTTTCTAAA 360
 ACTAATATCT TTAGTTTGTG GTGTATTGAA AAAAGCTTTA CTGTCATAAA ATTTTGTCTC 420
 5 AATGATATTC ATAATGAAAA GAACCTCCTT ATATGATTAT TTTGGAAAAA GCGATTAATT 480
 GATTTGAATG TTGTGGCCGT TAATTTTAAA TGGTCTTTTCG AATTATATAT GTTGAAAGTT 540
 GAAAATAGAG CGATGAATCG TGTACATAAT AATATTTATA ACTTTAATCA TAACGAAAAA 600
 10 GGTAGGAAGA AAACAAAAAT TTATACTCAA CATCGCAAAT ATTTTAAGAA AATGTAAAGA 660
 CAAAAGGGGA ATTGTATAGA AATCACTAAT CTGTGGGTTA GGGTAGCTAA AGGAATAAAA 720
 ACTACTATTG AAAAAGGGT GTAAATTAGT CAAACGTAAA TAAAAACAG TTCATTGAAA 780
 15 GTGAAATAAA TTCTACTTTA ATGAACTGTT AGTTAAATAC AACATGTCTA TAATTAGACA 840
 GTAATATAGT ATTATTTTGT TAATGCTTCA GTGATTGAG GTACGATTG TTTTTTTCGA 900
 20 GAAACGACAC CAGATAAGAA GGCCATGTCA TCTTCTAATT GAACATTGAA TGTCGCCAA 960
 CTTTATCTTT TTCAGCACCT ACAACTAAAA TTTTAGAATC ACTATTAATG ATGTCAGTAA 1020
 CAACAAGTAC AAATAAGTCA TATTTTCTT GTGCACTTAC AGCTAACATT TCTTTTCTA 1080
 25 TATCTTCTTT ACGATTTAAC ACTTCGTCOA GGTCAACAGC ATTAACCTGT GCAATACGAG 1140
 TCACATAGTC ACCCATAGTA AATGATTAG CATCCATGTT TAATAAGAAT TCAACTGATT 1200
 TATCAGTTGT TGAAGCACCT GCTTTTAAACA TATCTAAGCC GTACTTTTGA ATATCAACTT 1260
 30 TAGCAATATC TTTTAAATTCT TCAGCTGCTT TAACATCTTG TTGTGTACAT G 1311

(2) INFORMATION FOR SEQ ID NO: 647:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 1498 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 647:

CATATACCTT TTCATTTCTT TACGAGATAC TTTACCAGAG GATTAGACT TCATACGCTG 60
 45 ATCCATATGT GCTTGCGTTT CAGaATGTCC ACAAAACAAA CGATATACCG CTTCTTTCCC 120
 TTTACCAAAC AACGTTAATT TCTTTTACA GTTTGGACAT CTTGCATTG TTTTGCCTG 180
 TACATTCCTT TTCGTCTTAC AAGATGGATC TTGGCACACA AGCATCTGAC CATTTTGTAGT 240
 50 TTTAACTTTA ATCATGAATT TACCACACGT TGGGCATTCT GTGTTGTGA AATTATCGTG 300
 TTTATATTTA CGATCACTAT TTTTAATCCC ATTTACAACA TCTTTCGTAA AATCTTTCAT 360

CCATTGTGcA GTTAAAAGTG GCGACGTTAA TTCTTCTGGT GCTAATTCTA ATATTTGTTT 480
 ACCTTTTGAC GTTACTTTAA TTTTACCGTC TCTTGATTCA ATGGCATTCA TATTAAATAA 540
 5 TTTATCGATA ATGTCGGCCC TTGTTGCAAC TGTGCCGATA CCACCTGTTT GTTTTAAAGT 600
 TTGCGCATAT TTTTATCCT TCAATTGAAT AAAGTTCTGA GGGTTCTCCA TCGCTTTTAA 660
 TAACGAACCT TCATTAAAAT ATTCTGGAGG TGTGTTTCA TGTCTCTAA TATTTGTTTT 720
 10 TGAAATCTTC ACTTCATCGC CTTCTGAAAA AGGCTGTTGC ATCTCTGTAA TAGATTCACC 780
 TTGTCTAATA GATTTAAAAC CTAAAACAGT TGTTACATTC TCTTCAAAA CAAATGTGTG 840
 15 CCCTGCAACC TCTAAAGTTA CAGTTATCGC GTCATACTCG TGCGGAGGCA TTAAAGCTTC 900
 TAAAAAACGC TCGACAATCA TATCGTATAA CTTTAATTCT CTATTACTTA AGTCTGACAT 960
 GACAGGTCTC ACTTCTGTAG GAATAATTGC ATGGTGATCA GATACTTTTT GATTATTAAA 1020
 20 TATCGACATT TTTGATGAAA ATGTTTTAGA CATTAAATGGG CGTGCTTGGT CTTTATATGT 1080
 TGTGCCATC GTCACCTGAA TACGTTCTTT CATAGTATCT ACCATATCAG TTGTTAAATA 1140
 GTTTGAATCT GTTCTTGAT AGGTTACGAC TTTATGTCTC TCATATAAGC TTTGAAGTGT 1200
 25 ATTCAATGTT TCTTTAGGTC CAATTTTATA ACGTCTATAC ATATCTTGTT GTAAATCTGT 1260
 TAAATTGTAC AGTGATTGCG GATACGACTT CTTATGTTTA GTAGCAACAG ATTTAATCTT 1320
 ACCATCGACA TTTTCAAAT TATTAACCAT CTGTTCTAAA GTTCTTTAw TGGtATATCG 1380
 30 CTGATTTGaw TCTAGCTGAA AATCAAACCC TTTTACCGTT AATGATAATG TAAAGTATTG 1440
 TTGTGGnTTG AACTGATTAA TCTCTTGTTG GTCGTGTAAT TTAATAAATT GAAACGGn 1498

35 (2) INFORMATION FOR SEQ ID NO: 648:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1044 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 648:

TAAAAAATT CGTTTTATTA ATACAGTTTG TAAAAGTATT CTCTTAGTAC TAATAGCTAA 60
 TGTAATGATA GAAAATACTA GTGTTATTAA AGATTTGAAT AAAATAAAAG AAAGTGAAG 120
 50 ATATTGGAAT GTATTAGATG ATTATTACAC GATTGAATTT GCACCTTATC ACGAAACAAA 180
 ACAAAGTTTG ATTGATAATA TGGTGCGATC AGAACAATTA GTAAAGGCTA GTGAAGCAGA 240
 AAATAATGCG ATTTTATTCA AACCAAAGGG TGAATCCGTT GACAATGACA ACTTTTCGCC 300

55

TCAACCTGAT ATTCCGATAA AAAATCAAAA AAATAATGTC GAAGTAATTA TTCCACAAAA 420
 GTTTCATGCA ATGCGTAATG AAATCAATCA AGCATATCAT TCATGGTTTG AATTTGTACA 480
 5 AAATAAAAAT AATAAGAGA ATAAGTTATC TATACAGTTT ATCAACAAAA ATGATTGTCTG 540
 AATTTTTTCA TTTGATGCAC GAGATAGTCG CCATTTGTCA TTTATAGAGG CGCCAATCAT 600
 TGTGAATGTT CAGGCATCAG ATTTATCGaA TGATTTtAT TATGCCATGa TCaGTCaAGG 660
 10 CGGGTATTTa TTCaAAAATT ATGaCGCGCT AGTAAAAAAT ATTGGAAAAG TATCCATCTT 720
 GATGGGGAAA TCCAGTGGAA TAACCAATTA TAAAGATAGC GTGATGGAAA TGTATCATGA 780
 AAACAATTTG AAATTAACAG TACTCAACTT TTCACAAATC ATTATCGCAA TCATTTTAAT 840
 15 AATTATTATT TTATTTGATG TGAAATATTA TTTTGAACAG CATCGAAAAT TACTCGTAAT 900
 CAAAAAGCTA TATGGTTATT CAACATTAAG AGCCAATTAC CAATACTTAT TAATAAATAA 960
 20 TATAGTTGTT ATTTTTATTG GAATATTGAC GAATGTAATT TTACATTCTC ACTATATAAT 1020
 GATGTTATTT GCAACGATTC TTGT 1044

(2) INFORMATION FOR SEQ ID NO: 649:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 649:

35 GAACATATTG GGTATGCAA GGnGGTCACT CTTCACTT ATAAACAACA TTTTAATAAT 60
 GTAAAGTTTA ACCAGCTAAC ACTTTTGTTA GCTGGTTTTT ATTTTCCTTC AATTTTTTAAA 120
 TGGTTAAGTC CCCTTCTATA TCTTATAAGA CAATCATTAT AATCAATTCA ATTAATACAT 180
 40 TAACAACCAC AACTAATAAA TATAGTAACT TCAAAATCCA TATTTATGTC TAAAGATAAT 240
 CTCAATGTTG TTCACGTCAA TAAATTATC CCTAGGTTTT TAAAAATTGT ACATGTTTAA 300
 ACAATCAAAA GTGTACATTA TTAAATTATC ATTTCCAGTT AGATTTAGAA AACATTCACA 360
 45 CCACGCATGG ACCAACGTAT TCGTCTTcAT TcATTTTAG 399

(2) INFORMATION FOR SEQ ID NO: 650:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 747 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 650:

	GTACTAGGTT CTAnGATTCC CTACTGTAGT CATAGATATT TTCCAAACAG AGCCTGACAA	60
5	CTTTTCATTT TTAATAGTGA TCACTATATT AATCGCTATT GCAATTTATG CATGTCGATT	120
	TGTTTGGGTT TATTTCTGGT ACAAAGATTT TTATTTCCCG AAAAATATAC AATCTTATCT	180
	AGACGAGGAA CATGATTCAC ATGAAACACC ACCTTCTCGA GTGCGTTACG CATTTATTAT	240
10	GACCATGTGT GGTATTCACG GTACAATTC ACTTTCAATG GCACTTACAT TACCATTTAT	300
	CATTACAAAA GGACAAGCAT TCGAATACCG TAATGATTTA TTGTTTATTG CATCTTTCAT	360
	GGTATTAATT AGTTTAATCT TAGCGCAAAT TGTTTTACCT TTAATTACAC CATCTGCCGA	420
15	AGATACTACT TTAAAGGTA TGACTTATCA ATCTGCCAAA ATTTTCATTG TTCAAAAAGT	480
	GATCCAGCAT TTAAAAACG AAAGTAAAAA AGACAAAAAC GATACAAATT ATCGCCCACT	540
20	ATTAAACCAA TACTATGGAG AATTGTTATT TTTATTAAAT TCAGAACCTG ATAATCAAAA	600
	TACrAAAGAA CTCAAACGTT TAGAAGATAT TGCAAAAGTA ATCGAAACAT CTACACTTGA	660
	GCGTTTAATT GATAAAGGTA AGGCAACATA TCAGGATATT AATAATTACC GCAATATTGT	720
25	CGAATTAACA GAGACACACC GTACTGC	747

(2) INFORMATION FOR SEQ ID NO: 651:

(i) SEQUENCE CHARACTERISTICS:

30	(A) LENGTH: 1373 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 651:

	ATACAATACT CTTTTATTAT TCAATAAGCC ACTTCCTATA GCAAATGTTT AACTTTAAAT	60
40	ATTTTTCGAT GCTAACAAAA AATCACACTA TCATCTTTTA AAATGAAAGT GTGATTACAA	120
	GCAAATCTGT AAAATTTATA AAGCAGAAAC AATTCAACTT TATCATTATG ACATTTCAAT	180
	TAAACCTTCT ACATTATAGT TCCAAGCATC TTACACATGA ATGCAAGTAT TTAACGATTT	240
45	AATTGTGACA TAGCCTGTTG ATATTGTGTT TCATTGATAT AATTTTGTTG CTCATTTTTT	300
	TCTAAGTTCTG TGCTTACACG TTGCGTAAAA TTCTCTGACA TATTATTGAT ATTATATACG	360
50	CTAGGTGCAT TGACTTTACT AGCTAAAATA GCGCTTTGTA AAACTGTTAT GTGAGACATT	420
	GTTGTACTAT TTTTATTCAC GGTTGTTCCA AAGTAATGGT TTGCTGCGCC CTCAAGCGTA	480
55	TATTGATTAT CCCCAAAGTA AATATTATTT AAATAAAGC TTAAATTTTC GTTCTTATTA	540

TCATTATCAT AAAAATAATT TTTGACAACT TGTGTGTAA TGGTACTACC ACCTTGCACA 660
 TCTCTGTGCG TAATCGTTGA AAATAAAGCT CTAGTTGTAC CTTTCAAATC GAATCCATGA 720
 5 TGATTGTAGA ATCGTTCATC TTCCATTGAA ATAAAGGCAC CTTTAACATA CTCTGGCATG 780
 TTATCAGCTG ACACAAAACT ACTTTTATTT TCAATTTTTC TTAGTTCATC CACATTATCG 840
 CGTGTAGATA AAAAATACAT GATACCAATA AACAATGCCA TAATGATTAG AATGGTTAAT 900
 10 AATATTTTTA ATAGTATTCG TTTACTTTTT TTCTTTTTTC GCGGTTTGCC AACTGGTTGA 960
 TAATACGTAT TATAGTGAGG TTCGTGTTTC ATATGCTCAA AATGTTCAAT TGAGTTTGAG 1020
 TACCTATCGC TTCTTTTCAT GCGTTTGCTC CTTCTTTTAA AACTCACTTA GtATATACCT 1080
 15 TGaGTTTACC AGTACTATCA CAAATAGGCT ACACTTTTTG GGAAAATCAG TCCAAGGGCT 1140
 TACAATCGTA TACGCCATCA TACTTACTTT TTTGTTTTTT GAAAAAATTA TAGATAAATC 1200
 20 ATTGCAATTT TAAATATTAA TCATGTCAAA TATTGTTATA TTTTATAAAA ATAAAAGACC 1260
 ATCCCTATTA AATGCCAATA GAGACGACCT TTTATTTGTT ATTCATTTAT TAAAACTAAA 1320
 ATCCATATTT CATTTCAAAC GAAAATATAT AAATTTTAAC AATCGrTAAC CAC 1373

(2) INFORMATION FOR SEQ ID NO: 652:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 859 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 652:

GGCAGATAAT TTAGTCATTG TTGAATCGCC TGCAAAAGCA AnaACCATTG AAAAGTATTT 60
 AGGTAAGAAA TATAAAGTTA TAGCTTCAAT GGGACACGTC AGAGACTTAC CAAGAAGTCA 120
 40 AATGGGTGTC GACACTGAAG ATAATTACGA ACCAAAATAT ATAACAATAC GCGGAAAAGG 180
 TCCTGTTGTA AAAGAATTGA AAAACATGC AAAAAAGCG AAAAACGTCT TTCTCGCAAG 240
 TGACCCCGAC CGTGAAGGTG AAGCAATTGC TTGGCATTTA TCAAAAATTT TAGAGCTTGA 300
 45 AGATTCTAAA GAAAATCGCG TTGTTTTCAA CGAAATAACT AAAGACGCTG TTAAAGAAAG 360
 TTTTAAAAAT CCTAGAGAAA TTGAAATGAA CTTAGTCGAT GCACAACAAG CGCGTCGAAT 420
 50 ATTAGATAGA TTAGTTGGCT ATAACATCTC GCCAGTTCTT TGGAAAAAG TAAAAAAGG 480
 GTTGTACGCG GGTCGAGTTC AATCTGTTaG CmTTCGTTTA GTCATTGACC GTGAAAATGA 540
 nATTCGAAAC TTAAACCAG AnGAATATTG GACTATTGAA GGAGAATTTA GATACAAAAA 600

5 AAAAGATGTT GAGAAAATTA CAGCTGCATT AGATGGAGAT CAATTCGAAA TTACAAACGT 720
 GACTAAAAAA GAAAAAACGC GTAATCCAGC AAACCCATTT ACAACTTCTA CATTACAACA 780
 5 AGAGGCGGcA CGTAAATTAA AcTTTmaAGC AAGAAAAACA ATGATGGTCG CACAACAATT 840
 ATATGAAGGT ATAGATTTG 859

(2) INFORMATION FOR SEQ ID NO: 653:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 653:

20 TTCAACTTCG TTTGGAAATC ATGTTCTCTCA ATTGTTGGAC TTAAACGAAT TTCTTTAACA 60
 TTGATAATTT TTTGTTTCTT TTTCATTTCT TTTCTTTTT TCTGTTGTTT GAATTTGAAT 120
 TTACCGTAAT CCATAATTCT TGCAACTGGT GGTTCGCAT TCGGTGCAAC GACCACTAAG 180
 25 TCTAAATCTA CACGTTTCAGC CATTTCTAAA GCTTCACGCT TTGATTTAAC ACCAATTTGT 240
 TCACCATCTT GACCGATTAA ACGTAATTCT TTTGCACGAA TTTTGTCATT GATTTGAGTT 300
 TGATCTTTTG CTATGGTTGA CACCTCCAAA ATTTTACGA AATTGCAACC AAGCAAAAAG 360
 30 GAAGAGCAGG TATAAAATAC CCGCTCTTCC TTATACACAG TTATGTGTAA TGTGATTAAC 420
 CTGCCAACTG CTTTATGCGT CGCTACAGGT GAGAAGCGGG TGCTTCTACT TGGTTCGTTT 480
 35 CGTATTCAAC GTTATTAATC ATATCAACAA TTCACATTTA AGTCAACACT ATAACGTAA 540
 TTATTTTTAT TTTAACCTTT TATTTTCATCC ATTGACACGT CTTGACGTAA ATCTACTTGT 600
 TCTAATGGAA TTTTTTTCGT TTTATATCGA AGCTTATGAT AAATAAAGAA TGCTAAAAAT 660
 40 ACTGGAATTC CCATATACGT AATTAAGAAG CGACTAAAAT TAAAATCTCC TGTnTTAATA 720
 AAGTCAACAT CTTGCCCAAn AAnTACT 747

(2) INFORMATION FOR SEQ ID NO: 654:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 501 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 654:

TTGTGCTTCT TTTTAGCTT CTTGAACCTC TTGTGCCTCT TGTGATGTAT CACTyAAATT 120
 ATTTGCACTT GCTTCTTCTT TTATCGCTGC TTGTGTGCT TTCAATGCCA CTGCTTTTGT 180
 5 TTCTTyATTT GATACAGCCA CACTTTTATC CGCTTCTGCT TGTGCTTCTC TTTTAGCTTC 240
 TTGAATCTCT TGTGCTTCTT GTGATGTATC ACTTAAATTA TTTGCACTTG CTTCTTCTTT 300
 TATCGCTGCT TGTGTGTCCT TTAATGCCGC TTGCTCATTT TTAGATTTGT TTAAAAATCC 360
 10 TTCAACACGT TCTTTTGTAT AGGCAACCGT TTCTTCAAGT TCGTTTTTTC TTTCTTCAAA 420
 CTTTTGCGAC AGTTCTTGTh CTTTGACTTT nAAATCATCT GCTTTTGTAT AAACTTTATT 480
 15 TTAAATACC AACCTAAAGC C 501

(2) INFORMATION FOR SEQ ID NO: 655:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 655:

CACCTTGTCa TAATTAATTT TTTGATTTTT CATTTTACTG ATAATAGGTT CAGCATTAAAT 60
 CATGATTTAA CCTCCACAT TTAATCATTa ACTTCTATTA TATATGATTC ATATTAAATG 120
 30 TCAGTCAAAA AAGTTAGAAA TTCATTTTAA TGCATTTATA TTTCGAAAAT CCCTTATGTA 180
 TCTAAAAGCA TTTTTTAAAC TTGAATTTTA AAACACTAAA CAACACATAC GTCTCTGTGT 240
 35 CATTTTCATT TTTTGTATGT CATATATATG TTTACTTCAT TTAAATCAAT TTCATCTTAT 300
 AATTTATCGT GTATTTTACA AAAGATTGAC TTCAATTCAT CGTAAAAGTT ATACTTTTGC 360
 CATTTTTTAA TGTAACATGG TGTTAGTAAT AAAAATAATA CATTGAGGTG TTTTACATGA 420
 40 CAGCATTATT CCCTTATATC GCTTTTGAAA ATTCAAAGA AGCCCTTGCA TATTACGAAG 480
 AAGTATTGG TGCAACTGAC GTTAAACGTT TAGAAGTTGG CGAAGAACA GCGTCACATT 540
 TTGGTATGAC TAAGGAAGAA GCGCAAGAAG CAACTATGCA TGCTGAATTT GAAGTGCTTG 600
 45 GCGTAAAAGT GTTATGTTCT GATTCTTTTG GTCGCGCTGA CAAAATTaAT AATGGCATAT 660
 CATTATTAAT TGATTATGAT GTTAACAATA AGGAAGATGC TGATAAAGTT GAAGCATTCT 720
 50 ATGAGCAAAT TAAAGATCAT TCTTCAATTG AAATAGAATT ACCGTTTGCT GACCAATTCT 780
 GGGGTGGCAA AATGGGCGTC TTTACCGATA AATACGGTGT TCGTTGGATG 830

(2) INFORMATION FOR SEQ ID NO: 656:

- (A) LENGTH: 539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 656:

10 GTATCCATGG GCCCGTTnCG CACAACATTT GGnACAATTA GGTACACGnG TTGTCaTCGG 60
 TCGTTTCGGT ATAATTTTAT CGAATGaAGG CGGTGCGTTA CAAACAATGA AACTACCATA 120
 CGAATATTAC ATTGGTGGTA AATTAGGTTT TGGTCAACAA TGGTATTCAT GGATTTCATAT 180
 15 CAATGATTTA ATTCAAGCTA TTTTATTTTT AATAAATAAC GAGTCAGCTA GTGGTCCGTT 240
 TAAITTTAACT GCACCTATAC CTGAACGTCA AAATTTATTT GGCTACACTT TAGCAAGAGC 300
 TATGCATAAG CCTCATGAAA CTGGGGCACC AAGTCTTGCA ATGCGTCTCA TACTTGGTCA 360
 20 AATGTCAACA GTAGTATTGG ATACTCAAAA AGTATTACCT aATAAAATTC AAGCATTGGG 420
 aTTCCAATTT AAATATAGTA ATTTAAAAAT GGnACTTGAA GATTTAATTA AAGAATAATC 480
 AATACCATTA ATGAGCATTa GAAACAACAT ATGTACTAAA TGTAATGTCT AGAGCGACT 539

25

(2) INFORMATION FOR SEQ ID NO: 657:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1106 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 657:

ATCAAGTTGA ATATCATCCA TATTTAACCC AACATAAATT GAATTATAT TTGGCAGCAC 60
 AACGTATCGT gATGGaATCT TGGTCACCAT TGATGAATGC ACAAATTTTA AATGATGAGA 120
 40 CAATTAAAGA CATTGCTCAA GAATTAGGAA AGTCACCTGC CCAAGTTGTT TTAAGATGGA 180
 ATGTGCAGCA TGCTGTGGTT ACAATCCCTA AATCGGTGAC ACCAAACAGA ATCTCTGAAA 240
 ATTTCCAAAT ATTTGATTTT GAATTATCAG ATGAACAAAT GACGCGAATT GATGGTTTAA 300
 45 ATCAAGATAA GAGAATTGGA CCTGATCCAA AAAAAATTGA AGGCTAGATT AAAATCGCTC 360
 AACTGATGAA AAGGTTAGAT GAATTGTCAG GGCTTGGGAC ATTAAGTTCT TAGGCAATGT 420
 50 AAAAAAGCTG ATTTCTATTA ATTATTTGAT AGAAATCAGC TTTTTTGATA TGTATTTTAT 480
 AATGTACAGC TCGTTGAGCT GCTATTTTCC TTATATTAAG TGCCATTAAT ACAAACCTA 540
 GCTCTCGTTT AACTTTATTT AtTCCTCGAA CTGACGTTTC AGTTAAACCC AAAATAGCCT 600

55

CTGGTTCAGA AAGCTTTTGA TTAATTTGGA CTTTAAAGTA TTCCCAATTA TAATTCTTCA 720
 TGATTtTCTT ATTGGATTTT GAATTTGGTT TCATGCATTG TTGCCTCAAA GAACATGCTG 780
 5 AACAGTCATC GCATTCATAT AGTTTGAAGT CTCGTTTAAA ACCATATCTA TCATTACGGT 840
 ATGCATATCT TTTAAAACCT ATTCTTTTGT TATTAGGACA TATAAATTCA TCATTAAGTT 900
 10 CGTCATATTT CCAATTTTGA GTGTCGAAAA TGTCACTTTT AAACTTTCTA GTTTTATCtT 960
 TAATAAACAT GCCATACGTA ATAAGTGGCG TTTTATTAAA ATCATCTATA ATGGCCATAT 1020
 AGTTTTGGCT CACTACCCAT AACCTGGCAT CAGCTACCAA ATGAACCGAA GGGATTTTTG 1080
 15 GAATCCATTG GTTGAAAAAA TGGGAA 1106

(2) INFORMATION FOR SEQ ID NO: 658:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 658:

TTTTAACTTT ACTCTTTGAT TTAAAGAGTG ATAAATGTTT ACAGTTTAAT TAAAACTGCA 60
 TAAGAACTTC TAGCTTTTCT CTTCGTTCA AAGAGAAGCA GCTGTTGCA GTTTAATCAA 120
 30 AACCACATAA AGCTTTTAACT TTTACTCTTT GATTTAAAGA GTGATAAATG TTTACAGTTT 180
 AATTAAACT GCATAAGAAC TTCTAGCTTT TCTCTTTCGT TCAAAGAGAA GCAGCTGTTC 240
 35 GCAGTTTAAT CAAAACCACA TAAAGCTTTT AACTTTACTC TTTGATTTAA AGAGTGACAA 300
 ATGTTTACAG TTTAATTAAA ACTGCATAAG AACTTCTAGC TTTTCTCTTT CGTTCAAAGA 360
 GAAGTTCTAA TACCACCATA TCGTGCGATC GGGAACGGTA 400

(2) INFORMATION FOR SEQ ID NO: 659:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1899 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 659:

50 ATAATTACAA TGGCTACTGC AATTATTGTT GGAATTATAT ATTTCAAGTT AGTCATCATT 60
 TATGCTCCCT TGATTTCAAA TTCATATCAT TAGTTTACCA TATTGAAGAT GATATAATAA 120

	AACTTTAACT GAGCTTCATG GGGCACCAGG TTTTGAAGAA GAAGTAAAAA ATTATATGAC	240
	TCAGCAAATG GCGCCGTACG TAGATGAATT TATTGAAAAT CGTATGGGTG GATTTTTTGG	300
5	TGTGAAAAAA TCTAAAAATC CAAATGCAAA ACGTGTAATG ATTGCAGCAC ATATGGaTGA	360
	AATCGGATTT ATGATTACAA ATATCACTAA AAATGGAATG ATTCAATTCA CAAATTTAGG	420
	TGGTGTTCGA AATGATATTT GGCAAGGACA ACGCTTAGTA ATTA AAAATA GAAATGGCGA	480
10	TAAAAATTATC GGTGTTGTTT CTAATATACC TAAACATTTT CGTACTGGTA GTGAAGGTGC	540
	ACCGGAAATT AAAGATTTAA CATTAGATAT AGGTGCTCAA AATGAAGATG AGGTGCGTGA	600
15	GCGCGGAATA GATATAGGAG ATACAATTGT ACCTCACACG CCATTACACAC AGTTATCTGA	660
	ACATCGATAT AGTGCTAAAG CATGGGATAA TCGTTATGGT TGTGTCTTGG CAATTGAAAT	720
	ACTAGAATTA TAAAAGATA TAGAATTAGA TGTAGACTTG TATGTTGGCG CAAATGTTCA	780
20	AGAAGAGGTT GGATTACGAG GTGCGAAAGC ATCTGCAGAG ATGATAGACC CAGACGTTGC	840
	ATTTGTAGTT GATTGTTTAC CTGCCAATGA CGTTAAAGGA AACCAACCAT TATCTGGTGA	900
	ACTTGGTAAA GGGACGTTAA TTCGCATAAA AGACGGTACA ATGATTTTAA AGCCTGTATT	960
25	TAGAGACTAT TTATTAAAGT TAGTAGAAGC ACATGACATT GAACATCAAT ACTATATGTC	1020
	ACCAGGTGGA ACAGATGGTG GAGAAATTCA TAAAGCTAAT ATTGGTATTC CGACTGCAGT	1080
	TATTGGTGTA TGTGCACGAT ATATTCATAG TACAGACTCA GTATTTGATA TAAGAGACTA	1140
30	TTTTGCAGCT AGATCTTTAC TTTCAGAAGC CATTGTGAAT TTAGATAATA ATCAAATAGA	1200
	AACATTACAA TATAAATAAT CGGGTAATAA CAACTATTAT CTCTAAATAG TTATATATAA	1260
35	TCATTAATTA AGGAGACATA AAAATGAAAC AACTTGAATC AGAACAACAA TTTGAATCTT	1320
	TAAAACAAGG TGCTACAGTA TTTGAATTCA CTGCAGGCTG GTGTCCAGAT TGTAGAGTGA	1380
	TAGAACCAGA TTTACCGGAA TTAGAAGCGA GATATCCTAT GTTTGACTTC GTATCAGTAG	1440
40	ACCGTGATAA ATTTATGGAT ATTTGTATTG AAAATGGTAT TATGGGTATT CCAAGTTTTT	1500
	TAGTATATAA AAATGGAGAA CTGCTTGGAA GTTATATTGG AAAAGAACGA AAATCAATTG	1560
	AACAGATAGA TGCATTTTTA GCTCAATACG TGTAATTTAG ACTAGAGAAA AACGGGGTAA	1620
45	TACTCGTTTT TCTCTGTTAC TATGTGTTGA TTTATTGTAA ACTATTATAA GGTGCGAAAT	1680
	TAGGAGTGTT ACATATGAAT ACCTTTCAAA TGAGAGATAA ATTAAAGGAA CGTTTAAGCC	1740
50	ATTTAGACGT TGATTTTAAA TTTAATCGTG AAGAAGAAAC TTTGCGTATT TATCGAACAG	1800
	ATAATAACAA AGGTATCACG ATTAACTTA ACGCTATAGT CGCAAAATAT GAAGATAAAA	1860
	AAGAAAAAAT TGTAGATGAA ATTGTTTATT ACGTTGATG	1899

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3774 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 660:

10	GTATAATATC CTTTCAATCT GTTTTCATAT TTTATATATT TTTTAAATAT AAGTGCTAAA	60
	TGTTTTAACT AAAGCATAGA TTGACAAGAT GTTATACAGA ATTTCAAATT CTATCCAATA	120
15	TTGTTCGAAG TGTAAGTATCA CTGGATTGGT ATTAAACAAT GTAAAGGAGA GATTGCAAAT	180
	GCCGTATAAT TACAAGAAAC AAAATGGAGA GTTAATGTCT GTAATGAGCC AAGGTGAAAA	240
	GTTTATTCAT CAATCACCCG TTAATGATGA ACTTAGTGCA TTGATTAACT TATTAATTTT	300
20	TAAAATTAAC GGTGTGCATT ATTGTGTTGA TATCCATAAA AAAGAATTAA AGGAATTGGG	360
	TGTAACACAA ATGAAAATTG ATGAAGTCTT GAGTTTTAGA CATTTAGATT TATTTACTGA	420
	TCAAGAAAAA GTGACGCTTG AATTTGCAGA AATGTTAAAT TCAATCAAAG ACTTTAAGAA	480
25	GTTTGAAATT ATTGACCGGC TAAAATCATT TTATGATGAA GAACAAATTA TTGATCTTGT	540
	CTTTGTTGTA AACCAAATTA ACGGTTGGAA CAGATTAAAT ATTATTAGTG ATAGACTATA	600
	ATTGTTTCATA TAAATGCAGA GTTTCATCTC GAACGCTATA TCATAACAAA TCATGCCACT	660
30	ATACAGGTCA AATCTTGTAT AGTGGCATT TAAATTTATCC CTTTGAATAC TGTTATTTAA	720
	CGAATATCGG TCCACCTGGT CCAACAGGGA TACCTAATAG GAACCAAATG ATGACAAATA	780
35	CTGTCCATAC AATACTTAGC GCGATTGAAT ACGGCATTAA ACTAGAAAGT AAGGCTCCGA	840
	GTTTCATGCG TTTATCGTAT TTTTGTGCAT AAGTTAATAA TAAAGGTAAG TACGGCATCA	900
	TCGGTGTAAT TGGATTGGTA ATTGAATCGC CTACACGGTA AATGACTTGT GTGAATGCGG	960
40	GATGAAAGCC GATAAGGATT AACATTGGTA CGAATATCGG TCCTAAAATA CCCCATTTAG	1020
	CCGATGCGCT TCCGATTAAAC ATGTTGACCA TTGCACTCAG TACAATAATA CCTAGTATCA	1080
	ATACAATACC GTTTTGATGT TCTAATAATT TGGCACCTTT AACAGCAGCG ATAATTCCTA	1140
45	AATTACTCCA CTTTAAATAC GCAAGTAGCT GTGCTGCAA AAACACAATA ACGATAAATG	1200
	TTCCCATTGA TCCTACAGCA TCGCCGAACA TTTTACCTAA GTCTTTTGTA TTTTTAATTT	1260
50	CTTTGCTTAA AATCCCATAA ACTAATCCAG GTACTAAAAA TACGACAAGA ATAATTAATC	1320
	CGACACCGTT AATTAATGGC GCATCGTCTA GTAAGCTGCC TGTTTTAGCA TTTCTTAAAA	1380
55	AGCTATGTTT AGGAATGGCT GTAATAATTA ATAAAATAAT TGTGACTATG AAACTGATAT	1440

	CATCATGCAT TAAACTGTCA TCATATTTTC CTAATCTAGG AATAATGAGC TTAGTTGTAA	1560
	CTAGCAATAT CGTAGGAAGT AATACAACGA CACTCGCTGC GATAAAGTAC CAGTTCATAG	1620
5	CAACGTTTGT TTTAATAGAA TCTGAAACGA TACGTGTTGC CGGTTCTGTA AATGAATAGA	1680
	CCAAAGCATC TTGCATACCA ACAACTATAT TTGCTGCAAA TCCTCCAACA GCGGAAGCAT	1740
	ATGCCATCGT TAGTCCAGCG ATAGGGTGAT AGCCAATTTT AATAAAAAGC ATTGCTGCAA	1800
10	GCGGCGGCAA GATAATTGTC GCAGCATCGC CGGCTGTACT ACCTAAAATA CCAATTAATA	1860
	TAATAGTCGG TAAAATTAAG AAACGTGGTG CGCGATTAC AACAGAAATC ATTAAC TTAT	1920
	CGAAGTATCC TGTTTTCTCT GCAACACCAA TACCAATCAT CACTGCTAGT ACTAAGCCTA	1980
15	ATGCTGGGAA CTCTGAGAAA TTTTAAATCG TATCATTCAT TATCATCGTA AATCCATCAT	2040
	GGCTAATTAT ATTTTAAATA TAAATGGTTT GATGCGTACC TGGATGCTTA ACAGATACAT	2100
20	TAAATAATGA GATAACCCAT GTCATAATGG CTAAGCCAC ACACATTAAA AAGAATAAGA	2160
	CGCTAGGATC TGGCAATTTA TTTCCGATTT TTTCAACACT ATTCAAGAAA CGATTGACGA	2220
	TAGACCCCTT TTGTTGATGT TTTGATGTCA TCAATTATTC CCCCCTTTGT TAAATATTTA	2280
25	AAGTSTAACA AAAAATACTC TCAAAAGTAA CAATTTTCAG GAAATAAAAA AACTAATATT	2340
	GTTAAATATT TTGAGTTATT CAATAGAAAG TGTATAGCAG AGTAGTTAAG ACTGCCTGAA	2400
	GACTTATCTA TTAGGTTTAT GAAGCATCGA ACAGTGGAAA ATAAGGACTG TAAGTTTAAG	2460
30	ATATGTTGTA TAGGAGTGAC TGAATGAAAC GTTTGGAAAA TAAAGTAGCA GTCGTAACAG	2520
	GAGCAAGTAC AGGTATCGGT CAAGCTTCTG CAATCGCTTT AGCTCAATAA GGTGCGTATG	2580
35	TATTGGCGGT AGACATAGCT GAAGCGGTAT CGGAGACTGT CGATAAAATT AAAAGTAATG	2640
	GTGACAATGC GAAGGCGTAT AATGTGGATA ATGCAAGCGA ACAACAAGTG GTAGACTTTG	2700
	TGTCTGACAT AAAGGAACAG TTTGGAAGAA TCGATGTGTT GTTTAATAAT GCCGGTGTGG	2760
40	ATAATGCGGC TGATAGAATT CATGAGTATC CAATAGATGT GTATGACAAG ATTATGAATG	2820
	TAGATATGCG TGGGACATTT TTAATGACGA AAATGATGTT ACCTTTAATG ATGAATCAAG	2880
	GTGGCTCTAT TGTTAATACG TCTTCATTTT CCGGACAAGC AGCAGACTTG TATCGCTCTG	2940
45	GATATAATGC TGCGAAAGGT GCACTGATTA ATTTTACAAA ATCAATCGCA ATTGAGTATG	3000
	GCCGTGATAG CATTGAGCC AATGCGATTG CACCAGGTAC AATTGAAACC CCGTTAGTAG	3060
	ATAAACTGAC AGGTACGAGT GAGGATGATG CAGGTAAAC ATTTAGAGAA AATCAAAAAT	3120
50	GGATGACTCC GCTGGGACGT TTAGGTAAAC CAGAAGAAGT TGCTAAATTA GTAGTCTTCT	3180
	TAGCATCTGA CGACAGTTCA TTCATAACTG GAGAGACGAT TCGAATTGAT GGTGGTGTGA	3240

55

AATTATTTAA AATCGATATC CAGTGGAAAA GAATTTGGCA TGAAGTAGGC AATGTATGCA 3360
 TATTGTCGCA ATGAAAGACA TTAAAGCGG AGATTAAATTA ATCTTTGAAA ATCACATATT 3420
 5 GTTCATTGA AGTGATTGC TTAAAGCTTG TTTAACGTA TTGTAGGTCG TGCATCTAAT 3480
 TTATCAGAAT CTCGAGTTCA AACTTTTGGG GTATCTTTAA AATAAGCTGT ATTTTGTCTA 3540
 TTCTAATAAA TTAAGGAGAA TTTTATGTTA AAAGAAAAAG AAAGTTTTAG ATTGCTATAT 3600
 10 CAAGCTATAA GAGAGATTGC AGATAAAATT GGAGATAATC AGTTAGAAAC TAATTCCGTT 3660
 AGTTTATTAT TATTGGACTT TGATTTTGAA CATGAAGTAT TTGATGAATT GTATCTTGTG 3720
 15 ATTTTAAAAAT ATTTAAATAC AGTAAGTATA GAGAACATAA GTCATAGTGA GCTT 3774

(2) INFORMATION FOR SEQ ID NO: 661:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1078 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 661:

TACTGGTTTT GGATTTTGGG GGTAATACAG TACCTAATAG TAATnATGGT GTCGTCAnAA 60
 TTATATCTTT CCGTGATGTT CTTACTCATT AGAACATCnG CCTTCAGAGG AATCATGATA 120
 30 CGAGGAATAA GAAATTTAAA TGTGAGCGAA GTCAATATAG TATTTGCGAT TATTTTTATT 180
 AACTCCATTA TTATTGTTAG TTTGATTTTT CGAGGATAAC TTCAATTTTT GCATTTTGAG 240
 35 GTTTTTTAAC ATATCTATTT GCATCAGTTG ATGGCAACCT TTTACTTAAA TCTATTGTGT 300
 AGTTATTGTC TGCACCTGTT ATTTTAATTT GTCCTTTATT ATAAGAATTA TTATATAATT 360
 TTTTACTTTT AATTAATGTT TGACGAATAC GAAAATCTAA TTCTTTTAAA GTTAAACAG 420
 40 GCTTATTGCC TTCATAAACT GGAAATCCGC CAGTAAACGT TTCTGCTTTA TCTTTATATG 480
 TTACATTCAG TTTATAGTGT TTATCGTTAG ATGTTGCTGC AGGAGTAACA CCACCAGTAA 540
 ACGTTTCTTG AGATAATGCA AAAGaATCAA TGGTTTCTTG GTCTTTTATG CyAAAAATAT 600
 45 CAACGCTTTT ATTTCTTAAT TGGTTGATAT TGCCCCAACT TTCAGGTCCA TAAACTTGAA 660
 TATGACTATA CCAAGaAAAC TGTAACAACG TTGCATGAAT CgTACCGTTA TCTTTTTGCC 720
 ATAACGTACT GTTAGAGAAG GTTAAATATT TTTGCGAGTA ATATTTAGTT AACTCATTA 780
 50 CGTTAGTTTC GTTTTGATTT ATATAaAAG cTTTCGCTTC AGATGAAGAA TTGATAkGTG 840
 TATTAGGAAA TTGTGTAGAT GCTGTACCTA ATAGTAACAA TGTTGTTGAT AAAATAATTT 900

55

ATAAAAAGGG GTTAATTAGA TAATTGAAAT TATCCGCATT TACAAAAGGT AATAGGTTAG 1020

TTAGATTTTT CGAGTATGAC TcAATTTCTG CATTACGAGG ATTTTAAACA TAACGGTT 1078

(2) INFORMATION FOR SEQ ID NO: 662:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1398 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 662:

AAC TTGCTTT TTAGTATCTA CGTTAATATT AATAGCTTTA ATTTCACTTG TATTAATTAA 60

ATCAGCTGTG TAAATACCTG CTTTCAAATC GATAACTTTC TTGTTCCAT TTTTAAAGTA 120

AACAGTATAT TTCGCTTGCT TCGATAGTCT TAAATCTATA TCACTAATAC CTCTGTCTGA 180

TTTTAAACT GATTAACTC TATCCTCTAA ATCTTTATAA CTAATATTTT GATTCTTATT 240

AAATGTAAAG CTTGATAAAA TATTTTGGCT TGTACCGTTC ACAGTGaTTG CATATGGAAC 300

ATGGACTTTA GAATATCCAT GGTGTAACGA ACTTGATGAT TTATCTAATG GCTTAGCTGC 360

GGCAGACGCT TCATTATTAT TAAAGTTTGC ACCTGTTGAT GCTAAAACAC CTAATGCTAA 420

AGTTGTTGTA ATCAATGACT TAAATTTTCAT AAATTATCTC TCCTTTTTTG TGTAATTCGT 480

ATTTGCAACT TAATTATAGC CAGACTTTCT CTATTTTTTG AATTAAGTGA ATATTAATAA 540

TAAATTATCT TTAACAATAA TTTTTTAAACA CTGTTAAAAG TTCTTTTAAT TTTGATTAAAC 600

TAATTAATTT ACAATACCTA AAATGTTGTT TGGTTTTGTT TATACCAAGC TTCAAACCTTA 660

AATGTCATAA CAACATTCAT TTCTTAATTC CTATTAGATT TGTCGATTAT ATTTACAGCA 720

TCCTTATACT CAAAAACAT TTAAGTAAAA ATATAAATTC GATTTAATAA TTAATTTAAA 780

TTTAGTTAAT CAATTTTGCA TCTATTTTGT TGTAAGCTAT ATAAAAGGAG TGATAATGAT 840

GGTGAAAAAA ACAAATCCA ATTCACTAAA AAAAGTTGCA ACACTTGCAT TAGCAAATTT 900

ATTATTAGTT GGTGCACTTA CTGaCAATAG TGCCAAAGCC GAATCTAAGA AAGATGATAC 960

TGATTTGAAG TTAGTTAGTC ATAACGTTTA TATGTTATCG ACCGTTTTGT ATCCAAACTG 1020

GGGGCAATAT AAACGCGCTG ATTTAATCGG ACAATCTTCT TATATTAAAA ATAATGATGT 1080

CGTAATATTC AATGAAGCAT TTGATAATGG TGCATCAGAC AAATTATTAA GTAATGTGAA 1140

AAAAGAATAT CCTTATCAAA CACCTGTACT CGGCCGTTCT CAATCAGGGT GGGACAAAAC 1200

TGAAGGTAGC TACTCATCAA CTGTTGCAGA AGATGGTGGC GTAGCGATTG TAAGTAAATA 1260

CAACAAAGGC TTTGTTTATA CAAAATAGA GAAAATGGT AAGAACGTTT ACGTTATCGG 1380
TACACATACA CAATCTGA 1398

(2) INFORMATION FOR SEQ ID NO: 663:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 663:

TTGTAATTGG AACGTACGAA CTTTTCTAGG TAAGAAACGT CGAATCTCGT CCTCATTATA 60
ACCAACTTGT AGTCGTTTAT TATCTAAAAT AATTGGACGA CGTAATAAGC CAGGATTATC 120
TTGAATGATT GAATATAAGT CTTGTAATGG TAGTGAATCA ATATCAACAT TTAATTTTTG 180
GTATGTTTTA GAACGTGTAG AAATGATTTT ATCAGTACCG TCTTCAGTCA TTTTAAATAT 240
TTGCTTAATT TCATCAATTG TTAAATGTTT AGAAAAATA TTACGCTCCG TATACGGAAT 300
GTCCATGTTT tTGkTAACCA TGCTTTCGCT TTTACGGCAA GATGTGcAaC TTGGtGaAGT 360
aAATAATGtT ACCATACATC TCACTCTCCT ATTTGAATGA ATAAAATTCA TTGCTTAAAA 420
TTTAGTTATA GATCAAGAAA AAACATTTTT TTCTAAAATT CTTAATCGTT ACTATTTATT 480
ATACTATCT AACATTAAAA TTAAATGAGA AAAACCTAAT TTTTCAGATA AGTTTCTACA 540
CTTATAAAAA AGATTATTAA TCCCTTTGTT AGTAGTAAGT TATACGTATA TTCTAACACA 600
TCTTACATTT TTAAGAAATA CTGTTATAAT GATAATTATT AAAATATTAC TAAGAAAGTA 660
GGCATTTAAA TGGAGACATT ATTTTCAGGC ATCCAACCTA GTGGAATTCC TACTATTGGA 720
AATTATATTG GCGCACTAAA ACAATTTGTT GATGTGCAAA ATGACTATGA TTGTTATTTT 780
TGTATCGTAG ATCAACATGC AATTACAATG CCACAAGATC GTTTAAAATT ACGTAAACAG 840
ACCAGACAAT TAGCAGCGAT TTATTTAGCT TCTGGTATAG ATCCAGACAA AGCAACATTG 900
TTCATACAAT CTGAAGTCCc TGCACACGTA CAAGCAGGAT GGATGTTAAC TACGATTGCT 960
TCTGTTGGAG AATTAGAGCG TATGACGCAA TACAAAGATA AAGCTCAGAA AGCAGTTGAA 1020
GGTATACCTG CTGGTCTATT AACATATCCA CCTTTAATGG CAGCTGATAT TGTTCTTTAC 1080
AATACTAATA TCGTTCCAGT TGGAGATGAC CAAAAGCAGC ATATCGAATT GACTCGTAAC 1140
CTTGTAGATA GATTTAATAG TCGCTATAAT GATGTGCTTG TGAACCTGAA ATTCGTATGC 1200

(2) INFORMATION FOR SEQ ID NO: 664:

(A) LENGTH: 787 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 664:

10	TGTACCTTTT GCTTTAATAA ATACTGTTTC TTTATCATAA TTAGGTTTCAG TTAAGAAATT	60
	AAATTGTAAG CTTTGAGTAA TATTTTTTTG ACTATCACTT GTTGTAGCTG TACGTGTATA	120
	CATTTTAGTA TCACCATCAA GATTCTTCTC AGAAACTTGT TTGATTTCAG AATTAATCTT	180
15	TGCAAATGAA GTAGCTGGTA ATACAGTGAA AGTAGTCGAT AGTGCTAAsG WACAAATTGT	240
	GATATTTTTTA CATAGTTGTT TAATCATTAG TAATCCGCCC TTTCAATATT ATCCTTCTTT	300
	ATAAGGTTTA TTGTCATCAG AATATTTATC AACGACTTTA ACTGTTTTAT TTTTCCAATC	360
20	AACCTTCATAA GTGACAATTA ATCTTTGACC ATCTTTATTT TtctCTAAAA TTGGAGGTGC	420
	ATAATGTATy CCAGGTCTGT TTTTCAAAAT ATCTTGATTT CGTGTGTATG TTACTTCAAA	480
25	TTGCGTTTTTC TCATTTGACT TTTCATTAGA TAAATAAGTT AAAAATTCTG GATTAAAGCC	540
	ACTTCTTACT AATGCTGGGT ATCTATATTT TGAAGCAAAG CTTAGTTCaG GGTTTTCTAC	600
	AGTAGCAATT CTCGTATTTT TATAGAATAA TAATTCATCA TTTCTATTTT TCACTTCTCC	660
30	ACCATACTTC AAGTCATTCG CAATAACTGA CCAGTGTACA TGgCCAGTTn ATTTATTTTT	720
	ACCGCTGGCA ATTGTGTCAT AATTTGCTGC ATCATCAACT AATCGTTTTG GAAGTAGCTA	780
	TTTGCTG	787

35

(2) INFORMATION FOR SEQ ID NO: 665:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 533 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 665:

45

	TTATCTGTCA TTAACCACTG GTACATGTGA ACCCGGCGnTn AATCAATATA TATATTTAAA	60
	AGCAAAGGCG CGCCATGTGC CTTTtTnTA tTTGTATAT CTGTATCAAA ATCGATTTGA	120
50	TTAAATCCG CTTTATTTTA TCATCTATTC AAATGATTTT AGTGCCTTA TTTTACTATG	180
	GCATTATTGC CTAACCTGTT TGAGATATAT TAATATTTGA TGATTCGTAT TCAAAGTTTT	240
55	CATTTAACAT GTATTTAGTA TCATGATAGC TGCTTCATTG ATGATATACT ACCTACTTTT	300

CAGAGGCAAT ATTGAACAAT TCTTCATCAT GACTATTTTC CATCACATAA CTATGCTTAG 420
 CGAACGCTAA CATATCTTTA TCATTATTCG CATCTCCGAA GGCCATGAGC TCTGAAGGAG 480
 5 ACATTTCCCA TTTATCTAAC AATCGTTTTA ATGCCTGnCC TTTAGTCATG TTT 533

(2) INFORMATION FOR SEQ ID NO: 666:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 666:

GACTTCTCTT GCACTAAAAA TTGCTCAAAT TGCTCGCTGT CATAACCAGG CGGTGTTTCC 60
 20 AACCATACAT ATATGCCACC TTTAGCATGA ACAAATGGCA AATCAGCTTT TGCAAGCATG 120
 GCTTCGAATC GGTCACGACG TGTTTTAAAT ACATGCTTT GTTCTTCTAA AAAATCATCA 180
 25 TAATGGATTG AAAGCATATA TTGCGGTCAT CTTGTAAATG CACCAAACAT CCCAGCATTT 240
 GTGTGCGTTT GGTACTTTTT CAAAGCTTGA AATCATATCT TTATTACCAA CTGCCAAAAC 300
 CGACTCTGAA AACCTGGACA TGTTGATGAA CCTTnAGAC AAGGGAAGAA ATTTGCAATC 360
 30 GCAACATCTT TGCCCATTTT nCCGAAGCAA GTGnGACTAG 400

(2) INFORMATION FOR SEQ ID NO: 667:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 412 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 667:

TCTCCAATGC TACTCAAATT AAAAAGGTTT TAAATATTGT TAATTCTGAT CCAGAGCGAA 60
 45 AGATTGTTAT CGTTTCTGCT CCAGGTAAAA GACATGATAA TGATATTAAA ACAACTGATT 120
 TGTTAATCAG ATTATATGAA AAGGTCATTA ATCATCTTGA TTATCATGAT AAAAAAGAG 180
 AAATTATTCA GCGTTATGAT GATATTGTAA AAGAATTGCA AATGGATGAA AGTATTTTAC 240
 50 GGACGATAGA TGTGACTTTG GAACATTATA TAAATCAATT AAAAAATGAA CCAAAGAGAC 300
 TATTAGATGC ATTACTTTCT TGTGGTGAAG ATTTTAATGC GCaYTGatA GCyTTATATA 360
 ATAAtAGTCa gGtTACCAAC AAmATwTATA TCCcCGAAGG AAGCGGGTAT TT 412
 55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 668:

TTACTCGTTA AAGATATTGT AAAAGATGAA GTGACAGAAT ATGACATTCA TCAAATGTTA 60
 CCGCATCCGA TTAATATGGT AAGGGTTAGA CTTTTTGGTG TGAAATTAAA AGAGATTATA 120
 GCTAAAAGTA ATAAACAAGA nTATATGTAT GAACATGCAC AAGGTTTGGG TTTCAGAGGG 180
 AATATATTTG GAGGATATAT TCTTTATAAT TTAGGGTACA TTCATTCTAC AGGGCGTTAC 240
 TATCTGAATG GAGAAGAAAT CGAGGACGAC AAGGAATATG TACTAGGTAC GATAGATATG 300
 TATACGTTTC GTcNTATTnc CCAACATTGA AGGATTACCA AAGAGTATTT AATGCCAGAG 360
 TTTTTCAGAG GTATATTTAA nGAAAATTAT TGGCTATTTA 400

(2) INFORMATION FOR SEQ ID NO: 669:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1236 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 669:

TGTTTCGATAT TTTTAAATTT ATCTTTTAAA TACAACAACCT GTTTCCTGTA TGATTTAACT 60
 GTTTTATGAC TAATGCCATT GAATATTTCT AGCGTTTTAT TTAACCTATC GATAATCGCA 120
 TGTAATCCT TCAAATGTC TTTTGTTC AAGTAAATA CATTATGGAA GCGATGAATA 180
 TCATCATCAT AAACATCAGA ATCATTGATA ATCGTAAATA TCGTTGAGAA CAATTGCTCA 240
 TTTAACTCAT GAATCTCATT CATACTAGCC TTCAAGCCAA AAATATCAAT TGGTGCAATA 300
 TCTAATTTTT CCAAATTCG CTGCTTTTCC AGTTGATCAA TTGCCTTTAA CAATTTTTCA 360
 TTTTCGTTTT TACCAATCAA ACCAAGCTGA TATTTAATAT CAGCATAACT CAACTCATTT 420
 GTCACITGAT TTAAGGCATA GTCTGGTAAG CGATGTGCTT CATCCACTAT ACAATCATCA 480
 AACAATTGAT ATATTGAATT TTCAACATCA GAATGaATTA AATGTGCATG ATTTGTAATA 540
 CCAATTTGaA TGTTCTGTGC ATTTGCTTA ATAAAATTAT aATAATGAAC ATCGTGACGT 600
 GCCGGTACAT ATGTTTCAAT TTTCTGGTca AAATACATCT TTTGACCACC TTTTAAATTT 660

5 ATATTCACTt CGTAATTACT TGTGTCATCT TTTAAAATTT GACTAATAAG CCCCAATGAA 780
 ATGTAATCaC TTtTACTTTT AATCAATAGT GCATTAATTT TAAAATTCAA CGCTTCATTC 840
 ATTGCTGGAA TATCTTTTTT TAACAATTGA CTTTGCAGTA ATTTAGTATT GGTAGAAATC 900
 ATGACATGCT TCCCAGTTTC AATATTATAC ATCAAGGCCG CAAGTAAATA TGCTAATGAT 960
 10 TTACCACTGC CTAGTGATGC TTCAATCATT GCTTTTTTAC TATGCATGAG CTGATCTAAT 1020
 ATAGTTTCCG CTAAATATAA TTGTTGCGGT CGATATGTTA AGCCAAGTTG ATCTACAGCT 1080
 TTGCTATATA AAGACTTCAA GCTGCCATTA TAATTTGTTG TCGGCTTTTT AAAATCAACT 1140
 15 TGCTTACGAT AGATAATCTG TTCGAACTTT TCGTACGATT TATCCAATGG CTTTGCATCA 1200
 TATTGCCTAA CCATCTCAA GAAAATATCA TACAAA 1236

(2) INFORMATION FOR SEQ ID NO: 670:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1819 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 670:

30 ACAsAATTAT TGACCAATAT GACTCGTGGA CTGATATGTT TAAAGCACTA CTGCATGAAA 60
 CATTTAAAGC ATATGGCGTT CTATTTATAG ATGCGCAGTT TGAGCCGTTA AGAAAAATGG 120
 AAGCGCCTAT GTTTAAAAAG ATTTTGAAAA AACATCAGTT GCTTGATGAT GCTTTTAGAG 180
 35 CAACACAACA ACGTACTCAA AATCAAGGCT TGAATGCGAT GATACAAACA GATACAAATG 240
 TTCATTTATT CTTACATGAT GAAAATATGC GTCAATTAGT tTCGTATGAT GGTAAGCAYT 300
 TTAmATTAAA TAAAACAGAT AAGACATATA TAAAGGAAGA AATTATAAAT ATTGCGGAAA 360
 40 ATCAACCKGA ATTATTTTCT aATAATGTAG TGACAAGACC ATTAATGGnA GAATGGTTAT 420
 TTGAACACGG TGGCATTGTG TGGAGGACCG AGTGAAATTA AGTaCTGGGC TGAACATAAA 480
 45 GATGTATTTG AACTATTTGA TGTGAAATG CcATCGTGA TGCCAAGGCT TAGAATTACT 540
 TATTTAAATG ACCGTATAGA AAAATTACTT TCGAAATACA ATATTCCATT AGAAAAAGTG 600
 TTAGTCGATG GTGTTGAAGG AGAAAGAAGT AAGTTTATTA GAGAACAAGC ATCACATCAA 660
 50 TTTATTGAAA AGGTAGAAGG TATGATTGAA CAACAGCGTC GTCTAAACAA AGACTTATTA 720
 GATGAAGTGG CGGGGAATCA AAATAATATT AACCTTGTGA ATAAAAATAA TGAAATTCAT 780
 ATACAACAGT ATGATTATTT GTTAAAACGT TATCTTTTAA ACATTGAAAG AGrAAACGAC 840

55

5 GAAAGAATAT GGAATCCACT TCAAATTTTG AATGATTTTG GGACAGATGT GTTCAAGCCC 960
 TCCACCTATC CACCACTTTC TTACACTTTT GATCATATTA TTATAAAACC TTAATATACC 1020
 AAGGGTTTAG CCCGATTTAT CTTAATGATA AATCGGGCAT TTTTTGTTT TTTAAATAA 1080
 ATTTACAAA TTTTGTATAA ATAGTGGTGG ATAGTGGGGA GATGTGGTAA ATTATATATA 1140
 10 AGGTGAGGTG ATAAAAATG TTCATGGGAG aATACGATCA TCAATTAGAT ACAAAGGAC 1200
 GTATGATTAT ACCGTCCAAG TTTCGTTATG ACTTAAATGA GCGTTTTATT ATCACAAGAG 1260
 GCCTTGATAA ATGTTTATTC GGTTACACTC TAGACGAATG GCAACAGATT GAAGAGAAAA 1320
 15 TGA AACCTT ACCTATGACA AAAAAAGACG CACGTAAGTT TATGCGTATG TTCTTCTCTG 1380
 GTGCTGTTGA AGTAGAACTT GATAAGCAAG GCGGTATTAA CATCCCTCAA AACTTGAGGA 1440
 AATACGCTAA TTTAACTAAA GAATGTACAG TAATCGGTGT TTCAAATCGT ATTGAGATTT 1500
 20 GGGATAGAGA AACTTGGAAT GATTTCTATG AAGAATCTGA AGAAAGTTTC GAAGATATTG 1560
 CTGAAGATTT AATAGATTTT GATTTTTTAA ATGGAGGAAT TGAAGtGTTT CATCATATCA 1620
 25 GCGTTATGTT AAACGAAACC ATTGATTATT TAAATGTAAA AGAAAATGGT GTGTACATTG 1680
 ACTGTACGCT AGGTGGAGCG GGACAnGCCC TTTATTTACT AAATCAATTA AATGACGACG 1740
 GAAGATTAAT AGCAATCGAT CAAGACCAAA CTGCAATTGA TAATGCTAAA nGGGTATTAA 1800
 30 AGGATCATTT GCATAAAng 1819

(2) INFORMATION FOR SEQ ID NO: 671:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 609 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 671:

45 ACCAATTATT GGACAGGAAT TAGCAAAAAA TGCAATGCTT GCATTAATCT ATGCATCGAT 60
 AGGTATCATC ATCTATGTAT CATTACGATT TGAATGGCGC ATGGGTCTTT CATCTGTATT 120
 GGCATTATTA CATGATGTAT TCATTATAGT AGCGATTTTC AGTTTATTTA GAATTGAAGT 180
 AGATTTAACA TTTATCGCCG CTGTATTAAC AATTGTCGGT TATTCAATTA ATGATACAAT 240
 50 CGTAACGTTT GACCGTGTAC GTGAAAACCT ACAAAGGTT AAAGTGATTA CGACAACAGA 300
 ACAAATTGAT GATATCGTTA ATAGATCAAT tAGACAGACA ATGACACGTT CAATTAATAC 360
 55 AGTATTAACA GTTATTGTAG TAGTAGTTGC TATACTATTC TTCGGTGCTC CTACGATATT 420

TGCCGTTCCG CTATGGGGAA TAATGAAAAA ACGTCAGTTG AAAAAATCGC CGAAACACAA 540
 ATTAGTTGTA TATAAaGAAA AGAAATCGAA CGATGAAAAG ATTTTAGTTT AAAAtGaATT 600
 5 AAGCGGTAT 609

(2) INFORMATION FOR SEQ ID NO: 672:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2222 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 672:

CTTAAAACAG CAATTTGAAC GTTTACAAAA TGAACAAATC TTTGTTTATG TTTGTCATGG 60
 20 TAATCAGCAT CCTTTATCAT CAAAGATTTC ATCAAACCTGG CCAGATAATG TtntTGTATT 120
 TTCAAATAAA GTTGAGACGT ATGAAGCAAT TACTAAATCT GGTGAAACAA TTTATATTCA 180
 CGGATTTAGT TATGAAAATA GAGCAAGTTA TGAGAACAAG ATTGATGAAT ATCCATCAAG 240
 25 TCAAGGCCAA AAAGGCATAC ATATTGGTGT CTTGCATGGT ACGTATAGTA AATCTTCAGT 300
 TAACGAAAGA TATACCGAGT TCATTTTAGA AGATTTAAAC AGTAAATTGT ATCATTATTG 360
 GGCTTTAGGT CATATACATG AACGTCAACA ATTAAGTGAT ATGCCTGTAA TTAACCTATTC 420
 30 AGGTAATATT CAAGGTAGAC ATTTTAATGA GCAaGGTGAA aAAGGTTGCT TATTAATCGm 480
 GGGTGACCAC TTAAaATTAA AGACTAAATT TTATCCTACA CAGTATATTA GATTTGAAGA 540
 35 AGCAACTATT GAAACGGATA AGACATCTAA GCAAGGTTTA TACGAGGTCA TTCAAAACCTT 600
 TAAAGAACAA GTGAGAGAAG AAGGAAAAGC CTTTATATCGT TTAACGCTTG TTATTAATAG 660
 TGAGACATTA ATTTACCTC AAGATTTATT ACAAGTTGAA GAAATGATTA CAGATTATGA 720
 40 AGAAAACGAA AATCAATTTG TATATATTGA TGAGTTAAAA ATACAATATG CACAAAATGA 780
 TGAGTCACCT TTAGTTAATG AATTTTCAGC GGAATTATTA GTCGATCAAA CTGTTTTTGA 840
 TAAAGCGATG TCAGATTTAT ATTTAAATCC AAGGGCATCT AAGTTCCTAG ACGATTATGG 900
 45 AACATTCGAC CATACAGCAT TAGTTAATCG TGCTGAAGAA ATATTAAAAG CTGAAATGAG 960
 AGGTGAACAA AATGATAATT AAATCACTTG AAATTTATGG TTACGGTCAA TTTGTTCAaC 1020
 50 GTAAAATTGa ATTTAATAAA AAcyTCaCTG AAATTTTTGG TGAAAATGAA GCGGGTAAAT 1080
 CGACGATTCA AGCATtCATC CATTCGATAT TATTTGGATT TCCAACTAAA AAGTCTAAAG 1140
 AGCCAAGACT AGAACCACGT CTAGGTAACC AATACGGTGG TAAATTAGTA CTTATTCTTG 1200

TATATTTACC TAATGGTGCT GTGCGTGATG ATGCTTGGTT ACAAAGAAA CTTAATTATA 1320
 TTTCTAAAAA GACATATCAA GGTATCTTTT CATTTGATGT ACTAGGGCTT CAAGACATTC 1380
 5 ATAGAAATCT AAATGAAAAA CAATTGCAAG ATTATTTATT ACAAGCmGGG GCTTTAGGAT 1440
 CAACTGAaTT CACGTCAATG CGCGAAGTGA TTAATCGTAA AAaAGATGAA TTATATAAAA 1500
 aATCAGGTAA AAATCCGATC ATTAATCAAC AAATTGAGCA ATTAAAACAA CTAGAAAGTC 1560
 10 AAATTCGTGA AGAAGAAGCA AAGCTAGAAA CATATCATCG CTTAGTAGAT GATCGAGATA 1620
 AATCATCACG TCGATTAGAG AATTTAAAGC ATAATTTAAA TCAATTATCA AAAATGCATG 1680
 AAGAAAAACA AAAAGAGGTT GCTTTACATG ATCATTACA AGAATGGAAG TCTCTAGAAC 1740
 AACAGTTAAA TATTGAGCCA ATCACATTCC CAGAAAAAGG TGTGGATCGT TACGAAAAAG 1800
 CACGAGCGCA TAAGCAATCG TTAGWAAGAG ATATTGGTTT AAGAAATGAG CGTTTAGCTC 1860
 20 AACTTAAAGA AGAAGCGACT CAATTAGAGC CAGTTAAACA ATCTGATATT GACGCCTTCA 1920
 TTAGTTTGAA TCAACAAGAA AATGAAATTA AAAATAAAGA ATTTGAACTT ACTGCAATCG 1980
 AAAAGGATAT TGCGAATAAA CAACGTGATA AAGATGAATT GCAATCAAAT ATTGGTTGGT 2040
 25 CTGAAACGCA TCATGACGTA GATAGTTCAG AGGCAATGAA AAGTTATGTC AGTGAGCAAA 2100
 TCAAGAATAA ACAAGAACAA GCTGCATACA TTAAACAATT AGAACGTAGT TTAGAAGAAA 2160
 ATAAATCGA AGATAATGCG GTTCATAGCG AACTAGATTC TGTGAAGAA AAATAGTTCC 2220
 30 TG 2222

(2) INFORMATION FOR SEQ ID NO: 673:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 406 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 673:

45 TTGCAGGTAT CATTTTTTTA ATGCCATATG GATTATGTTT TCTACCGTTT TATAAGCAAA 60
 AAAAGAAAAA ACAGACATTT AAAAAATACA TGGTTTACAC TACGATTGGT TTGTCAATTT 120
 GTCTAGGCTT ATCTCTAGTT TTGGTTCACA CTACGAAAAT TTATATGGAC GAAGGTGGCG 180
 50 TAAGATACTA TTACGGTAGT TTTGTAATGA AACAAGCGGG CGGTTATGCT TATTTAGCTT 240
 TAGCGGTACT TTCAACGTTG TTAATTGTTG CGAAAAAGC TACAAATAAA AATAAAGAAA 300
 TCGAAACCGT CGACAATACA AATATAACGG AAAGATAATT AAGGGAGTGC TCATTCAGGA 360

(2) INFORMATION FOR SEQ ID NO: 674:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 927 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 674:

```

CCACTTTGTG GATTTTCTTT ATATTCTCCA CGCTTGtACT ACAACTTCTT TCGTTTCTGT      60
TTCATCACCG ACTGCCGCAT TCGTTAGCAC ATGTAATAAC TCTTTTGCGG TTAATACATT      120
CTCATCTATA ATCTTATCTT TTTGTTCTTG TATATATTGC TTGATGTGCG GCTTTTTCAA      180
TAACCTACAC GCTGTCACAT GTGCGCTATT TGCCTTATAT CCTGyTTTtA TGGCACTTTG      240
TGTTACATTC AGTGTtCTAA TATACTCATT CACAAAACGT GCTTGCTTTG CAGTTAACTC      300
ACTCATTtTA TCACCCCCAC AATTTTATCT AATATGGTTT CATACCATAA TATTACAGAT      360
TGTTCTGAAC AATCTAAGGC ACTACTAATA TCTTGATAAC TAAGTCCTTG TATAAGGGAG      420
TCAAAAATAT AAAACTCTTT ATCGGTGCGT AATCTGTCAA CAATCATtTC TATGTGATTC      480
TTTATAATAT GATCATTGAC ATTATCGTCT GTCATCAATT CGTCAGAATC TTCATCACCT      540
ATTAAAAAGA AATCATCAGT ATTTATTTCA TCATCGCCCC GTTAACTAGC TTTGAAGTCT      600
TTAGCACACT TGCATATACC GGCTGTGCTG CTGGCKAGAT ACTAGCATTG AAGTGGTCTG      660
ATATTGATTT TGAAAACAAC ACGATTAGTA TTAATAAAAC ATATTACAAT CCAAATAATA      720
ACAAGAAGAA ATATCAGATA CTTCCCCCTA AACTGAAAG TTCTATCGGT AAGATTtCCG      780
TTGACCCAAA TGTAATAAAG GtGtTGCGtG ATTATAAGAT AAACGTTCAA AATAATTGGA      840
AAAACGAATT ATATAATGAT AACCATTGTA ATGAAAAAAC TATCAATGTG GATACAAACC      900
ATTATGTCAC GTACCAGCAT ACTCAA                                     927

```

(2) INFORMATION FOR SEQ ID NO: 675:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1100 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 675:

```

GATGTTAAaA TAAGTACATC ATTTTGACTC CCAAATATTG GCTTTAGACC TTGAAATGCT      60

```

CTCATAATGG CATCAGGTAC cAGGGGTTGG GaCCAGGTGT TAATAACAAC GGTTGATGAT 180
 AATACATTTA TATGCCTCCT ATAAATAAGA TTCTACTATT TTAGCAAATT TTCTGAAAAT 240
 5 TTAAGAGCCT AAAAAgtGAT AACGCTATAT TGTCGAACAA ACGTTATCAC TTAACGATTT 300
 TATTTGGCCT GAATGATGGG ATAGATTTTT ATGTCCACAT TATTTCTTAC AGCATTGAA 360
 ATCATGCAAT TATTATCTGC AATTGTTATC AATTTGGTA ATCGCTTTTC TAATTGTGCT 420
 10 ATTTGATCAC TTGGAATTTG AATTGAGGA TGGTGACAA TTTTGACAT ACTGAATTTT 480
 CCGTTATTTA AACAAGCTGT TCCAATCGAT TGTGTkCAA TTGAAATATC TGTGAACTTT 540
 GCACGTTCAA GAGTAGCTGC TAATGAGATG ATATAACATG ATGAAGCGGC TGATACTAAC 600
 15 ATTTTCATCGG GATTGTGTC TATACCAACA CCACCTAAAG AAGCAGGTAT AGAAATATTC 660
 TCTGAAAGTA TGTCGCCTTG AACGTTCCG ACATTGTTAC GACCACCTTG CCAAGAAGTT 720
 20 TGGACTTTAA AGTCATGTTG ATGCAATTGC TTTAACCTCC AATATAATTG TGATAGTTTA 780
 ATTTTAGAAT AGTTTATCAA AGTTTAACAG AAAGGTGACT TATCAATGAC TCTGAATAAA 840
 CTGAAAGATG AATTACAAAT TGTTTCGCAC CGTGGATTGC CGAGTGATTT TCCTGAAAAT 900
 25 ACAATGGTCG GTTATCGAGA GGTAATGGGG CTCAATGTTG CTATGTTAGA AATAGATGTT 960
 CATTGACCA AAGACCAACA TTTTGTGTG ATACATGaTG AAACAATTGa TAGAACATCG 1020
 GaTGGtArGG GGCgTAWTgc TGaTTACACA TTATCGCAAT TAAAATCATT TGATTTTGGT 1080
 30 AGTTATAAAG ATGTTGCTTT 1100

(2) INFORMATION FOR SEQ ID NO: 676:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 460 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 676:

ATTAATTTCA TATGGAAATA GTTGAACATA CTGCGCTGTA TAAGCTTCAA CAGTTTGATG 60
 45 AATTAACGAT TGATCTTCAA TATATCCGTA GAACAAATCT TCAGTACAAA CTACTTTACC 120
 TTTATCAGGT TTAATTGCAC CTGCCAACAA TTGACCTACC AACGCTTTGG AAGATTGAGG 180
 50 TTCACCAATT ATACCTAATG CTTCTCCTTG ATAAATATGT AAATAATAT TGTTTAAATC 240
 GATATCTTCA GCATCATATC CAAAAGGTAA ATACCATTTC TTATTCTGTT TATTCCTATA 300
 GTAGTGTGTT ACTTTTAGTA ACTTTAAAAC AATTGAACTT CCCATCTATT TTCATCCTTC 360

55

CCCCACGCaA AAATACCTTT TAATCTTnCT ACTTTAAAAT

460

(2) INFORMATION FOR SEQ ID NO: 677:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 677:

15	GTCAAGTTTT CGGCTAGATT TTAAACGCTC ATCACTATGT AGAATGAAAC TTTCGAAAAA	60
	CTGTAAATCA TCATAACCTT TTACATAAAC ATAACCTTCG CCACCAATTG CTTGAATTAA	120
	ACATTGGGCG GCCATTTGAA TTTCTAAAGA TTGTTTTTCT AGCCTATTAA AGATACCTAT	180
20	TAGTTGTGTG TTAAAGATTT TTGACATCTT TATCCTCCAA TCTACTTATA AAATATTGTA	240
	ATTAATGACT ACATATTATG CAACGGCTTA AATTGTATAA AAATGTATAC GTTTGCATTT	300
	AGTATAACTA TCGCATTTTT CAAAAAATAC ACATTTAATC TGCAGTATTT CAATGCATTG	360
25	ACGCTATTTTT TTGATATAA TTACTTTGAA AAATACGTGC GTAAGCACTC AAGGAGGAAC	420
	TTTCATGCCT TTAGTTTCAA TGAAAGAAAT GTTAATTGAT GCAAAAGAAA ATGGTTATGC	480
	GGTAGGTCAA TACAATATTA ATAACCTAGA ATTCACTCAA GCAATTTTAG AAGCGTCACA	540
30	AGAAGAAAAT GCACCTGTAA TTTTAGGTGT TTCTGAAGGT GCTGCTCGTT ACATGAGCGG	600
	TTTCTACACA ATTGTTAAAA TGGTTGAAGG GTTAATGCAT GACTTAAACA TCACTATTCC	660
35	TGTAGCAATC CATTTAGACC ATGGTTCAAG CTTTGAAAAA TGTAAAGAAG CTATCGATGC	720
	TGGTTTCACA TCAGTAATGA TCGATGCTTC ACACAGCCCA TTCGAAGAAA ACGTAGCAAC	780
	AACTAAAAAA GTTGTTGAAT ACGCTCATGA AAAAGGTGTT TCTGTAGAAG CTGAATTAGG	840
40	TACTGTTGGT GGACaAGAAG ATGATGTTGT AGCAGACGGC ATCATTITATG CTGATCCTAA	900
	AGAATGTCAA GAACTAGTTG AAAAACTGG TATTGaTGCA TTAGCGCCAc ATTaAGGTTCA	960
	GTTcATGGTC CATAcAAAGG TGAACCAAAA TTAGGATTTA AAGAAATGGA AGAAATCGGT	1020
45	TTATCTACAG GTTTACCATT AGTATTACAC GGTGGTACTG GTATCCCGAC TAAAGATATC	1080
	CAAAAGCAA TTCCATTGG TACAGCTAAA ATTAACGTAA ACACTGAAAA CCAAATCGCT	1140
50	TCaGCAAAAG CAGTTcGTGA CGTTTTAAAT AACGACAAAG AAGTTTACGA TCCTCGTAAA	1200
	TACTTAGGAC CTGCACGTGa AGCCATCAAA GAAACmGTTA AAGGtAAAAT TAAAGAGTTC	1260
	GGTACTTCTA ACCGCGCTAA ATAATTAATA TTTAGTCTTT AAGTTATTAA TAACGTAGGG	1320

AATAAATAAA ACAGTTTGAT TTTAAATGA AAGCGTAAAA ATGGTAAAT ATATCAAAAT 1440

TGATTGTGAT A 1451

(2) INFORMATION FOR SEQ ID NO: 678:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 668 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 678:

NGTATTGAAG CGGTTAAACA AACACCTAAT GCAACTGACG AAGAAAAGCA GGCTGCTGTT 60
 AATCAAATCA ATCAACTTAA AGATCAAGCA ATTAATCAAA TTAATCAAAA CCAAACAAAT 120
 GATCAGGTAG ACACAACTAC AAATCAAGCG GTAAATGCTA TAGATAATGT TGAAGCTGAA 180
 GTAGTAATTA AAACAAAGGC AATTGCAGAT ATTGAAAAAG CTGTTAAAGA AAAGCAACAG 240
 CAAATTGATA ATAGTCTTGA TTCAACAGAT AATGAGAAAG AAGTTGCTTC ACAAGCATT 300
 GCTAAAGAAA AAGAAAAGC ACTTGACGCT ATTGACCAAG CTCAAACGAA TAGTCAGGTG 360
 AATCAAGCAG CAACAAATGG TGTATCAGCG ATTAAAATTA TTCAACCTGA AACAAAAGTT 420
 AAACCAGCTG CACGTGAAAA AATCAATCAA AAAGCGAATG AATTACGTGC TAAGATTAAT 480
 CAGGATAAAG AAGCAACAGC AGAAGAAAGA CAAGTAGCAC TAGATAAAAT CAATGAATTT 540
 GTAAATCAAG CCATGACAGA TATTACGAAT AATAGAACAA ATCAACAAGT TGATGATACA 600
 ACAGTCAGCG CTGATAGCTT GCTTTAGTGA CGCCTGACCA TATTGTTAGA GCgCTGCTAG 660
 AGATGCGT 668

(2) INFORMATION FOR SEQ ID NO: 679:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1906 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 679:

GATCCAAATT TAAAGGAAA AATAGCCTTT AACGAATTTA CGAAACAAA TGAATGTTTA 60
 GGGAAAGTGC CATGGAATAC TAATTTTAAG ACACGTCAAT GGCAAGACGG TGATGATAGC 120
 AGTTTAAGAA GTTATATCGA AAAGATTTAT GACATACACC ATTCAGGTAA AACAAAAGAT 180

	ATCGTGGGAT GGACATAAAC GCCTTGAAAA GTTATTTATC AAATACTTAG GTGTTGAAGA	300
	CcTGaAGTGA ATAGAACAAC TACCAAAAAA GCATTGACTG CTGGAATCGC TAGAGTAATG	360
5	GAGCCTGGAT GTAAATTGTA CTATATGCTT ACACTTTATG GTCCTCAAGG TGTAGGTAAA	420
	TCTGCTTTGC TAAAAAAATT AGGTGGTGCA TGGTTTTCTG ACAGTTTAGT TTCTGTTACA	480
	GGTAAAGAAG CTTATGAGGC CTTACAAGGC GTTTGGCTAA TGGAAATGGC AGAACTTGCA	540
10	GCTACAAGAA AAGCTGAAGT TGAAGCTATT AAGCATTTCa TATCTAAACA AGTTGACCGA	600
	TTTCGTGTTG CTTATGGGCA TTATATTGAA GATTTTCCAA GGCAATGTAT TTTCATTGGT	660
	ACAAC TAATA AAGTTGATTT CTTAAGAGAT GAACTGGTG GAAGACGTTT TTGGCCAATG	720
15	ACTGTAAATC CAGAGAGAGT TGAAGTGAAC TGGTCTAAAC TAACCAAAGA TGAGATTGAC	780
	CAAATTGGG CAGAAGCTAA ACACTATTAT GAACAAGGAG AAGATTTATT CCTTAACCCT	840
20	GAACTAGAAG AAGAAATGCG TTCAATACAA AGCAAACATA CTGAGGAATC TCCATATACA	900
	GGCATTATTG ATGAATATCT TAACACACCM ATTCCtAGCa ATTGGGATGA CTTAACTATC	960
	TTTGAACGAA GACGATTTTA TCaAGGTGAT GTTGATATGT TACCaACAGG AAATGTaGAT	1020
25	TACGT TAAAA GAAATAAAGT CTGTGCGCTT GAAGTGT TTG	1080
	AAGGGAGATA GTAGAGGATC TATGGAAATT AGAAAGATTT CAAACATCTT AAGACAATTA	1140
	GACAAATTGGT CTGTATATGA TGGTAATAAA AGTGGGAAAA TTCGATT TGG AAAAGATTAT	1200
30	GGTGTACAGA TAGCTTATGT AAGAGATGAA AGTTTAGAGG ATTAAATATA AGAAATATTG	1260
	AATAAATATG CATTTTAGAG TGTGTATCA GATGTTGCAT CATTTTTTGA GTGATGCAAC	1320
	ACGGGAGTGT AAAAGTAAT CGTAGGTGTT GTATCATTTT TGGTGATGCA ACATTGATGC	1380
35	AACAAATGAT ACAACACCTC TTTCTTTTCT AGCTGTAGGG TTCAACCCTG TTGTTTTCCA	1440
	ATGTTGCATC AAATTCAC TAAGTTTAA AAAGTAGTGT TAGGGAGTAA AGGGGTATAG	1500
40	GGGTAACCCT CTAACAGCTA TTTTAAAAG TTTGGCAAGA ATTGATACAA CATCGGAACA	1560
	CAAATATAAA TTTGTATAC AAGGTGAATA AATGAAAGAA TCGACATTAG AAAAATATTT	1620
	AGTGAAAGAG ATAACAAAGC TAAACGGTTT ATGTTTAAAA TGGGTGCGAC CTGGAACAAG	1680
45	AGGTGTGCCA GATAGAATTA TTATTATGCC AGAAGGAAAA ACATATTTTG TAGAAATGAA	1740
	GCAAGAAAAA GGAAAGTTGC ATCCTTTACA AaAATATGTG CATAGACAAT TTGAAAATAG	1800
	AGaTCATAAA GTATATGTGT TATGGAATAA AGAACAAGTA AAaACTTTTA TCAGAwTGGT	1860
50	AGTGAACATT TGGCGATTGA CTTTCAAACC ACATAGCTTt CCAAAG	1906

(2) INFORMATION FOR SEQ ID NO: 680:

(A) LENGTH: 948 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 680:

10 AATTATTTGA ATAAATTAGC AATTAAAGAG TTGATTGTC AGTTTAAGyA TTTGAGTGCA 60
 TTTGAAAAAG ATGTCATGTA TTTAATGTGT GAACAATATA AGCCGAGAGA AATTGCTCAA 120
 TTGATGCATG TAAAGAGAA AGTGATTTAT AATGCCATAC AACGATGTrA AAATAAAATA 180
 15 AAACGTTATT TCAAAATGAT TTGAAAAGCG CCTTAGGACG TGAATTGAAT TATAACGTGT 240
 TACTTACTGA TGGTTTGACA TTTGTTATAA ATTTTATGTA TAGTATACTG GTATTATAAT 300
 GAATAAAGGT GAATTATTGT GAGAAAAATA CCTTTAAATT GTGAAGCTTG TGGCAATAGA 360
 20 AATTATAATG TTCCTAAGCA AGAAGGCTCG GCAACAAGAT TAACCTTAAA GAAATATTGT 420
 CCAAATGTA ACGCGCACAC AATTCATAAA GAATCGAAAT AAATACATTC GAAATAATAC 480
 TTTGATAATA TGTTCAAAGG ATTTGGAGGT TGAGCAGATG GCTAAAAAAG AAAGTTTCTT 540
 25 TAAAGGCGTT AAGTCTGAAA TGGAAAAAAC AAGTTGGCCG ACGAAAGAAG AGCTATTTAA 600
 ATATACTGTA ATTGTAGTTT CTACTGTTAT ATTCTTCTTA GTCTTTTCTT ATGCCTTAGA 660
 30 TTTAGGAATT ACAGCATTGA AAAATTTATT ATTTGGTTAG AGGAGTGAAG ACATGTCTGA 720
 AGAAGTTGGC GCAAAGCtTG GTATGCAGTG CATACATATT CTGGATATGA AAATAAAGTT 780
 AAAAGAATT TAGAAAAAAG AGTAGAATCT ATGaATATGA CTGAACAAAT CTTTAGAGTA 840
 35 GTCATACCGG AAGGAAGAAG GAAACCTCCA GTAAAnAAGnT GGCCAAGCCT AAAACCGCCT 900
 GTTAAAAAAA ACCATTCCCC TGGGnTAnGG TTTTAAGTGG GAATTTAA 948

(2) INFORMATION FOR SEQ ID NO: 681:

40

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 863 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 681:

50 ACAnATAATA ACAAAGCGCT TGCTAGTACC TCTTAAAAAG ATGATGCTAG CAAGCGCTTk 60
 TCTATACTAT ATATTATTTT TCTAAAATTT TAACACCCTC TTGAGTGCCT ACAATAACTT 120
 GATCTGCCAT ATCTAAGAAG TATCCTGTCT CAAACACACC TGTCAGATGA ATTAAATACT 180

55

TATCAGTTAT AAATGCGACA TCTTCGTTTA CACGACGTTT TACTTTTATA TCAGCGTATG 300
 ATTCAATTTT ACCTAATATG TGATACCAGT TAAATTTATC CACCTCTACT GGTAACCTAA 360
 5 ACGTCTCACC TAAGTATTGA ACTATTTTCG TTTTCATCGAC AACCACAACA AAACGCGATG 420
 CCATTTTCATC TATAACTTTT TCTCTGAACA GCGCACCACC GCCACCTTTA ATTATATTTA 480
 AAGATGGATC TACTTCATCA GCACCATCAA TTGCTAAGTC GATATGATCA ACATCATTGA 540
 10 TTTACACATAT TTTAATACCT AATTCTTTTG CTAAAAATGC AATTTTATTA GAAGTGCATA 600
 CACCTGTAAT ATTGTAACCA CGTTCTTTAA TTAGTTGCGC CATTGAGGT AAGAGTAATT 660
 15 CCATTGTAAT TCCTGTACCA ATTCCCAGCG TCATGTCACC ATTGATTkGA CTTAAAACAT 720
 CATTTAATGT CATTAACTTG AGTGCTTTGA CATCTTTCAT GAAGGTAGCC TCCCATATTT 780
 ATGTAATCTA TTCAATTCAT ATTTTACATG ACTCGTATAA ATTAACATAC CCTTATnGCT 840
 20 AACCATTTGT GTTAAACATA TCG 863

(2) INFORMATION FOR SEQ ID NO: 682:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 480 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 682:

TTCATTTTTG TTGTTTCAGCA ATTTTGGTCC GAATTTCAAT TCATCAGATG ATAGCTCTAT 60
 35 TAATTCTGTA GAATATTCTG CcACAACAAT TTCATAAATA TGGCCTTTT CTTCCATTAT 120
 TATTTTCATCa ATTATTTTCAT AATTCAATTG TTGTAATGTT TGTCTTAAAT TTTAGTTTG 180
 GATATTACTT TGTAATATCA ACCTTGGATG TTGACTTAAC TTATCTTGCC CATCTTTTAA 240
 40 AATTTTAGCA ATAAGTGGTC CGCCCATACC ACAAATTGTG aTATTATCGA TTACGTCCTC 300
 AGGTTGAATA AACTTTAAGC CATCCCCTAA ACGTACATCA ATTCTATCTA CTAATTGGTT 360
 TGCAGCTACA TTTTTCACAG CAGCTTGAAA AGGGCCTTGA ATAATTCTC CAGCAATACC 420
 45 GaTTGCGATA AATGGTTTTG AATTGCATAG ATTGGCAAAT AAGCATGATC TGAGCCAATA 480

(2) INFORMATION FOR SEQ ID NO: 683:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 689 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 683:

CTGCAAAAAA TATTGGTATA ATAAGAGGGA ACAGTGTGAA CAAGTTAATA ACTTGTGGAT 60
 5 AACTGGAAAG TTGATAACAA TTTGGAGGAC CAAACGACAT GAAAATCACC ATTTTAGCTG 120
 TAGGGAAACT AAAAGAGAAA TATTGGAAGC AAGCCATAGC AGAATATGAA AAACGTTTAG 180
 GCCCATACAC CAAGATAGAC ATCATAGAAG TTCCAGACGA AAAAGCACCA GAAAATATGA 240
 10 GTGACAAAGA AATTGAGCAA GTAAAAGAAa AAGAAGGCCA ACGAATACTA GCCAAAATCa 300
 AACCACAATC CACAGTCATT ACATTAGAAA TACAAGGAAA GATGCTATCT TCCGAAGrTT 360
 15 gGcCCAAGAA TTGAACCAAC GCATGACCCA AGGGCAAAGC GACTTTGTTT TCGTCATTGG 420
 CGGATCAAAC GGCCTGCACA AGGACgTCTT ACAACGCaTA AcTACGCACT ATCATTcAGC 480
 AAAATGACAT TCCCACATCA AATGATGCGG GTTGTGTtAA TTGAACAAGT GTACAGAGCA 540
 20 TTTAAGATTA TCGGAGGAGA GCGGTATCAT AAGTAAAAct AAAAAATTCT GTATGAGGAG 600
 ATAATAATTT GgaGGGTGTT AAATGGkGGA CaTTAAATCC mCGTTCATTC mATATATAAG 660
 ATATATCACG GTAATTGCGC ATATAACTT 689

(2) INFORMATION FOR SEQ ID NO: 684:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 858 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 684:

TTATTAAATT GGTATGTGTT CATTATACAT ATGacAAATA TGaATGTAAA CCGATAATTT 60
 AGATTTTTTG GAATAACCTG AAAATTCAAG TTaTAGCGTT GCTTATATTT TAAAAGGTGG 120
 40 TGATAATGAG ACTTTTTGAA AAATAAAATT CAAAATACTT ATAGCATAAT CAATATGCAC 180
 ATTAAATAAA TGTACTCTTT TAATGCGTTG ATAActGTAT TTGTAAATTTA GAGAAGGGGT 240
 GTTCACTATG CTGATGTGTT AAAAAATAAA ATAAAAAGGA CACCTCGATG CTATAAATAT 300
 45 TAGCATCGAG ACGCCCAgTT AATGTCTATT AAATTGAATA TAGTCTCGGA CATGAATCAA 360
 TGCCCTAGGC CCTGCAATGT TATATTGACA GTAGTTGACT GAATGAAAAT GACTTTGTAG 420
 50 CTAGCTTTTT TCAATCCTTG TCGGTGCAAC ACATAGAGAA ATTGGATTCC TAATTTCTAC 480
 AAACAATACA AGTTGCGGAA TAAGTCCCAA TATAGAAGGT GACAGTAAGC CAACTTACAA 540
 TAATGTGCAA GTTGGTCGGG CCTCAATACA GAGATTTTCG AAAAGAAATT CTACATATTA 600

GCCACCTTGT TTATTTAAAT CGATAACACG GTTTCGATT GTATTGATAA ATTCAAAGTC 720
 ATATGAAGTA AAGATAATAG AACCTTTGAA TGATTAAAGT CCATCATTAA CAGCAGTAAT 780
 5 ACTTTCTAAG TCTAAGGGTT GTGGGTCATC AAGTAAAGA CGTTGCCCTG ATAACATCAT 840
 TTACTAGCAd CACGACTT 858

(2) INFORMATION FOR SEQ ID NO: 685:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1747 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 685:

20 TTAGTATTTT CAGGnTGGAA ATTGATGTTT GAAATTCAC T GAAATGGGCC AGAACCTgAA 60
 ATGaaACTTT ACCAACAATT TGGtCTTCAT CAATGAGGCC AAACGCACGG CTATCTTTAC 120
 TTACTTCACG ATTATCTCCA AGCACTAAAT ATTTACCTTT TGGAAATGACA TTTGATTAG 180
 25 GATTTCGATT CGGTAAATCT TTAAGTTGGA AAGTCCAGT AATGTAATCA CCTTGTTTAT 240
 GTTTTAAATT GTAGTTTAAA TATGGTTCAT CTGTTTTTTT ACCATTGACA TATAATGTAT 300
 CATTTTTGTA TTCTACTTTA TCACCAGGAA CACCGATGAC ACGTTTAACA TAGTCATCAT 360
 30 TTTTGTTCG ATGGAAGACA ACTACATTAC CTTTTTCCAA ACCACCTGTT TTATATCCAA 420
 CAATGTTTAC AGCTACTCGC TCGCCATCTT TCAAAGTTGG ATCCATTGAT TCACCTTTAA 480
 35 TTGTATATGG CGTAACAATA AATTTACCTA CTATAAATAA AATGACAAAA GCGACTGCAA 540
 TTGAAATAAT CCATTCCAAT ATTTCTTTTT TCAATTTTGA CACCTCTTTT TAAGATTGGA 600
 ACTGAACAGT CCATTTTGAA AAAGGATAGT ATCGTAAACt AACATTACCA ATAATatCCT 660
 40 TTTTATCGAT TAAACCAAAT TGTCTTGAAT CGTGCTTGtK ATTATCTTGA TCATTTAGCA 720
 CAACAAAATT GTTTGGCGGA ATAATATCAC CATCTAATTC TTTAAAATTG CGCAAACATA 780
 AATCTTTAAT TTTTCTGTTT TTGGCATAAG ATGCGTCAAC CGGTCGGTCA TCACGGTATA 840
 45 ATTGTCCCTG ACGAAACGCC ATTGATTGAC CAGGTTTGGC AATAATTGCA CTAGTATATA 900
 TCTCGTTACC ACGCCTATAT GTAATGATAT CACCATTATT CAATTGATTA AATGTAACCT 960
 50 TAATTTTATT TACAATAACA CGATCCCCTT TGTTAAGGGT TGGTGACATA TCATTATTCG 1020
 GAATGACATG ACCAACTATT ACAAAGTTT GTACGAACAG TACAATGATA ATAGCAAGTA 1080
 TCAATGAAAT CAAATATTTT ACAACTTTTT TCACGATGTC ACTCCTTTTT CGATCCCAT 1140

55

ATCCCTCTTA AAGGATGCGT TAAAATTGTA GTAATTTCTT TACCTAAATA ACCTAAAATA 1260
 ATTGTTGAAA CTAACCTTGA TGATGCCAAA ACAATGAAAT AATATTTAGG TCTAATATGA 1320
 5 GATAGACTCG CTACAAAATT TATTAATGTA TTTGGCGTAA AAGGAAAACA AAGTAAAATA 1380
 AACAAATGGGA TTAATCCTTG GCGATCAATA AACTAATCA AGCGTTGAAC AGCAGTACGT 1440
 TGTTTAATTC GCTGCATCCT CTCAGTGTTC ACCAATCGTT TACAGATCAA ATAGACTGTA 1500
 10 AATGTTCCAG AAATTAATCC AAGCCAACTA ATCAATATAC CTAAAATAGG TCCATAAGCT 1560
 TGAATGTTAA TTAAATATA GAGTGCTAAA GGAAATACTG GAATTATAGC TCTAATATAT 1620
 AACAAATATA ATCCAGGTAA ATAACCAAAC TGTCGAAATA TCTCAAACCA TTCTTCTACT 1680
 15 TGATGAAACG ACAAATCATC AATCCCTTTC TTTGGTTGAA GATAATTATT CTTACATTAT 1740
 AAAGTTA 1747

20 (2) INFORMATION FOR SEQ ID NO: 686:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 645 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 686:

CGTAAAGATT ACTATTTAGG AGGGTGACTA TGAAAAAGAA ATTAGGTATG TTAATTCTTG 60
 TACCAGCCGT AACTTTATCA TTAGCCGCAT GTGGGAATGA TGATGGAAAA GATAaAGATG 120
 35 GCAAGGTAAC AATTWAaGc CaGTTaTCCm TTGCAATcAT TTgCAGAGCA AATTGGTGGa 180
 AAACACGTGA AGGTATCATC AATCTATCCA GCAGGGACAG ATTTACATAG CTATGAACCA 240
 ACACAAAAAG ATATATTAAG TGCAAGCAAG TCAGACTTGT TTATGTATAC AGGGGATAAT 300
 40 TTAGATCCGG TTGCTAAGAA AGTTGCATCT ACTATCAAAG ATAAAGATAA AAAACTGTCT 360
 TTAGAGGATA AATTAGATAA AGCAAAGCTT TTAAGTATC AACACGAGCA TGGTGAAGAG 420
 CATGAACATG AGGGACATGA TCATGAGAAA GAAGAACATC ATCATCATCA TGGTGGATAT 480
 45 GATCCACACG TATGTTAGA TCCTAAAATT AACCAAACCT TCGCTAAAGA AATTAAAGAT 540
 GAATTAGTGA AAAAAGATCC AaAACATAAA GATGACTATG AGaAAACTA CnaAAATTAA 600
 50 ACGACGATCT TAAGAAAATT GATAACGATA TGAAGCAAGT TACAA 645

(2) INFORMATION FOR SEQ ID NO: 687:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 956 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 687:

	TTTGTTACTG CTTCTAAAAT AATATCCTTT AATTGTTTAA CATGTTGGAT TGTCATATGA	60
10	GGTGATGGTA CATTAAAAGG ATTTAATTCA TCTATTTGTG CATATTGATT TATGACATCT	120
	TGATGCATTG AAATAGGGTT GATATCATTG GTTACTACTT TATTAGATTG GTCTTGTGAC	180
	ATACTAATGG TGCCACCAGT ATGAATAACA AGTAGATGTT TCATATATTT CCTCCTATAT	240
15	TTAATTTACC TAATTATGAT AAAATATTAT TCATAAAACG ACAAGGAAGG GAAATGACGC	300
	ATGAAAGCCA TTAATATTGC ATTAGATGGT CCAGCTGCTG CCGGAAAAAG TACAATTGCG	360
	AAACGTGTAG CCAGCGAACT ATCAATGATT TATGTCGATA CAGGAGCAAT GTATCGTGCA	420
20	TTAACATACA AATATTTAAA ATTAAACAAA ACTGAGGACT TTGCAAACT AGTTGACCAA	480
	ACAACATTAG ATTTAACTTA TAAAGCAGAT AAAGGTCAAT GTGTCATTTT AGATAACGAA	540
	GATGTAACAG ACTTTTTAAG AAATAATGAT GTGACGCAAC ATGTTTCATA CGTTGCATCT	600
25	AAAGAGCCAG TACGTTTATT CGCCGTTAAA AAACAAAAG AGTTAGCTGC AGAAAAAGGT	660
	ATCGTAATGG ATGGTCGCGA TATCGGAACT GTAGTGCTAC CAGATGCAGA TTTAAAAGTA	720
30	TATATGATTG CATCAGTTGA AGAGCGAGCA GAAAGAAGAT ATAAAGATAA TCAATTAAGA	780
	GGTATCGAAT CAAATTTTGA AGATTTAAAA CGTGATATTG AAGCTCGTGA TCAATATGAC	840
	ATGAACCGTG AAATATCACC ATTAAGAAAA GCAGATGATG CAGTGACATT AGATACGACm	900
35	GGCmAGTCGA TTGAAGAAGT TACTGACGAn ATTTTAGCGA TGGTGAGTnC AATTnA	956

(2) INFORMATION FOR SEQ ID NO: 688:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 1166 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 688:

	AAAAGGGAGG AAAGAGAACA GTTAAATATG AATACAAAAT AATTTTGTTT TCGGACAGCA	60
50	GGGGTATTAG ACGCGATTGA CAATGTCGTG TTAATTAAAC GTAATGTTTA TTTAAGCGAT	120
	GAATATTAGG TGAAAAGTTT TTGAATTTGA ATGTAATTGA GGTTTATTGA TTAGACATTT	180
	TATTGAATTG CGTGTTATTA TATAAATGTA AAAATAAGAC GACATGCGCG AACATGTCGT	240

55

5 TAATGCGGAA TGGTTTTTTT ATTTCCGCT AATTGAAATA AAAATGACGT TTTAATATAT 360
 TATGGGCTAG GTGGTTTGTA AGAAAGGGTT AGTTATTAAT GTTTTATGAA TTAAGGAAAT 420
 10 TTGAGTTTAA GGTTTAATCA ATTGTGATTT TGTTGATGAA GCGTTTAGTT AGAGTATTTT 480
 CGCCACCACT AGTTACTCCT TCTCCCACTT TACCCGAGAC TGGAGAAGAg CTATCTGAAG 540
 AATAAATAGA TACTTTTTTG CCATTTTGTA GTAAACCAAG ACCTTTTAAC TkCTCGGTTA 600
 GAGAATTCCA TGTATTTTTA GCATCTAGCT TTTTGTTAAA GTCayCGTAG ACATTTTCCT 660
 TAGTTAAATC AATTGTGTTT AATCCTTTAA AGTCTATAGA TTGTGTTAGA TGGCCTCCAT 720
 15 CATCATTTCC AGGAGCAGAA ACGCTGCTAG AATATCCATT GCTTAATAAA TAAGTAACGT 780
 TGATTGTTTC GTACTCGTTA CTTAAAATAA TATCAGAATC ATGTAAGaAT CTTTAACTTT 840
 TTTCCATAAT TGACCATCTG TCATTTTTTC TTCTGCTTTA GCCGTTTAA CAACTTTATT 900
 20 TGTATCTAAT CCTAAGTATG AAGAATGTAA GCCTGTTTCT AATGTTGTTA ATACTAAAGC 960
 ACTTGCTACT aATGTTTTAC CTAAAAAtT TGTATTCATT TTTATTGCTC CTTtTTTTAT 1020
 ATTGTAAACG TTTACAATGA AAATATAATA ATAATTTTTT AAAAGAACAA TTAATAAAT 1080
 25 ATCAAAAATG TATTAATAT CTATTAATA AAAAATAGAA TAATTTTTTA ACATAGTTTT 1140
 GTTGTTTTGA ATTATAAAAA CTAAAG 1166

(2) INFORMATION FOR SEQ ID NO: 689:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 689:

40 AGCACTGAAG GATGGCTAGT TGTCATAGCA TATGCGGTCA TTGGTATAGT TATAACGAGC 60
 GTTTCCTTATA TGTTGTCTAT TAAAATTTTT AACAAACAAG AACTAKAATT GTCGATTTTG 120
 GTGATGATAG TATTTGAAAT AATATTTAGA GCAGGTGATA AATCTTTACG ATTGTCATCT 180
 45 GTTCTTTTkg GTGTGGAATG AAATGTGGGG GATAAGTATA GGTGACATAT CTATATTGAT 240
 TTATTTGTTT TGAGGTGGTT ATGTTGTGTG GGAATTATTT CCTTTTAGAT AGCGGGGATT 300
 AGAGGATATA TGTTATTTAT AAGTATCATT TGATGATTGT ATAGGCTAAC GATTTCCCTCG 360
 50 GAAATATTTA AAAACCTCGA TCATGTAGCA TAACTGAAGT TTGTCACAAA AGTATAATGT 420
 GAAGTTCGAC ACTTTTGGaT TCAGTTCAAA TACTTTGACC GAGGTAAATA CTATTTATTC 480

5 TGATACTGAG ATAATCATTa CATGGTCGTG ACCTTTAAAT AAAAGGCTGA CAATATAAGA 600
 CATAACGAGT ATACCTAGTG AATATGAAAT ATACTTCGCG TTTGTCAGTT CATTATGGAA 660
 10 ATAAGGCGTG ATTAACCATA ATCCAATATA GAATATTAAA AACTGATAT ACATCATATT 720
 AATTTCAAAC AAGTCATTa GTTTATTGTT ATTACTAAAA ACAATTGCAG CATTAAATCAC 780
 ACCTAAAGCG ATATTGATTA ATAGATGCGT ATACGATAAA CGGAAACCGA TAGATGTTAA 840
 TTTATGATTA ATATAATTTT CAGTAATGAT CCAATATACA CCGAAAAGAC TAATTAAAAT 900
 CATAAATTGG AATATATAAA TGTAATAAA ATGATCAATG CTAAATGATG ACGAAGCTAA 960
 15 ACCAACCAGT ACCTCGCCa AGWTATAATT GTTAGTAACG AAAAACGTCT ACTAAATGCA 1020
 TCATATTAAc AGGTnTAATA CAAGTATTTc TGAAATGGAA TAAGnCTGTC GCTGCATGAT 1080
 ACG 1083

20 (2) INFORMATION FOR SEQ ID NO: 690:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 627 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 690:

30 TTAATCATCT GGATGTATTT AGTATTnAGA ATAATAAnAA AACGATCATG TTGTATtTGA 60
 GTAGCGATTG GTTTGCGGAA TTAGGCTTTA CTTTCTTTAA TTACCACTAT ACAGCAAAGT 120
 35 TGATTAAATC ATCCTATAAT TTGAAATGTC TACTATTAAA ATTGACATAT CGATACCTTG 180
 ATAATCAGCC TCTTAATGAC GCTGATayTA GAAAATTACA GGATATTATT AAAATCATTG 240
 CAAAAGAAGC AAGTATGGAT AAAAAGATTG CACAAAATCA ATATCGATAT GCGTATTATG 300
 40 GTGATTTGCG TGATGAGCTC GAATATATTT ATCAAAATGT AAATCAACGA TTGACATTAA 360
 AAAGTGTGCG TGATAAATTA TTTGTCTCAA AGTCAAATTT GTCATCACAA TTCCACTTAC 420
 TTATGGGCAT GGGTTTTAAA AAATATATTG ATACTTTGAA AATnGGTAAA TCGATTGAAA 480
 45 TTCTACTTAC TACTGATAGT ACTATTAGCA ACATAAGTGA nCATTTAGGT TTTAGTAGTA 540
 GCTCCACTTA CTCTAAAATG TTTAAAAGTT ATATGGATAT CACACCGAAT GAATATCGTA 600
 50 ATTTATCAAA ATATAATAAn TGTTTAC 627

(2) INFORMATION FOR SEQ ID NO: 691:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 641 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 691:

	TTGAACGTAA TGCTAGCAAA TGACTTTGTG CCATAAAATA TTCTTCCCAT TTGATTCTTT	60
10	CCAAGATGTT CACCTTCCAT ACTTAAAATT TAGTAACATT TTCTAATAAT ATAAGATTAA	120
	TCACAAAAAA TAAAATTTGC AATTAAAATA ATCCATTATG TCGTGAAATA AGATTTTCAGT	180
	TTATCAAAAG TTTTACTTCC AAAACCTTTT ACTTTTTTCA AATCGTCAAT TTCTTGAAAT	240
15	GCACCTTGTT GGTGCGATA TTCAACAATT GCATTAGCTT TAGCTTGCCC AACTCCAGGA	300
	ACAGACATCA ATTCTGATAC AGATGCCGTA TTAAATTTA CTTTAGTATT ATTTGTGTTT	360
	CCATTTTTTT CGTGCACT GTTACTTCA ATTTGTGGTT CAACATTCTT TTGTCCTTTA	420
20	TGAGGTATGA AAATCATT TTGATCTGTT AATTTTTCAG ACAAATTAAT TTGACTTACA	480
	TCTGCATCCT CCAATAATTG TGCTTTATCA AGTAAATCAA CTACTCTATC CTTAGATGTC	540
	ATTTTATAAA CATTAGGATG TTAAACAGCA CCTTTTACAT CGACATATAC AGGACCCTTA	600
25	TTTTTGAAT TATCTCCATC TTGACCTGG ACATCTTCTA C	641

(2) INFORMATION FOR SEQ ID NO: 692:

30

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 631 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 692:

	TATTAAAGnA CTTGATGAAC CAAATCATAA AAAGCnATAT ATGTTATTTG CAGCTGGCAT	60
40	TGTGTTnGCA ACTATTTTAC TTATTTTCGGC ACATTTTATAC AGCAGAAAGA GAGGTAACCA	120
	AGTTTGAGAA TCATAAAGTA TTAAACCATT TTAGTGATAA GCGTCGTTAT CTTAACCAGC	180
	TGTCAATCTT CCAGTTCTCA AGAATCAACT AAATCCGGCG AATTCAGAAT CGTACCAACA	240
45	ACTGTTGCAT TGACAATGAC ATTGGACAAA TTGGATTTAC CAATTGTCGG CAAACCCACG	300
	TCATATAAGA CATTGCCTAA TCGTTATAAA GATGTACCGG AAATTGGTCA ACCAATGGAG	360
50	CCGAATGTTG AAGCTGTTAA AAAGTTAAAA CCAACACATG TTTTGAGTGT GTCAACGATT	420
	AAAGATGAAA TGCAACCATT TTACAAACAA TTAAATATGA AAGGCTACTT TTATGATTTT	480
	GATAGTTTAA AAGGGATGCA AAAGTCGATT ACACAATTAG GTGatCAATT TAATCGTAA	540

55

GCAGCTAAAC AAAAGAAACA TCCCAAAGTA T

631

(2) INFORMATION FOR SEQ ID NO: 693:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1111 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 693:

15	AATTTAACTA TGTTTTCCAC ATTGTTTCAT GTCACGAAAA GGACAACGCG CGACTATAAG	60
	TATCAACTAT TTCCACAAGT TTTATTGGTG TTTTATTAT TCATCGATAC GCTTCATTTT	120
	CATCTCTCCA ACACAAAAAA GAAGCTAAGC AACTTATGTT GCCTAACTCC TCTATACTAT	180
20	CCATATTTTA CTATTATCCA TATTTTCATTG AATTATCTAA TGTGGCCTTC TATTTTTTCA	240
	ATATTTCTAC CGTCAATGAC GTCACCTCATG CGATTTGTTT GTAATTTTTT ATTAAGTTCA	300
	AACGTATAAT AGCCGCCATC TTTCAATTATC ACTTTTATCT TACTATCTTT AGGAACTTT	360
25	TTATACAGAT CAAAATTTTG AATTAAATAC TGTCTCAATT TAAAGTCGAG TTCTTTAAGT	420
	GAAATCTCTT CTTTATAAAT GTAGTGTACT CTACCGTACG TAGCAATACC GTCACCTTCA	480
	TCTCTCTTGA TTTGAAATCT TGGTGCCTTT ATATAATCAT AATAAGCGTC TTGATTTTTT	540
30	TTAGTGACAC CACCATATGA AAACACTGTG CCATTACGGT TTTCCGCTTC TTTAACAACA	600
	AATATGTCTA ATCCCGGATT TTTACGTGCT TTAAATCTTT CAATATCTTT ACCAAATATC	660
35	TGTACTCTTG TGAATTTTCT ATTTTATCA AAGATAAGGT AATGCTTGCC ACCTTTGCTA	720
	TAACGATAAC CAGTAACATT TTAAAGTTCC TTAAGTGGC CACTATAGTA ATCTCTTAAG	780
	TCAAAGATAT CTTTTGTCAC ATTTTCATAT TTTGCTTTAT GTTCACTCGC ATTTACAGTT	840
40	TGATGCAATG ACGTTATTGT TCCTGTTGCT AAAATACCTA ATGCTAACT TGCTTTGCA	900
	ATTGCTGTCA TTTTCATAGT TGTATGCTCC ATTCGTAATT ATTAGATTG TTCGCTTACG	960
	TCTATTGAAT CATACAGCTT TATTATAGTT AGCGTATTTG ACCTTTCACA TTAAACCATG	1020
45	TTTAATAATC ATTGAATCAT TATTAAGTAA ATTAAGGATC TATAATGTTT GTTAAATAAA	1080
	CTGANCCCGT TGTGCTTCAC ACCCGnTnGA T	1111

(2) INFORMATION FOR SEQ ID NO: 694:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 426 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 694:

5 TTATGGATGG ATTAAGAGGT CGTGTGAAA AAATCAACGA TAACTCTGTT ATTGTTGACT 60
 TAACAATTAT GGAAAATTTT AATGACCTTG ATTTACCGGA AAAAAGTGT ATCAATCATA 120
 AACGATATAA GATTGTTGAA TAAGAAGGTA AGTTATAATG AATAAAATCT CGAAGGCTTT 180
 10 AACTTGTTTT ATTATAAGTT TCATTATATT TCATCTCATA TTATTTATTA TGTGGGGCGA 240
 ACACCAAGAA TACTGGTATT TATATACAGG TATAATGCTA ATTGCTGGTA TCAGTTATGT 300
 ATTTTATCAA AGAGATATTG AATCTAAGCG GTTGCTTACA TCAATTGGTG TTGGTATTAT 360
 15 TACGGCAATT ATTTTAATTA TGCTTCAACT TTTATKCTCA CTTATAACTT CTAATTTAAG 420
 TTATAG 426

20 (2) INFORMATION FOR SEQ ID NO: 695:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 737 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 695:

30 AAAAGGGGnG TAAGGTTTAG CTCAAGTACG AGAAGTCTTT GGTGATGAAG CAATTGATGA 60
 AAATGGTGAG ATGAATCGTC GTTATATGGG TGATCTAGTG TTTAATCATC CAGAAAAACG 120
 CTTAGAATTA AATGCTATCA TACATCCTAT CGTGCGAGAT ATTATGGAAG AAGAAAAGCA 180
 35 AGAATATTTA AAACAAGGAT ATAATGTAAT CATGGATATT CCATTATTAT TTGAAAATGA 240
 ATTGAAAAT GCAGTAGACG AaGTGTGGGT TGTATACACT TCTGAAAGTA TACAAATGGA 300
 40 TCGTTTAATG CAACGTAATA ATTTGTCATT AGAAGATGCG AAAGCACGTG TCTATAGCCA 360
 AATTTCTATT GATAAAAAAA GCCGAATGGC CGATCATGTT ATCGATAATT TAGGGGATAA 420
 ACTTGAATTA AAACAAAACC TTGAGAGATT GTTAGAAGAA GAAGGTTATA TTGAAAaGCC 480
 45 GAATTACGGA GAAGAAGATT AATATTACAC TATAAATAAG TCATTACTTT ACGTACGCGT 540
 TGATGTATGT AAGTAATGAC TATTTTTTAT AAAAAAGATA AATAAATCAA CGGAAAACGC 600
 TTTCAAATTT CATATAATAT GCTATACTAA TTCCATAAAG TATAACACAT AAAGATCAAG 660
 50 GGGTGCTTTT AATGTCAACG AATATTGCAA TTAATGGTAT GGGTAGAATT GGAAGAATGG 720
 TATTACGTAT TGCATTA 737

55

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 432 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 696:

```

10 AACGATAAGA ACAAAGATA TATTACAAAG CGTTTATTTA AAACGTTATT TATTACGCGC      60
   GATGATGGCA GGATTTATTA TCGGGATTAT TACGGTCTTC GTATTATCAG TTAAGCAAC      120
   ACACGAACCA GATTTACCGC CAGgCATTGT GAATATGGCC AGTGCCATTA CATTGAGCTT      180
15 TCGGTTAGTA CTCATTTTAT TTACAAACTC CGAACTACTA ACCAGTAACT TCATGTACTT      240
   TACTGTAGGC CTGTATTmTA AAGTAATTAA ACCAACTAGA GTATTGaAAA TATTTTTATT      300
20 ATGCTTTGCA GGAAATATTT TAGGTGCTGC TATTTyATTT AGTTTCATGC GTTTTTCAAA      360
   TGTAATGACG CCAGATAwGt TAAAyCAGTT ATCAGCAGTT ATAGAGCATA AAACGTTGTC      420
   TACTGGTTTT GT                                          432

```

(2) INFORMATION FOR SEQ ID NO: 697:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 782 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 697:

```

35 CTTTTATACG AAAGTTAAaA TCAAATATAA CATTAAATGTT TGATGGGGAT TTTGCGGGTA      60
   GTGAmGCAAC ACTTAAAACA GGTCAAAATT TGTTACAGCA AGGGCTAAAT GTATTTGTTA      120
40 TACAATTGCC ATCAGGCATG GATCCGGATG AATACATTGG TAAGTATGGC AACGATGCAT      180
   TTA CTGCTTT TGTAaaaaAT GACAAAAAGT CATTTCACA TTATAAAGTG AGTATATTAA      240
   AAGATGAAAT TGCACATAAT GACCTTTCAT ATGAACGTTA TTTGAAAGAA CTAAGTCATG      300
45 ATATTTGCT TATGAAATCA TCGATTTTGC AACAAAAGGC TTTAAATGAT GTTGCACCAT      360
   TTTCAATGT TAGTCCTGAG CAATTAGCTA ACGAAATACA ATTCAATCAA GCACCAGCCA      420
   ATTATTATCC AGAAGATGAG TATGGCGGTT ACATTGAACC TGAGCCAATT GGTATGGCAC      480
50 AATTTGACAA TTTGAGCCGT CAAGAAAAAG CGGACnAGCA TTTTAAaAC ATTTAATGAG      540
   AGATAAAGAT ACATTTTTTAA ATTATTATGA AAGTGTGAT AAGGATAACT TCACAAATCA      600

```

TATCAGTGAT GCTGTGCAGT ATGTTAATTC AAATGAGTTG AGAGAAACAC TAATTAGCTT 720
 AGAACAAATAT AATTTGAATG ACGAACCATA TGAAATGAA ATTGATGATT ATGTCAATGT 780
 5 TA 782

(2) INFORMATION FOR SEQ ID NO: 698:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 714 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 698:

AAATCATTGG CAAAATATAC GAATTGCATT TAAATCGTCA ATTTACATAT ATTTTTCGCT 60
 20 AGTAATCAAT CGTTATCATT GTTATTTATC GTTACATTAT TTCGAGTATC AGTATGTATT 120
 TCGGGCTTCG TTTGATAACG ACATTTCTTT GTGACATCGC TTCATCAGTG TAACAACAAA 180
 TACAATGATT TCGTGATGTT AGTTACCCAT TTTATGTGTT GCATAAAATA TGTGTGTATA 240
 25 AAACATTTTA AATCATTTTA TATAAACAAT CTATATATTT TTGGCATTTC CAAAATATCA 300
 CTTGTTATAT TAAAAACCGA CAAGACATTT TATCTTATCG GTTGAAATTT GTTATTGTTA 360
 TTTGTAATGT TTTTAGGTTT CTTTTTAATA TAATATATTT CAGTGAAAAT ACATGATTGA 420
 30 TTGTGATTC ACTGAAACAT GGTAAATTGC GTTGTGTATG AATAACTTTA GCATAAATAT 480
 AGGAAGTTAT TTTGTACATC GCCATATATA GAAACGAAAT TATAATGACA GCTAGTACGT 540
 AACTTGTTAA AAATATATGA TGGTTATTAA TACCTATCAT ATTTAGTAAC GTATATACAA 600
 35 TGTTACTAGA AATTAATGTG TGAATCAGTG CTA CTGTTAT TGGTATGCG AACAGAAAG 660
 TCATTTGATT TCGTGTATC TTTGCTATTC TTCCATTATC TAAACCAAGT TTTT 714

(2) INFORMATION FOR SEQ ID NO: 699:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 361 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 699:

ACTGAGAGCA ATAATTTAGT AACTTCTACT CAAGGAATTA TTAAAGAAGC ATTGCATAAA 60
 TTGGGATTTG nntTTTAATTA AAGAACCTTT AAGAATGTTA CAAGTGCGTA TCCCTGTACG 120

TGGACCAACA AAAGGGGGCG TGC GTTTCCA CCCAGATGTT GATGAAGAAG AAGTAAAAGC 240
 ATTATCAATG TGGATGACTT TGAAATGTGG CATTGTGAAA CTTACCATAC GGTGGTGGTT 300
 5 AAGGGTGGTG ATCGTTTGTG GATCCACGTC AAAGGAGCAT TCCATGAAGT TGGAACGTTT 360
 A 361

(2) INFORMATION FOR SEQ ID NO: 700:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 943 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 700:

20 GTCCTTAATT GGTATCCACC TTTAGCACCG CGTACACTTC GAATTAACCC CGCATTTCCT 60
 AAAGGACCTA CAAGCTGTTC TAAATATAAA TCACTCAAAT TATTTTCTTC AGCAATTGAC 120
 TTTAATGATA TACATCCTTG CCCCTCTTTT TTAGCAAGAG AAATCATCAA TGTAAGTCCA 180
 25 TATCTCCCTT TAGTAGAAAT TTTCATTGTA TAACCTCACT TAATTCGAAT ATTGATATTC 240
 CCATTTTAGC ATTTTTTGAG TTAAGATAGT ATAAGAAAGG TGTGACAAAT GTGAGTACAG 300
 AACCATTAGC ATCGAGAATG CGCCCAAAA ATATAGATGA AATCATTTC CAACAACATT 360
 30 TAGTTGGACC AAGAGGCATT ATCAGAAGAA TGGTTGATAC AAAAAAATTA ACTTCAATGA 420
 TTTTTTATGG TCCACCTGGT ATAGGCAAAA CAAGTATTGC CAAAGCAATT TCGGGCAGTA 480
 CGCAATATAA ATTCAGACAA TTGAATGCTG TAACTAACAC TAAAAAAGAT ATGCAACTTG 540
 35 TTGTTGAAGA AGCTAAAATG TCTGGTCAAG TTATCTTGTT ATTAGATGAA ATACATCGAC 600
 TAGATAAAGC TAAACAAGAC TTTTATTAC CTCATTTAGA AAATGGCAA ATCGTCTTGA 660
 TCGGTGCTAC AACTTCAAAT CTTATCATG CTATCAATCC AGCGATTCGT TCAAGAGCGC 720
 40 AAATTTTCGA GTTATATCCT TTAAATGACG AAGATGTGCG CCAAGCGTTA ACTCGTGCAA 780
 TAGAAGATGA TGAGAATGGT TTGAAACAT ATCAACCCAA AATTGATGAA GATGCCATGA 840
 45 CCTACTTTTC TACACAAAGC CAAGGTGATG TTCGTAGTGC GTTAAATGCA TTGGAATTAG 900
 CTGTATTAAG CGCAGATAAT GACAAAGACG GTTATCGACA TGT 943

(2) INFORMATION FOR SEQ ID NO: 701:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 445 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 701:

5 CATTAAAGTGA AGTTGTTGAT ACACCCATGC ATCAAGTCAA TTGTTCTGTT GATTTAGATA 60
 CAGAAAGCTT ATTAGGCTTT AAAACAATTA AAACAAATGC GGAAGGTCAA CAAGAAATTG 120
 TCTTTGTAGA TGGTCCAGTT ATTAAAGCTA TGAAAGAGGG GCATATTTTA TATATTGATG 180
 10 AAATAAATAT GGCTAAACCT GAAACATTGC CTGTATTAAA TGGGGTCTTA GATTATCGTC 240
 GTCAAATTAC GAATCCATAC ACTGGTGAAG TAATCAAAGC TGTACCAGGA TTTAACGTTA 300
 TAGCAGCGAT AAATGAAGGT TATGTTGGtA CTTTGCCAAT GAATGAAGCA CTAAAAAaT 360
 15 CGCyTTGTTG TTaTTCACGT kGATTATaTt GATGGGGaCa TTTAAAAAAT GTGAnTAAGG 420
 AGCAAGGTTT ATTACAAGAT GTTAA 445

(2) INFORMATION FOR SEQ ID NO: 702:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 752 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 702:

30 TGCAAGATAA ACAAATTTGT CATTGTTTTG GTTGTAaaaa AGGTGGCAAT GTTTTTCAAT 60
 TTA CTCAAGA AATTAAAGAC ATATCATTTG TTGAAGCGGT TAAAGAATTA GGTGATAGas 120
 TTAATGTTGC TGTAGATATT GAGGCAACAC AATCTAACTC AAATGTTCAA ATTGCTTCKG 180
 35 AyGATTTACA AATGATTGAA ATGCATGAGT TAATACAAGA ATTTTATTAT TACGCTTTAA 240
 CAAAGACAGT CGAAGGCGAA CAAGCATTAA CGTACTTACA AGAACGTGGT TTTACAGATG 300
 CGCTTATTAA AGAGCGAGGC ATTGGCTTTG CACCCGATAG CTCACATTTT TGTCATGATT 360
 40 TTCTTCAAAA AAAGGGTTAC GATATTGAAT TAGCATATGA AGCCGGATTA TTATCACGTA 420
 ACGAAGAAAA TTTCAGTTAT TACGATAGAT TTCGAAATCG TATTATGTTT CCTTTGAAAA 480
 ATGCGCAAGG AAGAATTGTT GGATATTCAG GTCGAACATA TACCGGTCAA GAACCAAAT 540
 45 ACTTAAATAG TCCTGAAACA CCTATCTTTC AAAAAAGAAA GTTGTTTATAC AACTTAGATA 600
 AAGCGCGTAA ATCAATTAGA AAATTAGATG AAATCGTaTT ACTAGAAGGT TTTATGGATG 660
 TTATAAAATC TGATACTGCT GGCTTGAAAA ACGTTGTTGC AACAAATGGGT ACACAGTTGT 720
 50 CAGATGAACA TATTACTTTT ATACGAAAGT TA 752

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 703:

nCGTTTTTGT nAAACnGCTT ATATGTATAA CTTTTCCCAA TTTATCATAA GTTTTGATAG 60
 AAGGAGTTGG ATCGCTAAAG AAAATATCTC TAAAGATATC ATCTTCAATG ATAGGAATAT 120
 TATGTCTTTC GCTATAAGTA ATTATATTTT TCTTTTGCTC ATTCGTTAAA GAACGACCTG 180
 TCGGGTTATT AAACCTAGGT TCTATATAAA TCGCTTTATT TTTAAAATTA ATAAATCTAT 240
 CAATGATGGT ATCAATTTCA TTAATTTGAT TATAAGGAAC ATCAATATGT CTAAAATTCA 300
 ATTGCTCAAA AACATTTGTA GAGTGAATAT ATGATGGTGT ATTCGAAATT ATTATGGCAT 360
 CTTGACCTAA AAACCCAATA GATAAAAGTT GAATGGCATG TAAAGCGCCT GAAGTGATCA 420
 TTACATTTTC TCTACCTACA TTTATACCTT GCTTTGACAT TCGTTCAACG ATAATATCTC 480
 TTAACCTTGAT ATAACCATAG CCATTATTAT AACCAAAGA TAAGTCTTCA ATATGACTGG 540
 CTGTATTAGA CATGGCTTTT TTCAATTGAA TATGTGGCAT TAACGATATA CCCAATTCAC 600
 CTTTACTTAT ATGTATATAC GAATCATCTG TCTCAATTTT ATTAATTAAT TGCACCGTAT 660
 ACTGACTTCT TTGTTGAGAG GACCATAACA TCATTTCAGA CCACTTATTT GTAATATGTG 720
 CTTCAATCAA ATAGTCATTA ACATATGKTC CACTACCTAC TTAGTATAG ATAAATCCTT 780
 CAGCTTCTAA TAACTCAATA CTTTAAATAA TCGTTACTCT ATTTACGTTG 830

(2) INFORMATION FOR SEQ ID NO: 704:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 659 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 704:

AGTGGTGTG GAAAAGCTAT TATGAAATTA TTACGTGAAC AACAAGTTTA ATAAAAAAG 60
 AGGGGTCAA TATGAAAGGA TTAATTATTA TTGGCAGTGC ACAAGTGAAT TCACATACAA 120
 GTGCACTAGC AAGATACTTA ACTGAGCATT TTAAACACA TGATATTGAA GCGGAAATAT 180
 TCGATTTAGC AGAAAAACCG TTAAATCAAT TAGATTTTTC AGGAACAACA CCGTCTATTG 240

TATTAGGAAC GCCAACTAT CATGGTTCAT ATTCTGGAAT ATTGAAAAAT GCATTAGATC 360
 ATCTAAATAT GGATTATTTT AAAATGAAAC CTGTAGGCTT AATAGGAAAT AGTGGTGGTA 420
 5 TTGTTAGTTC AGAGCCATTG TCACATTTAA GAGTAATCGT CAGAAGTTTA CTAGGCATTG 480
 CTGTACCAAC TCAAATAGCA ACACATGATT CTGATTTTGC TAAAAATGAA GATGGTTCAT 540
 ATTACTTAAA TGATAGTGAA TTCCAATTAC GAGCAAGATT ATTTGTCGAT CAAATTGTAT 600
 10 CTTTGTGAA TAATAGTCCA TATGAACATT TAAAATAATA TTAAnAAATA TGTAAATnT 659

(2) INFORMATION FOR SEQ ID NO: 705:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 705:

nCACCATATA GTAACGGCA CCAACTACAT TACCGTCTT TAAAAAGATT TTTTATAGT 60
 25 TATTATCAAC ACTATTAAAT ATTTCAATAC CTTTAATTTC TGCATTTTCT ACAATTGAC 120
 CAGCACTATA CAAGTCACAC CCAGAACTT TTAATGACGT AAATGTTGTT GATCCCTTGT 180
 ATCCGTTTCTGTT TCTTTATTT GTTAAATGAT CAGCTAATAC TTTACCTTGT TCATATAGTG 240
 30 GTGCAACGAG TCCATAAACT TTGCCGTTAT GTCTGCACAT TCACCAACTG CATATACATT 300
 GCTATCACTT GTTTCATCA CATCATGAC AACAAATACCA CGATTACATC TAGACCTGAT 360
 CTTGGCACTC CTGGGAAGGC GGAACCACTG CATACTACTA 400

35 (2) INFORMATION FOR SEQ ID NO: 706:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 380 base pairs
 (B) TYPE: nucleic acid
 40 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 706:

TTAAAAATAC AGCTACAGGT AATTTTAATG ATTTTTCATC AATATCAAAT TTGGGATTAT 60
 GGTGTGGCGC TGTAATACCT TTAACCACTC AGAAAGAATG CACCTGGTCG TACTTTCAAA 120
 50 TAATGTGAAA AATCTTCTCC AATCATCATT AAATCTGATT CATTAAAGCG TACATGTAAG 180
 TCATTGTGTG CTTCTTTAAT AACTTGGATA TGCTTTCTCG TTTATTATGG ACAGGCAAAT 240

55

GCTTAATCCA TTTTGTCCAT ACATGATTCT GTATATCTGA AATCGAAAGT TCTnACTGTA 360
 CCTTTACAAA ATGCCTTGnn 380

5 (2) INFORMATION FOR SEQ ID NO: 707:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 707:

CAnAAAGGAT CAAAGTGATC GGTATAGTGA TAATTAGCTC CAAAGAAAGA ATATTCTAAA 60
 TCAAATCCAT ACCAAGCAGA AAGTATTAGC GAATATCAAT TTAAATGGTA CCGATTCAAA 120
 20 TAAAGAAACA CGACATATAG AATTTTACT TGATGATTTT AGTGAATCAT ATGAACCAGG 180
 AGATTGTATA GTAGCATTAC CGCAAAACGA CCCTGAATTG GTTGAAAAAC TAATATCCAT 240
 GTTAGGTTGG GATCCGCAAT CTCCGGTGCC AATTAATGAT CATGGTGATA CAGTTCCTAT 300
 25 TGTGAAGCA CTAACATCAC ATTTTGAATT TACTAAATTA ACATTGCCAT TATTGAAAAA 360
 TGCAGATATC TATTTTGACA ATGAAGAATT ATCTGAnCGT 400

(2) INFORMATION FOR SEQ ID NO: 708:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 447 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 708:

40 TAATGCTGGT ACGGGTCATG CAGCATTATG TGAGTTGAAC TACACAGTTT TACAACCTGA 60
 TGGTTCTATC GACATCGAAA AAGCGAAAGT GATTAACGAA GAGTTTGAGA TTTCAAACA 120
 ATTCTGGGGT CACTTAGTGA AAAGCGGTAG CATCGAGAAC CCAAGAGAAT TTATCAATCC 180
 45 ATTACCACAC ATCAGTTATG TTAGAGGTAA AAACAATGTT AAATTCTTAA AAGATCGTTA 240
 CGAAGCGATG AAAGCTTTCC CTATGTTTGA TAATATCGAA TATACTGAAG ACATCGAAGT 300
 AATGAAAAAA TGGATTCCAT TGATGATGAA AGGCCGTGAA GATAACCCTG GTATCATGGC 360
 50 GGCAAGTAAA ATTGACGATG GTmCAGATGt AAmCTyCGGT GAATTAACAC GTAAAATGGC 420
 TAAAAGCATT GAAGCACATC CAAATGC 447

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1448 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 709:

10 GTGCTACCT GTTTTGTG TTCAATTGTT AACTTTCTT TTTGAATAGT AGTATTCGAT 60
 AATCTTTAT CGCTTATAAA ATCATCTAGT GGATGGCTTT CTTTGAAACT TTTATTTTCA 120
 15 GTCATCAATC ATCTCTCCAA TAGTTGCTAT TTTACATCAA TCTTGATGAT GTTTTGATG 180
 TAACTTATTT AATTCCATTT CAATATCTAA ACGTTCATAA TCATCTTCGT TGAGACGCTT 240
 TAAATCAGCG ATTAATGTTT GTTTGACCTC ATCCAAAGTA ATTCGTGTTT GTTCTAACTT 300
 20 TTGCTGTTCA TTAATTGATT TTTTGGGCAT TTTTGCTAGA CGTGTATATG CATCAACCAA 360
 ATTTAAAGCA TTATCAATAT GAGAATAAAA AAAGCCTTCA ACTTTATAAA ATGATGCAGG 420
 TCTCTGTCTA ACTGTCGTAT AAATAGAACG TGAAATTTGG TATATATCAT TAATCTGCCT 480
 25 AAAATCTTTA ATTGATCTTA TATTGACATA CGTTTTTAAA ATACCTCTAA GTTTTTGGTG 540
 TGTATGATTT AACTGATTTT GAATATAGCG ATAGTCTTTT CTAGTCAAAC CAATTCGTT 600
 TAAATATTTG CGTGAAGTGA GTTTTTGTAT CGGTAGGTAT GTCATTAAAA AGCCAACAAT 660
 30 ACCAATAGAC ATATCAATTA AAAAGATAC ATCAAGTGCA ATCATCCCAA ATATGCTTGT 720
 TAAAAACGCT ACAGGAATTC CCACTAACAC CCAAATATA TGAGAAATAT TATATCTCAC 780
 TGTCTCTTC CTTTATTTAG CATTTTATAT TGATCGAAAA TCCGATAATT TTTGATTTAG 840
 35 TTCTAACTCT TCAAGTTGAT GGCTTGTTAC ATTTGATGCT GGTGAGGCAC CTTCAATTAC 900
 ACCTTGAATA AATCTCTCTA TATCTGCGTC ATCCCCTTGT GCATATATCT CTACATAGTC 960
 40 ATCTACATTT TGAACAGTAC CGACAATGTT ATAGTTCATT GCAATGCGTT GTGTAAAATA 1020
 TCTAAATCCG ACGCCTTGAA CGCGTCCGAA TACTTGTAAG TGTATATGtC TCaTTTTTAC 1080
 CACCTCATAA TGTTATTATA CGTAGTTTTA CTTAAAAAAA CTAATAATTA CTATAGTTAC 1140
 45 TACTTTGTTT GTTCAAGTC GTCAAAGTGT ATTTTCAGAG GATAAAGGTA TAAAAATAAG 1200
 TATAGAGTTT TTGAAGTATG GAAGGGGTCT TTAATAATGT GGACAGTTAC CAAAATTAGA 1260
 GCCGATTATG AGGGATGGTG GTTATTCAGT GACTGGCCAG AAAACATTGT TGAAAAATAT 1320
 50 CAATATCAAG ATTTTGATGA CATGTTTAAG CACTATCAAC AATTGATTAA TCAATGTAAA 1380
 GTTCAGTTCG ATAACATGT CACAGGCAAA TATAATATTT ATGCATTTTA TAATAATTGT 1440

(2) INFORMATION FOR SEQ ID NO: 710:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 519 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 710:

GCTGnAAAAT ATGTTAGAAG CAGATGCATG GGCAAAGCTG GGGTCCTTAT TTATTGTCAG 60
 GTAAAGATGT CTTCAATTCA ACTATTGGAA TATATGGTAT GGGAGATATT GGTAAGCTT 120
 TTGCAAGAAG GTTGCAAGGG TTTAATACTA ATATTCTTTA TCATAATCGA TCAAGACATA 180
 AAGATGCAGA GCGCGACTTT AATGCAACAT ATGTTTCTTT TGAAACGTTG TTAGCAGAAA 240
 GTGATTTTAT CATCTGTACA GCGCCACTTA CAAAAGAAAC ACATCATAAA TTTAATGCTG 300
 AAGCATTGTA ACAAATGAAA AATGATGCAA TTTTATTAA TATCGGTAGA GGACAAATTG 360
 TAGATGAAAC AGCATTAAATC GATGCACTAG ACAATAAAGA AATTTTAGCA TGTGGTTTAG 420
 ATGTATTAGC AAATGaACCG ATTGATCaTA CACATCCATT aATGGGaCGT GaTAATGtTC 480
 TGaTTaCACC aCACATTGGG TAGGCGCATT CAGTTAACh 519

(2) INFORMATION FOR SEQ ID NO: 711:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1349 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 711:

CTATTACAGG TGGCGAACCA ATGTTTTCTA AAAAGTCTAT TAGAAATGTT GTTAAACCTC 60
 TATTAAAGTA TGCACATCAT CGAGGTATAT ATACACAAAT GAATTCAAAC CTAACATTGC 120
 CTCAGATCG TTATTAGAT ATTGCTGAAT ATATCGATGT TATGCATATC TCACACAACT 180
 GGGGAACAAC TGATGAATTC GCAATGTTG GCTTTGGCGC AATGAAGAAG CAACCACCGT 240
 TAAAAGCTAA GTTAAATTA TATGAACAA TGATTTGCGA TGCACGTACA TTATCAGAAC 300
 AAGGAATGTT TGTATCTGCG GAAACAATGC TCAATCAAAG TACGCTACCA CATTTACGAA 360
 AAATACATCA AGAAGTCGTT CATGATATGA AATGTAGCAG ACACGAGATT CACCCTATGT 420
 ATCCAGCTGA CTTTGCAAGT CAATTAAATG TGTTAACTCT AGCGGAAATG AAAAAGACAA 480

	TGTTTCCATG CTTAAAGGAT GATGAAGATC AAAAGTTACT ATCACGTTTA AGAAATGCTA	600
	AAAATGTAAC GACTAGAAAT GACCCGGATG GCCGTAGTCG TTTAAATGTC AATGTATTTA	660
5	CAGGTAATGT AATCGTAACT GATTTCGGAG ATGAAACAGG TACAATTTCG AATATACAAA	720
	AAGATAAATT AACAGATGTA TTTGATAAAT GGTTATCCTC TGATCTTGCT AAATCATTAA	780
	ATTGTCAATG TTCCGAGTTT AGTTGTTTAG GGCCAAATGT TCTTGTTAAA AATATGTACT	840
10	ATCCGAATAT GGATTTTAAA GATAATGAGC GTCATATGCA CAAACAACCA CAAATTATAC	900
	AATTTTAAAA ACTCTTAATT ATGCGGAGAA GCACTTTATC GATAAGTAGT CTCCGCATAT	960
	TTTAATGCTA TTATAAATA AAAACAATT AATTGCTGGC AGTACTCTAC TTAAATAATA	1020
15	AAGGGCATT T AATAGGACTA ATAGTCTATA ATAAAGGGG TAAATTTTAA CTAAAGCAT	1080
	AAACGTGCAT AATCAAAAAG ACAGATTGTA GGTGGAATAT TCGAaCATAA CAGTTCAATT	1140
20	CATCCTTAAC AATCTGTCTT TATATTTTTA GTCTCTAATA TGTGCACTT GAGCTAAATA	1200
	TTCAATTGTT gTTTACTTtC AaTGCGaCGt GCTTtCtC GTtCaCaCG TgwGGTGCTG	1260
	TATCATAAAA CCATTTTTCa ACATCATCTT CTGGATATAC ACCAGGTACA TGTTTAGGTT	1320
25	GCCTTCATCA TCTAACGCAa CAAATGTAA	1349

(2) INFORMATION FOR SEQ ID NO: 712:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 712:

	GAAATTAAAA AAGCAATTGG ACAAGATGCA ACAGTGTCTAT TGTTTGATGA ATTTGATAAA	60
40	AAATTATACA CTTACGGCGA TAACTGGGGT CGTGGTGGAG AAGTATTATA TCAAGCATTT	120
	GGTTTGAAAA TGCAACCAGA ACAACAAAAG TTAAGTGCCA AAAnCAGGTh GGCCTGAATG	180
	GAACCAGGAG GAAnTTGAAA ATATGCTGGG GATTACATTG TGAGTACAAG TGAAGGTAAA	240
45	CCTACACCAG GATACGAATC AACAAACATG TGGAAGAATT TGAAAGCTAC TAAAGAAGGA	300
	CATATTGTTA AAGTTGATGC TGGTACATAC TGGTACAACG TCCTTATACA TTAGATTTCA	360
	TGCGTAAAGA TTFAAAA	377

(2) INFORMATION FOR SEQ ID NO: 713:

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 528 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 713:

	AATGAAnCGA AAAATGACTA TAAAAAGTTT AAAGTGT TTT CACTTATTTT AACACTTGTC	60
10	ATTGTCATTT TAGCAATTAT AAGATTTGTT CATAAAATGA TGTAATTAGA GTGAGACATT	120
	GTTTTATGTC TCAGGATCCA GTTATTCATT ATATCTACAA TATTACGAT TATATAAATA	180
	ACCCGAGATT TTAGTATGAT TCATTtCACT AAAATCTCGG GTTTCTATTT GATAATTTTT	240
15	AATGGGATAT GGCATGTATA CGTTCCTGCC TTTTATCTCA TTTCCAATGA TTAATCTGGA	300
	TATTGTTCTA AAAATGCTTT CGCTTCTTTA TTAACGT TTT TAAAATCAAT ACCTTGTTGC	360
	ATCGCTGCAA AGACACATCC ACAATAACAC TGCCTAAAGA TATTATAGTC ATTACACATT	420
20	TCTATGGaAC GCTCATAACC TTTACTTTTC yTAAATCAC TTGGcAAATA GTTCACaTCG	480
	TATATTTTTT GGACATCCAT ACCAAGTTCA TTGATTAATT GTGCGTTC	528

25

(2) INFORMATION FOR SEQ ID NO: 714:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 731 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 714:

35	TTTAATGGCA TGAAGTCACA GTATATAAAG AATCTGGTGT GACAATCAGT ATGACTAAGT	60
	ATATAAAAAA AGAGCAGGTG cGAAATAATG GCGAAAGAGT CGAAATCAGC TAATGAAATT	120
	TCACCTGAGC AAATTAACCA ATGGATTAAA GAACACCAAG AAAATAAGAA TACAGATGCA	180
40	CAGGATAAGT TAGTTAAACA TTACCAAAAA CTAATTGAGT CATTTGGCATA TAAATATTCT	240
	AAAGGACAAT CACATCACGA AGATTAGTT CAAGTTGGTA TGGTTGGTTT AATAGGTGCC	300
	ATAAATAGAT TCGATATGTC CTTTGAACGG AAGTTTGAAG CCTTTTTAGT ACCTACTGTA	360
45	ATCGGTGAAA TCAAAAGATA TCTACGAGAT AAAACTTGGA GTGTACATGT TCCGAGACGT	420
	ATTAAAGAAA TTGGGCCAAG AATCAAAAAA GTGAGCGATG AACTAACCGC TGaATTaGaG	480
	cGTTcACyTT CTATcAGTGA AATAGCTGAT CGATTAGAAG TCTCAGAAGA AGAAGTGTTA	540
50	GAAGCAATGG AAATGGGACA AAGTTATAAT GCGTTAAGTG TTGATCATTG CATTGAAGCT	600
	GATAAGATG GTTCAACTGT TACGCTATTA GATATTATGG GGCAACAAGA TGACcnTTAT	660

55

CGAGAAATCA T

731

(2) INFORMATION FOR SEQ ID NO: 715:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 830 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 715:

15	TTTTAGrGGa AAGCGaATTa GtCaTATwCg CaGCAGATwG aATGATTGaA AAmGaAATTG	60
	aTTCGATTCC AATTGTaAGA AAAAAAGATA ATCAAAAGTA TGAAGTAATT GGAAGAATTT	120
	CCAAAACAAC AATAGCTaAG TTATTAGTAG CATTATATAA AGAATAGGTG AGAAGTAATG	180
20	GAAAAAATTa AAATTATCGT AGCTTCAGAT TCTATAGGTG AAACGGCAGA GTTAGTTGCT	240
	AGGGCAGGTa TTTCACAATT CAATCCTaAG CAATGTAAAA ATGAATTATT AAGATATCCA	300
	TATATTGAAT CTTTTGAAGA TGTTGATGAA GTGATTCAAG TTGCAAAAGA TACAAATGCT	360
25	ATCATTGTTT ATACACTTAT TAAACCTGAA ATGAAGCAAT ATATGAGTGA GAAAGTAGCA	420
	GAATTCCAAT TGAAGTCTGT CGATATCATG GGGCCATTAA TGGATTTATT ATCTGCTTCG	480
	GTTGAAGAAA AACCTTATAA TGAGCCAGGT ATCGTTCATA GATTAGATGA TGCATATTC	540
30	AAGAAAATTG ATGCGATAGA GTTTCAGTT AAATATGATG ATGGTAAAGA TCCTAAAGGA	600
	TTACCTAAAG CTGATATTGT TTTACTTGGT ATTTGAGAA CTTCAAAGAC ACCATTATCT	660
	CAGTATTTAG CGCATAAGAG TTACAAAGTT ATGAATGTAC CGATTGTACC AGAAaGTGAC	720
35	ACCGCCAGAT GGCTTATATg GATATTAATC CAAAGAAATG TATCGCACTT AAAATAAGTG	780
	AAGaAAAATT AAATCGCATT AGaNaAGAGC GACTAAAACA ATTaNGACTA	830

(2) INFORMATION FOR SEQ ID NO: 716:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1159 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 716:

50	nGAAGAACAA GTTTAnCATC TAAATGCCCG TTAAAACTA CAACTACTTG ACGATGTTAA	60
	ATCAGTGTTT AAnTCTCAAA TGACGCAAAA TAGTGATTTT AATGAAGAAA AGAAAGTGTC	120

AGAACGTATA AAAAAATACT TTAATAAGCA ACTCACTGAG CAAATTGCAC CAATCGTTCA 240
 ACAATTAGCA GATTTACATG TCATTATTAA TCCTCAGTTT AACTTTGAAT CAGCTAATAT 300
 5 AGAGCAACCA TTATTGCACA TCGATTTCAA CGATATGCTA AATGCATTGC CTAAACAATT 360
 AACAAAACGT AAAATTTTGA ATCCAAATGG GCAAAGAGAT ATACATGAAT CAATTTGTCA 420
 AAGTACGTTA GGATTATTAC AACCACAAAT GGGATTATTG AGGCAACAGC TTGAATTATA 480
 10 TGTAAGCAA ATGGCTGTAG AAGCTGAATC GCAATTTGAA AGTTTTGAAG CTAATATTCA 540
 AACGCAAATA AACGATTTAT TAGCATTTGA TTTAGATACA ACACTTATCA ATCAATTGAA 600
 AGATAAACAT CAACAACTGA AACTATTTT ATATTAAGAA AGAAGGAACG TTTTAAATGC 660
 15 CTAATAAAAT ATTACTTGTA GATGGTATGG CGCTATTATT TAGACATTTC TATGCTACAA 720
 GTCTTCATAA ACAATTTATG TACAATTCAC AAGGTGAACC TACAAATGGA ATACAAGGAT 780
 20 TTGTGCGTCA TATCTTTTCG GCAATACATG AAATACGCCC TACACATGTA GCTGTATGTT 840
 GGGATATGGG ACAATCAACT TTTAGAAATG ATATGTTTGA tGGTTATAAG CAAAACGTT 900
 CTGCACCACC AGAAGAATTG ATACCACAAT TTGATTATGT TAAAGAAATT TCAGAGCAAT 960
 25 TTGGCTTTGT AAATATTGGC GTTAAAACT ATGAAGCGGA TGATGTTATA GGTACATTAG 1020
 CACAACAATA TTCAACTGAT AACGATGTCT ATATTATTAC GGGCGACAAA GATTTACTGC 1080
 AATGTATTAA TGACAATGTT GAAGTTGGCT AATTAAAAAA GGTTTAACAT TTATAATAGA 1140
 30 TATnCATTAAC ATCGTTTTn 1159

(2) INFORMATION FOR SEQ ID NO: 717:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 574 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 717:

GATTTAATCA ATGCAGTTGC AGAGCAAGCT GATTTAACTA AAAAAGAAGC TGGTTCAGCA 60
 45 GTAGATGCTG TATTCGAATC AATCCAAAAC TCACTTGCTA AAGGTGAAAA AGTACAATTA 120
 ATTGGTTTCG GTAACTTTGA GGTACGTGAA CGTGCTGCAC GTAAAGGTCG TAACCCTCAA 180
 ACTGGTAAAG AAATTGATAT CCCAGCAAGT AAAGTTCCAG CATTCAAAGC TGGTAAAGCA 240
 50 TTAAAAGATG CTGTAAAATA ATTTTACTTA AAAAGCCCTG AATAAGGGCT TTTTATTTTG 300
 CTTTAAATAC TTACAACCTG TACATAAATT GTAATGTCT TCTAAGTTT TAATCTTTGG 360

AACTACAAAA TATACATATG aATATTGaGa TTAATTGTTA GCGTTGaATT TACTTAAAAG 480
 GTAACCATGT CTACTATAGT ATTTTACGTT ATTTAAAAAG ATGAATAATG TAAATGAAGT 540
 5 AAAGGTTIATT ATGAGAATTA CAAAAGCTAC ATAA 574

(2) INFORMATION FOR SEQ ID NO: 718:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 493 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 718:

AAGaTATTGA AAmTGmAaTG GaGCACGCCC TTATTTGATA GAAGTAAAAG ACATTTAATT 60
 20 CTTACCGATG CAGGTCAAAT TTTTATGAG AAAAGTAAAG AAATTGTtGC ACTGTATGAT 120
 TATTTACCAT CTGAAATGGA ACGCTTGAAT GGA CTGGAAA CAGGACATAT AAACATGGGC 180
 ATGTCGGCAG TCATGAATAT GAAGATTCTT ATCAATATTC TTGGTGCATT CCATCAACAA 240
 25 TATCCAAATG TTACATATAA TTTGATAGAA AATGGCGGTA AAACAATTGA ACAGCAAATT 300
 ATCAATGATG AAGTAGATAT AGGCGTGACC ACTTTGCCAG TCGATCATCA TATTTTCGAT 360
 TATACTACCC TAGATAAGGA AGATTTGCGA CTTATCGTGA GCAGAGAGCA TCGACTCGCA 420
 30 AAATATGAAA CTGTTAAACT CGAAGATTTA GCAGGTGAAG ACTTCATTTT ATTTAATAAA 480
 GACTTTTACT TGA 493

(2) INFORMATION FOR SEQ ID NO: 719:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1257 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 719:

GCATTGCCAA AACAAACGT TTATATTATT GAGGAAC TGTATAAAAG TAATGAATAT 60
 45 CAAAATAAAA AATCATATTA CGAAACACTT GTTAACCAAG TAATTGAACT TAAACAGGCA 120
 GATGATTTAA TTATTGGA CTGCTTATTC GTACAACGCT TAGTCGTCGA TCATTTACAC 180
 50 GTTGTGGGTG ATATTTATGA TCGTGACCA CAACCAGATA AAATTATGGA TACACTGATT 240
 AATTATCATT CCCTAGATAT TCAATGGGGT AATCATGATG TGCTTTGGGT TGGAGCCTAT 300

5 GATATTATCG AAGACGCTTA TGGCATTAAAT TTAAGACCAC TGCTTACTTT AGCTGAAAAA 420
 TACTATGACG CAGATAATCC TGCTTTTAAAG CCTAAAAAAA GACCTGACAA ACACGAACGT 480
 10 TTAAGTCAAC GTGAAGAAAG TCAAATTACT AAAATTCATC AAGCTATTGC GATGATTCAA 540
 TTCAAGTTAG AAATACCAAT TATTAAACGT CGTCCAAATT TCGAAATGGA AGAACGTCTT 600
 GTGCTTGAAA AGGTTAATTA TGATACAAAT GAAATTACAG TTTATGGTAA TACATACCCA 660
 15 TTGAAAGACA CATGTTTCCA AACTGTCAAT CGTGATAATC CAGCAGAATT ACTACCTGAA 720
 GAAGAAGAAG TCATGAATAA ACTATTATTG TCATTCCAAC AATCTGAAAA ATTACGTGCT 780
 CATATGTCTT TCTTGATGCG TAAAGGCTCT CTTTACTTAC CATATAATGG CAATTTACTC 840
 ATTCATGGTT GTATTCCAGT TGATGAAAAT GGTGAGATGG AATCATTGTA AATTGATGGT 900
 CATACTTACA GCGGCCAAGA ATTATTAGAT GTGTTTGAGT ATCATGTCCG TAAATCATTT 960
 20 GATGAAAAAG AAAATACTGA TGACTTATCG ACGGATTTAG TTTGGTATTT ATGGACTGGG 1020
 AAATATTCGT CACTATTTGG TAAACGTGCC ATGACTACGT TTGAGCGATA CTTTATTGCA 1080
 GATAAGCTT CTCATAAGA AGAAAAGAAT CCGTACTATC aTCTTCGTGA AGATGTGaAT 1140
 25 ATGGkTCGTA AAATGCTCaG TGaTTTCGGA TTAAATCCAG ATGAAGGACG CATTATTAAT 1200
 GGTCAACAC CAGTGAAAGA AATCAATGGC GAGATCCTAT CAAGGCTGAT GGAAAGA 1257

(2) INFORMATION FOR SEQ ID NO: 720:

30

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 720:

40 TGAGCACATG ACAGTGGCTG AAAACATTAA ATTTTTTAAA TCACTTTGTA AAAATCCAAT 60
 TAACGATACA ACTATCAACG AATATTTACA GCAATTAAAC TTTGATGATA CGTCTGCCAA 120
 AGTATCTACA TTGTCCGGTG GGAATAAAG TAAAATTAAT ATATTAGTAG GTTTACTAGG 180
 45 TCAACCTCGA ATTCTCATTT TAGATGAACC GACAGTTGGT ATTGATTTAA AATCTAGACA 240
 TGACATCCAC CAACTACTTA ACATCATGAA ATCTAAATGT TTAATTATAT TAACTACCCA 300
 TCATTTAGAT GAAGTTGAGG CACTTGCGA TGGTTATCAA GTTAATTGGG CCAGGTnCCC 360
 50 TTTTnTTCAA CAGTTTTTGG GGGCCAACCA TGGGCTTATA 400

(2) INFORMATION FOR SEQ ID NO: 721:

55

(A) LENGTH: 570 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 721:

10 CTTATTTTATA TTATAATATA AATATTTTAT TTTTCAATAC TTAATTTTAA AAATCCCCTA 60
 TGAAAATGTT GATGAGCTAC ATCTTTAAAT CTTATTATAC ATTGTATAAA ATTATATTGC 120
 GAGGTAGTAA ATTGATATTA TACACTTTTG GTTGTGTATG CATCTCACCT ATCTTTTTTG 180
 15 CTTTTTTTCAA ATAAGAAAGT ATTAAAAATy AATGCCTATA CAGGAACAGA CATGGCTAAA 240
 CTTGATTTAA ATAGTCTTGA CGACGAGCAC GTAAAATTAT TAATAAATGA ATTAAAATAT 300
 CCAGAACTC ATATCGATGT AAATGAATTA AAAACAATAG TTGCTAGTCG AATAAATGAA 360
 20 AGGCAAGAAA TAATAAGTTT TAAGTTAGGA ATAAAGTACT TATTAACAAT AAAAAGAGGG 420
 AACATAGAAA AAGATAGGTT TTCAATTTCA ATCATTTTCA AAGATACCTA TCACACCCTA 480
 GTTAGAATAG ATATTAAACGG TGGTACTCAC GATAATCCAG ATGGAACAAT CGCTCCGAAA 540
 25 AGTCATATTC ACATATATAA TGATAAGTAT 570

(2) INFORMATION FOR SEQ ID NO: 722:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 722:

ATTTTAGTCT TAGGTGTTGA TTGCATGATG AATGCAGAAG TTATACCTGC AGCCATATTA 60
 40 GCACCTTCAT TATTGTGTAT AATTGGAATA TTATTGCTT TTAAAGTATG TCCAATTAAG 120
 TTTGAAGTCG TTGTTTTACC ATTTGTTCCA CTGATAAATA CAATATCATC AACTTGCTCT 180
 45 GCTAATTTTC TTAATATATC TGTATCCACT TTTCTAGCGA TTTGTCCAGG TAAATCTGTT 240
 CCTCTTTTAC CTACTGCTCT ACTTGCTTTA CGCGCCAATT TCGCTAGATG GATTGCCGTC 300
 CACTGTCTCA TGTGTTTCCT CCTCAAATt CCACTCGCAT CATTATAACA TGACAAGGCA 360
 50 ACTTCAAAAA AGTTTCTCAA TCACAAATnG ATACCAGTGT 400

(2) INFORMATION FOR SEQ ID NO: 723:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1113 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 723:

	GATTATATGC TGTGGCAAA ACAATATACC CAGATGAATT TTTATTTTGA GTAATAGTAT	60
10	AACCATGCTT TGTCTTACTA ACTTTGACAG CTTCTAATGA ATTTGAAGCG TCTTTAAGCG	120
	TGGTATAATC TTCGCCGTAT ATCCCTTTTA AATTTACTCG ATACTTACCT TTAGGCAATG	180
	ATAATCTAAT TCTATCTGGA GCTTTAATGC GTATCGTTAC GGGTGTACAC ACGCGTCGAT	240
15	ATTTATAAGT GAGTTTATTT CTTTCTGTG TATATTCATT CACTTTAACA TCATGAGCTT	300
	TATCCGGCGA AAGTAATTCT AATCCATTT CAAAATACAA ATCTTTAAAT TGATTAGAAA	360
	CTGATTTTGG CAACTGTACA GTTAGACCAC CATTATTTTG TTTAACTTGT AATAAATGTT	420
20	TTGTAGGAGA TTGCCAGGCT GCACTATTTA ATTTAATTGT TGAATCTGAT AGTAAATTTT	480
	TATTGGCTTT AAAATGTGTA TTAACATCTT TAATATTGTT AGAAACAATC CCTTGCAACA	540
	TTGCTTGTTT TTTATCTAAT GGAGATTTTA ATTCTTTATT GGAAAAGACC TTATTTGTAA	600
25	TATGTGCACT TGGATAATGG ATGGTATTTT TAGAATGAAT CCAACGAAC TATTTGTCTT	660
	TGTGTTTCTA CTAAATTTTA AATCCATATG GtAAGTTGTC ATCATGATtC ACTCTAATTC	720
30	GATCATTAAAC ATTCCAAAGT GATAGTaAAT TTgACGATgC CAAGTAATCT ATAAGTGCTG	780
	TTTTTATCGA TTGGCATATT AATTTGGAGT GTCTTGTCTAT AATATTTTAA AATGTCTCCA	840
	TTAAAAATAC TAGAATATAA TGAAATGCCA TTATAATGAT ATATAAATGG TGAATTTAAT	900
35	GCATAGTCTG ACATATAATC AATGCGATTA AATGAGCCTG TTGCATTTTG ATTTATCTTT	960
	TTTATAAGCT GGTTTACATA GTTACTATGG TAATCATGTT GTTTCACCGT TGATAATGAT	1020
	TGTTGATAAG GTTTGATTGC CATGTTTTTG TTGTTATCTA AAATGACGAT TTGTTGAATC	1080
40	ATaACGATTA ATACTAATAT TGCAACGGTT AAT	1113

(2) INFORMATION FOR SEQ ID NO: 724:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 464 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 724:

	TGCCGTTTAG CAAAATCTAG TATCGCTTGA TGATCAGATT CTGAAATTC AGTGTGTACT	60
--	--	----

55

GATTGATTAA GTTTATATGC AAGTGCATGT TCTCGTCCAC CAGCACCAAT TACTAATACA 180
 TTCATTTATG TTACGCCCCCT TTAAAATTAG TGTTTAAAAT GTCGAGTGCC TGTGACTACC 240
 5 ATTGCAATAC CATGTTTATT AGCCATATCA ATTGAATCTT GATCTTTAAT CGAACCACCC 300
 GGTGGGATAA TTGCCTTTAT ACCATGTTGT GCTGCAAGTT CAACTGTATC TCCCATAGGG 360
 AAAAATCCAT CAGATACTAA CGGCTACATG ATCATTGATT TCAATAGCTC TCTCTAACGC 420
 10 AATTTTAGCA GCACCGACAC GATTCATTGG CCAGCAnCTA TACC 464

(2) INFORMATION FOR SEQ ID NO: 725:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1440 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 725:

GTTAAATnCA nTATTnAAA TaTATTACCT TATTAGAAAA AGTCGTAATA TGAGGTGTAC 60
 25 AAATGACGCA AATTTTAATA GTAGAAGATG AACAAACTT AGCAAGATTT CTTGAATTGG 120
 AACTCACACA TGAAAATTAC AATGTGGACA CAGAGTATGA TGGACAAGAC GGTTTAGATA 180
 AAGCGCTTAG CCATTACTAT GATTTAATCA TATTAGATTT AATGTTGCCG TCAATTAATG 240
 30 GCTTAGAAAT TTGTCGCAAA ATTAGACAAC AACAACTCTAC ACCTATCATT ATAATTACAG 300
 CGAAAAGTGA TACGTATGAC AAAGTTGCTG GGCTTGATTA CGGTGCAGAC GATTATATAG 360
 TTAAGCCGTT TGATATTGAA GAACTTTATG CAAGAATTCG TGCAATTTTA CGTCGTCAGC 420
 35 CACAAAAGGA TATTATCGAT GTCAACGGTA TTACAATTGA TAAGAACGCT TTTAAAGTGA 480
 CGGTAAATGG CGCAGAAATT GAaTTAACAA AAACAGAGTA TGATTTACTA TATCTTCTAG 540
 40 CTGAAAATAA AAACCATGTT ATGCAACGGG AACAAATTTT AAATCATGTA TGGGGTTATA 600
 ATAGTGAAGT AGAAACAAAT GTCGTAGATG TTTATATAAG TATTTTACGA AACCAAGTTAA 660
 AACCATACGA TCGTGACAAA ATGATTGAAA CAGTTCGTGG CGTTGGGTAT GTGATACGAT 720
 45 GACAAAACGT AAATTGCGCA ATAAGTGGAT TATTGTTACC ACGATGATTA CGTTTGTCAC 780
 GATATTTTTG TTTTGTTTAA TTATTATTTT TTTCTTGAAA GATACACTGC ATAATAGTGA 840
 GCTTGATGAT GCaGAACGAA GCTCaAGCGA TATTAATAAt TTATTTCaTT CTAAGCCTGT 900
 50 TAAAGATATA TCTGCaTTAG ACTTGAATGC aTCTTTAgGT AAtTTTCaAG AGATAATTAT 960
 TTATGATGAG CATAATAATA AATTATTTGA GACATCGAAT GATAACACAG TGAGAGTTGA 1020

55

ATATTTAATT ATTAAAGAAC CAATTACAAC GCAAGATTTC AAAGGGTATA GCTTGTTAAT 1140
 TCATTTACTA GAAAATTATG ATAACATCGT AAAATCATTG TATATCATTG CGCTGGCATT 1200
 5 TGGAGTGATT GCAACAATTA TAACTGCCAC AATCAGTTAT GTATTTTCAA CACAAATTAC 1260
 TAAACCGCTT GTCAGTTTAT CAAATAAAAT GATTGAGATT CGACGAGATG GTTTTCAAAA 1320
 TAAATTGCAA TTAAATACAA ATTATGAAGA AATAGATAAT TTAGCAAATA CGTTTAAATGA 1380
 10 GATGATGAGC CAAATTGAAG AATCATTTAA TCAACAAAGA CAATTTGTTG AAGATGCGTC 1440

(2) INFORMATION FOR SEQ ID NO: 726:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 375 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 726:

TGATATTCCG ACTTGCGAGG ATTATTGCGC ATTAAGGATT AACGCAGTAT GAGTCCAAAG 60
 25 ACGCGCGACA CCTGGAAAAA GGATTACCTA ATGCCTTATT TACAGTAACC TTGTATGATA 120
 AAGATCGGTT AATTGGTATG GGTAGAGTGA TTGGCGATGG CGGAAGTGT TTTCAAATTG 180
 TTGATATTGC AGTTTTGAAA AGTTACCAAG GTCAAGGTTA CGGCATCTAA TTATGGAGCA 240
 30 TATTATGCAA TATATTAAAG GTGTGGCTGT TGAGAGTACA TACGTTATCT GATTGGCAGA 300
 CTACCCAGCG GGnTAAATTA TATACCAAAT TTGGGTTnA TACCTACCGA ACCAGAnTCC 360
 AGGCGGTGAT GTATG 375

35

(2) INFORMATION FOR SEQ ID NO: 727:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 471 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 727:

TAAAAAACTT AGAGATGAAA ATCATATTAT TTATCTGTTT TTGGGACAAT TTTTACTAA 60
 AAATGAAGAT CCATGGCATC AAATACTTAA TGATTTAGAA GTTACAAATT CTGTTGATAA 120
 50 TTTTTTAAGG TCAATAAGTA ATAAGGCCAA AGAAACAAAA AAAAGAGCTT TTATTATTAT 180
 TGATGCGCTT AATGAAGGTG AAGGTAAAAG GTTATGGGGA AATTATTTTC AAAGCTTTAT 240

55

AGATGTTATA TTACCCAAAA ACGCAATACA AGATAACAAT ATTGTAGTAT TTCAGCATGA 360
 AGGTTTTAGT AAGGAAGAAA ACTATAATCC AATTGTATCT TTTTGTGATT TTTATGGATT 420
 5 AGAGCTACCT AAGTTACCTA TATTAAATCC AGAATTCAAC AATCCATTAT T 471

(2) INFORMATION FOR SEQ ID NO: 728:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1750 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 728:

TAGGATGTTT CACCCTgGCA ATTGcTCTAA GACATctACa GCAGTAGAAT aycCTGTTTT 60
 20 CGTctTTTTA ATAACAGGTA ATTGTAATGT CTCAAACAAT ACAACACCTA ATTGCTTAGG 120
 AGAATTTATA TTAATCTCTT CACCAGCTGC ATCATGGATA TTTGGAATCA AGACGTCTAA 180
 TTTTCTTGA ATTTCTTTTT CCATTCTTC TAAATCATGA ACATCTGTAA ATATACCAAT 240
 25 TTCTTCCATT TCACTTAAAA TCTTAGCTAG CGGTAGCTCT AAATCAGCCA AGAGTTCTAC 300
 CTGATTGTAT TCTTCTAATT GTTTATCCAT ATTTGGTTTC GCAAAGTAAA TTGCATCAGT 360
 AATAGAAGCA ACATATGGAT TTAACATC ATCTTCAGGT ACCTTAAATT TCTTACCTTT 420
 30 TCCATATATA CTCACATCGT CTTTCACAAA ACTTTGACCG TACAATGAAA CAACTGATTG 480
 AACATCACTA ATCGTACGAG ATGGATCAAT TATATAACTG GCCAACATAA TATCGAAAGA 540
 AATATTTTGA ATATCAATCC CCAATCTATG TGATGCTACA TATGTTTTTT TAGCATCATA 600
 35 TACGACTTTT TTCGAATTCG GATTTTCTAA CCATGAAACT AGTTCGACAT AATTATTTAT 660
 GTCATCCGCA TTAATTACAA TATGTTTCTC ACCTGtAAAT aAAGAGAATT TTAAAAtATT 720
 40 ATTtCGCAAA TAGtTACCAC CGtCTAATTC GAAATGGATG GCCGCTTCTT TcAATGAAGT 780
 AAAATCcAAT ATTATCAAAA GACGTTTCCA ATTTCAAATG TCTTtCTAT TGCATCTTCA 840
 ACGtTtGCTG ATTGaTCAAT GTCAGCCAAc AATTGTTTGA ATTCTAACTT CTTAAACAAT 900
 45 TCGATTTTTT CTGTGTGTTT ATCTtGAKGA GTCATTAACG TATCTTCaAG TTTTACTTCa 960
 ATCGGACTAT CTACATTAAt CGTTGCTAAT TCTTTACTCA TTAATGCATC TTCTTTGCTA 1020
 TTTTGAAGTT TTTCTTTTAA CTTTTTACCT GAAATTTTCAT CTAAATGTTT ATAGACACCT 1080
 50 TCTACTGTGT CAAATTGGTT TAGCAATTTT ATTGCTGTTT TCTCTCCAAC ACCTGCAACA 1140
 CCTGGTATAT TATCAGAAGT ATCTCCcATT AATCCTTTCA TATCAATAAT TTGATTAGGT 1200

CCTTTTTTAG TGTAATAAAT GGTACATTA TCCGTTGCAA GTTGTGTAA ATCTCGGTCT 1320
 CCCGTAATAA TAATTGTCTG AAAGCCCGCT TTATCTGCTT CTTTACTTAA AGTTCCGATA 1380
 5 ATATCATCTG CCTCATAGTT ATCTAATTCA TAACGTTTAA TATGATAAGC ATCTAATAAT 1440
 TGGCGAATAT AAGGAAATTG CTCACCTAGT TCAGGCGGCG TTTTCTGGCG TCCACCTTTA 1500
 TATTCACAT ATTTTTCATG TCTGAAAGTC GTTTTACCTG CATCAAACGC TACTAAAAAA 1560
 10 TGATTTGGCT TTTCTCTTT TAAAATCTTC TCTAGTAACA TTGCAAAACC ATATACTGCA 1620
 TTGGTATGAA TGCCTGCTTT GTTTGATAAC AAAGGTAATG CATAAAAGC TCTAAAACTT 1680
 AAGCTATTAC CATCGATTAA TACTAATTTA TTCACAATTT TAACCTCCAG AACTAATTTA 1740
 15 TATATnTGTG 1750

(2) INFORMATION FOR SEQ ID NO: 729:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 439 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 729:

AAATTAATTT TAACATCCTT TCAAAATAGT TTTAACGGAT CcTCCcAAAA CGTAAACTCA 60
 30 CATCAGTTAC TTGTAACATG CATTTTCTCC TTTTTCAT TCGATATTCT AACGGAAGAA 120
 TTATATCATA TTATCGTCAC AGTTTCGACC TCATATAAGT TGTAATGATA GAATGACTCA 180
 CACATGTTAT AATAATAAAG AATACAAGAA TCGAAGGAGA ATAAACATGG CATTAGACAA 240
 35 AGATATAGTA GGGTCTATAG AATTCCTTGA AGTAGTAGGG TTACCAAGGT CAACTTACCT 300
 TTTAAAGGAC CAACGGTGAA ACGTAAAGTT AAACCAATCA GAATGAACGT GATGATTGAA 360
 40 TTAGAAGTAG GGGAGnATAT AGTTCCTCAT TTATCCAACC GGTGAGGTGG ATTATTGCAC 420
 TCAAATATGC nGnTATACG 439

(2) INFORMATION FOR SEQ ID NO: 730:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 481 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 730:

55

AGTTTAGGAT TGATGACAAA GCCATTACGA CCAACAAATC CTAAACCTGC ACGTTCTGCT 120
 ACTGCCCTAT CTGATAATAC ACCCGTATCT ACCATAGATT TGATTTCAAC ATCTGGAAC 180
 5 TTAGATTCAA TAAATGCAGC TAACATGTCT AATCGTTTAC GCATAATTGT ATGATAATCT 240
 TGACCCACG ATGCTCTAGC AAATAAGCCT CTGCGATCAC CTCTAACACT CTTAGGTGCA 300
 10 CCTTTCAGTT TGTTAGGATA ACCAACTGCA ATTGCTATGA TTGACCTTGC TGTGTAAG 360
 GATAATTTAG GCTCCGTTCC TAAAGCAATA TCAGATTCTT CAAATCCTGA GGCATAACCA 420
 TTTGCATGAT ATGCTTCTAG CTTTGTCTC AATTCATCAA AGGGnTCGGC AGTAGTAAAT 480
 15 C 481

(2) INFORMATION FOR SEQ ID NO: 731:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 828 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 731:

GTGATGAAAA ATCTTTTCT ATTGATTGAG AAAATAGTTA TTTATATAGT TTTTAAAGAA 60
 AAAAAGTGAT AATATTTAGC TAATCTAATG AAAATTGATT TATTGGACAA AAAATACATT 120
 30 TTAAATGAGT AGAGGAGGCT GCAAATGGGT TACAGAACTT TAAAAAGTAT TTTTCATGAA 180
 CACAATGAAA GTAAATGAA GGAAGAGTAT ATTAAAGAT TTAATCTTT mGCTTCTTTC 240
 35 AATACTAATA TTAATATCAT ACCTATGGAA AATGGAAAAA AAGTTAATGA TTTGGAATAT 300
 CCTCTATTCT TTATGGTGAC TAAAAATCTA TCAAAAAAAC AAGAATTAAT ATCAATTAAT 360
 AGTAGAAAAA TTGATAGAGC ACTTAATTCT TTACCATATG CAGCTAGAGA ACAATATTTT 420
 40 AATGATTTAT TAATCGATGA ATTACAAAGT ACTAATGAAA TTGAAAATGT ATTTAGTACT 480
 AAACAAGAGA TTGCACATGC GTTAAATAAC CAAGCATCAG AATTTCTTAA GTTCAGAGGC 540
 CTCGTGGATC AATATAAAGA GATAGAACTT AATAAAAAAA TTAAAGTTGA TAATGTAAGA 600
 45 GACATTAGAG CGATTTATGA TAAATTAGTT TCAAATGAAA TTAACGAACA AGATAAGTTA 660
 GATGGAGAGC TATTCGTAA AAATTTGTG GGTGTGCATG ATGGGTCAAC GAATAAATAT 720
 ATACATGTTG GGTTACAACC TGAAACCAAA ATTGTTGAAT TTATAGGTGA AATGCTAACA 780
 50 TTTTAAAT ATTTGATGC GCCTCAGCCG TTCAAATCA TGGCTAGT 828

(2) INFORMATION FOR SEQ ID NO: 732:

(A) LENGTH: 1622 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 732:

10	CGTATGTTTC GAATTTTATG nATTGCATTG GATAATAGTA CCAGTTAAAA GAACTAAGAT	60
	TGTAATTATC CCCCCTTAAA ATTCAAGAAT ATATTTTCTT AATCTATCAG CAAGAGGTTA	120
	TATTTCAAAG TGTTTATTTT TAAAACAACA TCTAAAAGCC ATTTcATAAt AAwGTAAAT	180
15	CaTTAGaATG TATAAGATT CCaATTAATA AAAATAGATA AAATGCAATA AAGTTCCAAA	240
	TACTTATTAT TATTCATCAC AAATTACGTG ATGCCCTCTA CAACACTAAA TCAAGGATGA	300
	TATTTATTTT ATATACATTA GCATTCATTG TTTGCTTACC CACATTATCC TGTTTAGTCT	360
20	TTTTTATAAC TCTAGTCTTG AATGTCATAT TTACAAAGAA AAAAACACTA AAATATTTAA	420
	AACAACTACT ATTTATTCAC TTTGTTCTTT TATAGCACCT TTTATGTTTA TTTTCTTGTA	480
	TTCAATAAAT AGTATGGAGA TGATAATTTA AAGTAGATAT ATTTGAAGTA TACTATAGAG	540
25	TAAATTATAT TGGGGAAATA TGTAACGAA CCGGTTGATT TTATAGTGGA TTAACACTTC	600
	ATAGAGAATA TAATCAAGAA CAAAAACAGT CAATAGGTGT GAATTTTTTTT AAAGACGGAT	660
30	CTGTTGCTAC TAAAAAAGA AGTATACAAT TCAATAGTTA AACCTTAAAA CAAGAAATAT	720
	TATTCAAAAT CAATGAATTT CCTATCTTAT TAGTTTTAAC AATATTTATT CTCATAGAGT	780
	TTGTCCAATT AAATATAGAT GATTCAAATA TTTATAAGCA TACAAAAGAG CAGTAAGACA	840
35	TTTTCTAATA GAAAATAACT TTAGTCTCTT TTGTACATTA CTCATCAACT ATTGTAAATT	900
	AAATAATAAA TAACTACCTA TTTTATTATT CAGCTAAAAA ATGCTCTGCT AAATATTTTG	960
	CTGCGCCATC CTCTTCATTC GTGTATGCCG TTACATCTGA AGTTAATGCT TGGATTCAG	1020
40	GGCGTGCAAT TTTCATAGCA ACTGTATAAT GACCAAATTC AAACATTGCT CTATCATTGT	1080
	CGCTATCTCC AATAACTAAC GTTTCTTCTT GATGAATACC AAAATGTGCA ATCATTCTTT	1140
	TAATGCCTGT ACCTTTATCA GTTTGATAAG CCATTGTTTC CGCATTAAAT CTTGATGAAT	1200
45	TTGATACACT AATCTGTAGT TGCACATGAT TTTGCTTTAA TTCATCTCTA AATGCTGTTA	1260
	TTTTTCTIAA ATTAGAACTG AATAGATAAA TTTTGAATA TTCACCTTCA GGAAACTTAG	1320
	TTACCCAATC TATCTTACCA GCAAGCGCAT CTTGTCTTGA AGACCATTCA CTATGACTTA	1380
50	CGCCATTAAT AGGATCTTGA CTACGAATCA TATCTCGCAT CCATGTTTCA TCTTCTTTTA	1440
	AAGAACTCT ATTACCTTCA AAAGGAAATA CCTCATAATA AATTGTTGG CGTTAGCTA	1500

55

CTACTTCTCC AATTGTTCCA TTTGAACTAA TGATGCCATT AACCGCAAAA TCTTGAGGTA 1620
CA 1622

5 (2) INFORMATION FOR SEQ ID NO: 733:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 base pairs
(B) TYPE: nucleic acid
10 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 733:

AATGCCAAT AAATCCAAAT TCTTTTAACG AACAAGAGTG CCCAAGTTAT TGCTGATCGT 60
TATGCCAGAT AAGCATATCA ATGATAATTA TGGTTTAGAA AGAATTTCTA AGACAAATCA 120
20 TGGATATAAT TATGTGTATT CCAATGATAA TTCAACTAGT AAGCAACATG TAAGTATTTT 180
AAATCAAGGC ATAATAACGA AATAATAGAT GGAACAGTGT ATTCTAATTG GATATACTGT 240
TTTTATTTTG CAATAATTTA ATTTAAAAAG GTGAATTCAA CTTATAAAAT GATGTAAATG 300
25 TTATGTCAAA ATCAACCAAT CCGTAATGTA TTTTAAAATG TTAATATAGT TCTGAAGAAG 360
TATAAATGAG GTGTTGAAAT GGCTAAAAAT AAGAAAACgA ACGCGATGCG TATGCTTGAT 420
CGTGCAAAAA TTAAATACGA AGTTCATAGC TTTGAGGTAC nTTAAGAACA TT 472

30 (2) INFORMATION FOR SEQ ID NO: 734:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 519 base pairs
(B) TYPE: nucleic acid
35 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 734:

CATAGCCCCA AATTTTTTCT ATTATTTGTT CACGAGTAAA GATTTGCTTA GGACGTGCTG 60
CAAGCATAAA TAATAATTGA AATTCCTTGT TCGGTAACGT CATCGTTTTA TTAGATACTT 120
45 GGAGTTCCAA ATAGGATTGG TTTAGCGTTA AGTTGCCAAT AGTCATTTCT GAATTTGAAT 180
TGATATTATA TCGACGTAAT ACAGCACGAA TTCTAAAAAT AAGTTCCTTA ACCTCAAAGG 240
GTTTGGTTAC ATAATCGTCA GTACCGCTTA TAAACGCACG CTCTTTGTCA CTAAGTGCA 300
50 CCCGCGCTGT TAACATAATA ACTGGTATAT CATAATCATT TTTAATGTA TTACATAATT 360
GAAAGCCGTC CATACCATCC ATCATAATAT CTACCACTGc AATATCGaCA CGCTGTTTTT 420

55

GTTTTGTdAA ATGGGCTAGC TATAATAATT TAGGGGATT

519

(2) INFORMATION FOR SEQ ID NO: 735:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 735:

ACTTGAAAAC GATAAACAGT ACTTTCAGC TACACATTGG AAAGCTATAA ATGGGATACC 60
 TTATGCAGGC AGTAGTGATA TTGATGGATT GCCTCAAGAC GGTATCATT CGGTAAATGA 120
 TAAAAATAAA TTAGATAATT TAAAAATAGG CGAACAGGAA TTATTCAAAA TAGCATTGTA 180
 CAGAAATCCC CAAACGGTAA ATTGTGGAAA ATAACAGTTG ACGATAGTGG GAAACTTGGT 240
 ACAGTGCTAT TTTATTAGAA AGGAAGGTGC ATTATGGAAA ATTTGTATTT AATAAAGGAT 300
 TTGGGAGCTT TAGCAGGTCG AGATTATAGA GCTAAGGAAA TACAAAACCTT ACAAGGAATA 360
 GGCATTGCGC TGGGCTGACC ACCAGGGTTT AnGTGnCCnC 400

(2) INFORMATION FOR SEQ ID NO: 736:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 780 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 736:

CTTTTTTATT GAAAAAGTGG TATTGATTA TCTATTTACT ATTTATATTA GCGGCACTCC 60
 TTATTACATT AACGACAATC CAACATGTAA CAGAAGATGA CAATCATnTT AATATAGGTG 120
 TCGTAGATAA AGATCAATCA AGTGAAACGA AATTAATCTT AAACCTCTATT GGTAAAGGGA 180
 GTAACCTAGG AAAAAACGTG AGCATTAAAG CATATGATGA TAAGCAAGCA CATACTTTGT 240
 TAAAAAACA TAACTTCAA GGCTATTTTG TTTTGTATAA AGGTATGACC AAGGCATTTT 300
 ATAAACAAGG CGAACTACCA ATTTCACTAT ATACATATGA TCAACAATCC ATGAAAAGTG 360
 TCGTGCTATC TCAGCTAACA GATTCTGTTT ACCAACGTCT TATGCGATCA ATGGGTGGCA 420
 TCTTAGCTTT TCAAGACTTA GCACCGAAAG CATCACATTC TGACAGTATC AATGTTATGA 480
 CTGATTGCT GATTACAGGA TTAAACCGTT CAGGTGCAT TAACTTAGAA CCGATTCATT 540

CACTATCTTT ATTTACAGTT TTGAAAATGA ATCAAGATAC TGTATTGAAA GCGCGATTGA 660
 AAATGTTTCA TTTTCTAAA GAGCGTTTAT TAATCATTGG TACGTTGATT ACATGGTTTT 720
 ATACTATGTT ATGGGnGTAT CGnTGGTGTA GTTTGGAATG TGGTTCAGTA TTCCGAATnA 780

(2) INFORMATION FOR SEQ ID NO: 737:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 946 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 737:

TATTTAATTT CTnATTnTG GTTTTAAATT nTAAGATGAC CATAAATGAA TTTCTTCTGC 60
 TGT CATAGAA TTAGAGCAAA ATACGAGCAT AAATTTAACT AATTTCTTGC TTGTCTTATA 120
 GTCTTGATTG CCTTGTA AAA ATGTTAGTAT TAATTGCTTC ATCGTTTCGA ATTGTTGTGT 180
 TGATTTAATA GCGATTGCTA AACCTATATT TTTTTCGAAA AGCTTTTCAA AACATCTATT 240
 TAACAGGTTG TAATTCGTTT GGCTAATCTC AATATCGTGA ATATCGTCAA TGAAAACGAC 300
 AGGTTTTGGT AAGTGCTCAA AGTTAATATC GTAATATTCA TTGAAAATAA ATTGAAATAG 360
 TTCATTGAAA TTAATAAATC GAACAAATAC TTTGGCACGC TGACTTCGAT CTTTCGGATT 420
 AAAATCATCT ATATGAATAT CTGTTGTATT TGTAGCCAGT TGAGCTGTTA AATCAATTGT 480
 TGAAATTAAC TCAGTGAATT GTGATACGTT ATCTTGTGTTG AAGTGTATTT GTGGTGGTTC 540
 AGTAATCTTG GAAACGAGTG AACGGAAGTG TTTAGGACTA AAATGCAAGT AATTCTTAAA 600
 TTGATTTGCA AAATTTGTAT GACTGTTAAA ACCGGCCAAT TCAGAAACAG TTGTAATAGA 660
 ATGTTTTGTA GAAAGTAATA AATTGATGGC GTTAACAAGT TTAATACTTG TAAAATAATC 720
 TTTAAAATTC ATACTCAAAT ATCGAACAAA AAGATTAGAA CAATATGATT CAGAGATGTT 780
 GCAATGCATC GCTATATCTT TTAAAGACAA GTGCGCATCA ATATkGTCAT GAATATAGTT 840
 TACACAGTCT ATGAACACCG GATTGCTTAA AGCAATATTT GGTAATATT CATGATCTAT 900
 TCTTATAAAA GCTTCTTTAA GTAAAGTGTC TATnATACTT TGACCT 946

(2) INFORMATION FOR SEQ ID NO: 738:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 738:

5 TTTATTTGAT GATGTACTAA AGCAAGATGA AAGATTTGTT ATTGTAGTTC AAGCATTAGA 60
 AGAAAAGAAC GGACAATTAT TAAAGAGAAC TTTAAGGGAA TATCCCGGTT TAAACCATAA 120
 ACAAATGAAT GATTTATTTA TGCACCTAAA GGAATTATTT TCCGAAGAAT CATTTGCTGA 180
 10 AAACCAATCA GCGTTCAGTA TTACAGTTTA TACAAATTTA GATTATACTG CTGACCAAAT 240
 ATATGCTCAT GTAAACGTT TCAGAGGTAA GCATGACTGG ACACAAACAG CTAAATAAAA 300
 TCAAGATGCT TGCACTGgAT TkTAACmATT AATGtTAGaT gTcmCyGcAT TTTTaATGCC 360
 15 ATTCGGTAAT TAATCCGGAT GGTCCAATTT AATTAAC 397

(2) INFORMATION FOR SEQ ID NO: 739:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 915 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 739:

CAAGGTATTT CATTTGACCC ATTCACACCA TTAATAGTGG CAGCAGCATT ATACTTTGTC 60
 30 TTAACATTTG TACTTACACG CATTATGAAC ATGATTGAAG GGAGATTGAA TGCCAGTGAT 120
 TAAAATAAAC AATCTTAATA AAGTTTTTGG AGATAATGAA GTTTTAAAs ATATCAATCT 180
 TGAAATCAAT CAAGGGGAAG TAGTAGCAAT AATAGGTCCA TCTGGTAGTG GTAAAAGTAC 240
 35 ATTGTTAAGA TGTATGAATT TATTAGAAGT ACCTACTAAA GGTCAAGTGA TTTTGAAGG 300
 CAATGACTTA ACGGAAAAG GGACACAAGT AGATAAACTA CGTCAAAAAA TGGGTATGGT 360
 ATTTCAAAAC TTCAACCTAT TTCCACATAA AAAAGTTGTC GATAATATTA TTTTAGCTCC 420
 40 TAAATTATTA AAGAAAGATA ATAACGATGA ATTACATAAG GAAGCATTGT CGTTATTAGA 480
 TAAAGTGGGA TTAAAAGAAA AAGCAGATGT ATATCCGAAT CAATTATCAG GTGGTCAAAA 540
 45 ACAAAGGGTA GCAATTGCAA GAGCTTTAGC AATGCATCCA GATGTTATTT TATTCGATGA 600
 ACCAACTTCA GCATTAGATC CTGAGGTAGT TGGTGATGTA TTAAAAGTAA TGAAAGACCT 660
 AGCCAAAGAA GGTATGACCA TGGTGTTGT GACACATGAA ATGGGATTTG CCAAAGATGT 720
 50 AAGTGACAAA GTCATATTTA TGGCAGATGG CGTTGTCGTA GAGTCAGGCA CnCCAGTCGA 780
 AATATTTGAA CAACCGCAAC ATGAAAGAAC ACAAATTTT TTAGCAAGAG TATTATAACA 840
 55 ACCTAACGAG GCTTGAATAT ATGATACGCA CCACAAAGTT ATATCATATA TTCAAGCTTT 900

(2) INFORMATION FOR SEQ ID NO: 740:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 435 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 740:

```

AGTAAATGCG ATGTTGCATG GAGGCTATAG CAATGAGCAA TAAAGCTTGG GCGGGTAGAT      60
TTGAAGTACA ACCTGAAGAG TGGGTTGACG ACTTTAACGC ATCCATTACT TTTGATCAAA      120
CGCTCATAGA TCAAGATATC GAAGGCAGCA TTGCACATGC AACTATGCTT GCGAATCAAG      180
GCATTATTAG TCAACAAGAC AGCGAACAAA TTATACAAGG ACTAAAATCT ATTCAACATG      240
ATTATCATCA AGATCAAATT CAATTTAGTG CATCATTAGA AGATATTCAT TTAAATATTG      300
AACATGgAAT TAATTAAACG TATCGGTGgA TGCTGGTGGT AAGTTGCATA CTGGACGCgT      360
AGtAAACGAT CAAGTTGcAA CAGACATGCA CTTGTACACT GAAGAAACAA GTGCAAGnnn      420
TCATCGCATT GGATG                                          435

```

(2) INFORMATION FOR SEQ ID NO: 741:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1198 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 741:

```

AATATCATAA CCAAAGCTTG TCATTACTTA TATGTAAAGA TTAGAGGCTA ATTACATAAT      60
CCTAACTTAT TTAACAGGAA GTCATGAAAA CTATATTACA ACTAAAAGTA TAACTTTCCA      120
TCACCAAAGT CAGAATTTAT TGAAACACTC AACTGATAAT ATAGAAGGCT AATTATAAAA      180
ATACCTGGCG ATTAATAAAC TGTTCGCTT TTTTAGCACT TAATATATTT AAAAAGACAA      240
ATAGAGCCCT TTCAGGCTCT ATTTnCTGTA TCCTTAACAA CAGAAGCTGT TAAAACTACC      300
CGTCTTTTcA CAaCCAAaGc TaCAAAAAaag aTgGCTAGTT ATTCTTTCGT CACCCGCTGA      360
ATCATTAGTG TTGTTATTTT CTTTAACTTG CACGTCTAAA TCAAGAAcTT TTTCCAAaGA      420
TaAAACCCCC TATTTAAAAG TTGAAGTTAA GACCCcTTCA ATTGTCATAT TAATCATTAC      480
CACACAATAA ATCAATGAAT TTAATAATAA TTATTTGTTA TTAATTTGTA AATCGTCATT      540

```

TGAATATTTT TTGTTTAACT CTAAATTTT GGGTTTAACT TCCTCTATTT CTAACAGTTA 660
 TACTCCAGGA TTAGTTTCTT TAGAATCCGT ACCGATTAAT TTAACAGTAT GGTTTTGCCA 720
 5 GTCAACTTCA TAAGTAGATG TAAACGTTAC TGTATTTTGA TTTTTGTAGT TATTCCAAC 780
 CCAGTGTAGT CGATTCCATT GATTAGTATA TCTATCCATT TCTCTTTGGT AAGTTACTTT 840
 10 GATTTTAGAT TTTTTGTAT CATTTTGTGT ATGAGAAAGT ACGCTTATAA ATTCTGGGTT 900
 AAAGTTACCA CGCGCCAATA AAGGCATTTG ATGTGTTGGC AAGAAATTTT GACCAGCATT 960
 TGAACTACTT TGTCTACCAC CTAAAAACAG TTCATTACCA TATGTTGGGT CATAACTATC 1020
 15 TCTACCATAT GGTCCCCAAC CATTATTCAT AATTTTGTGC GCCTCAACAC CCCAGCCAAT 1080
 TGATTTATGA TTTGTTTTTC TATCAATCGT AGTTCTGTAA CTTTCTTGTT TATAATTTAT 1140
 CGTTTCTGAA AATGATTTTG ATCCATTTAA TCCACCTGAT AAGCCATTAG ATATATTA 1198

(2) INFORMATION FOR SEQ ID NO: 742:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 818 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 742:

GTGGCATATn CAACTACTA CTGAAAGATC AATTACATAC GTACAGATTT TGCAAAATGA 60
 TCAAGATTAT CCAACAATTT TTAGATGAGa CAAAACGTAA ATCTGTAATT AATTcAGATG 120
 35 TTAATGTAAC GGTAAAAGAT AGAATAATGA CTTTATcAAC GTGCGAAGAT GCATATAGTG 180
 AAACGACGAA AAGAATTGtT GTTGTGCGAA AAATAATTAA GGTAAGTTAA ACAGAAAmGA 240
 GGATAATTAT GAAATTTATG GCAGAAAATA GGCTGACGTT AACaAAAGGA ACAGCAAAAG 300
 40 ATATTATAGA ACGATTTTAC ACGAGACATG GGATTGAAAC ATTAGAAGGC TTTGATGGCA 360
 TGTTTGTTAC ACAAACTTTA GAACAAGAAG ATTTTGATGA AGTGAAAATT TTAACAGTTT 420
 45 GGAAATCAAA GCAAGCTTTT ACGGATTGGT TAAAATCTGA TGTCTTTAAA GCAGCGCATA 480
 AACATGTTAG AAGTAAAAAT GAAGATGAAA GTAGCCCGAT TATAAATAAC AAAGTAATTa 540
 CATATGATAT AGGCTATAGT TACATGAAAT AATTGAATTT AAAGAGGTTG CAGTACTTGT 600
 50 TATGACAATT GGTATTGTTA GCTTTTGTAa CGACCATACA TGATACCGAT GATGGTCGTT 660
 TTTTTAATGa ACACAAACAT GCTAACAACA AATTGCTAAA ACATAGTTTg aTTTGAATGT 720
 GgCTTTGGAA AATAATCATC CAtTATATCa TTtTGGATTn TATTTTGGAA GAAGAAAAGA 780

(2) INFORMATION FOR SEQ ID NO: 743:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1209 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 743:

5	TGCTGGTAAA TCTTCCCTTA TCAAGTCTTT AATTGGTGAA TTTAATGCTA CCGGTACTAA	60
15	ATTGTTATAT AACAAACCTA TACAACAACA ACTGCAACAT ATTACATATA TTCCACAAAA	120
	AGCACATATT GATTTAGATT TTCCTATAAG TGTGGAACAA GTGATTTTAT CAGGTTGCTA	180
	CAAAGAAATT GGATGGTTTA GACGACCTAA TAAATCAGCA AGGGATAAAC TCAAACAGTT	240
20	ATTAAGCGAT TTAGAATTAG AATCTTTACG TCATCGACAA ATTTCAGAAT TAAGTGGTGG	300
	ACAATTACAA CGTGTGCTAG TAGCAAGAGC ATTGATGTCC GAAAGTGAAG TTTATTTTCT	360
	AGATGAGCCG TTTGTCGGAA TTGATTTTAG TAGCGAAAAA tTAATCATGA CAAAAATcGA	420
25	GAACTTAAaA CAACAAGGAA AACTTATTcT TATCATCCAC CATGATCTAT CAAAAGCAAA	480
	GCAATACTTT GATCGCATTa TTCTATTAAA TCAAACATTA CGATACTTTG GTGATAGTGA	540
30	AGAGGCTATG AGTGTCACTC GCTTAAACGA AACATTTATG AGTAGCACTG ACTGTAGTGA	600
	CCCTAGTCAA AGGAGCAATA TAACATGTTA GAGTTTGTG AACATTTATT TACATATCAA	660
	TTCTTGAATC GAGCATTGAT AACTTCAATT ATTGTAGGGA TAGTTTGTGG TACAGTTGGT	720
35	AGTTTAATTG TATTACGTGG TCTTTCATTA ATGGGAGATG CAATGAGTCA CGCAGTATTA	780
	CCTGGTGTG CCCTATCATT CTTATTTGGT ATTCCAATGT TTGTAGGCGC ACTTATAACT	840
	GGTATGATCG CAAGTATTTT TATCGGTTAT ATCACATCTA GTAGTAAAAC GAAACCTGAC	900
40	GCCGCTATTG GAATTAGTTT TACCGCTTTC TTAGCTTCGG GGATTATTAT TATTAGCTTA	960
	ATAAACACTA CGACAGATTT GTACCATATT CTGTTTGGTA ATTTACTGGC AATCACAAAT	1020
45	AGTGCAATTT TAACAACAT TGTGATTGGT TCAATCGTCC TTATTCTAAT CATTATTTTC	1080
	TATCGTCCAT TAATGATTTT TACATTCGAT CCAACGTTTA GTAGAATGAG TGGTCTAAAT	1140
	ACGACGTTAT TACATTACTT TGTGATGTTG tTACTCTCAT TAGnAACAGT AGCAAGTAtT	1200
50	CmAACGGTA	1209

(2) INFORMATION FOR SEQ ID NO: 744:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 675 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 744:

	CCGCCTTTTCG TAAACCTAAT ATAATACGAA ATTTTCGTAT TGTCAACATT AAATACGTTT	60
10	TTTTCGTAAA AAACCTTACT ATGATATGAA AATTTCGTAT AATAAGAAAA AAGGAGGTAA	120
	GTAATATGAA CAAAGAmAGm AATATTATTA TAGCCAAAAm CATTAGAAAA TTTCTCAACG	180
	ATTCAAATAT GTCTCAAAAG AAACCTTGCTG AaCTCATTAA CATAAAACCA TCTACTTTAA	240
15	GCGATTATTT AAATTTACGT TCCAACCCCT CTCACGGCGT TATACAAAGG ATAGCTGATG	300
	TTTTCGAGGT TGGTAAAAGC GACATAGATA CTACATACAA AGACGATAAC GACATCACTT	360
	CCATATACAA CAAACTCACA CCTCCCCGCC AAGAAAACGT ACTTAACTAT GCAAATGAAC	420
20	AATTGGAAGA ACAGAATTCT AAAGGAGATA ACGTTGTAGA TATTAATTCA TATAAACAGG	480
	AGAAAACCTCC AGTTAACGTC AATGGTTGCG TCTCTGCTGG TGTAGGAGAA CGTTTACACG	540
25	ATGAAACGCT ATTTACTGAA ATGGTTAAAG GACCTATCCC CACACACGAT TTAGCGTTAA	600
	AAGTAAATGG TGATTCTATG GfACCTATGT TTAAAGATGG CGAAATCATA TTTGTGGAGA	660
	AAACTCACAA TnTGA	675

(2) INFORMATION FOR SEQ ID NO: 745:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 924 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 745:

40	CAGGTGTCAT AGTTGAACTT GGtGGTTTTG GACTAATCGT TAAATCATCA ATTTGTGTCA	60
	TCCAAGGTTG GAATTTACTG CCTTGTGTCT CCAGTGCCT GAAAATACCT TTATCTTGAA	120
45	ATAAGTCAAC TAACTCTTGG ATACCTTTAA TTAATGCTGG GTTACCACCA GAAATTGTAA	180
	CGTGGTTAAA TAAATCGCCA CCAATTCGTT TTAATTCATC ATAAATTTCT TCAGCGGTCA	240
	TGAGTTTTAT ATCGCCTTTA GCACTACCAT CCCAAGTAAA TGCAGAATCA CACCAGCTAC	300
50	AGgATaATCA CATCCAGCTG TTCTCACAAA CATCGTTTTT CTACCGATTA CTCGACCTTC	360
	ACCCTGAATG GTTGGACCGA ATATTTTCGAG TACAGGAATT TTAGCCATTA GTTACACCTG	420
55	TTCTTTGGT CTAAATACGA CATAACTTGT TGGTGTCTTCT CTTACAAATA CTTGAATACA	480

TACGATTTCA GTTGAAGGGA TTTTGTTTTT AAAAGCAGGT AAGTTATTTA ACAGTTGATG 600
 GTCAAATTTA CCGTGTATCA TCTTTTTTCAA ATGGCTAAAG TTCACTAAGA AGCCAGTGTC 660
 5 ATCTAGTTTA TCACCGACAA TTGTTAAaya CAAAGTAAGT ATGACCATGG ACATTTTGAC 720
 AAATACCTGC TTCTTCACAT GGAATGTGAT GTGCAGCCGA AAAATTAAAA TCTTTATTTA 780
 10 ATTGGAATTG ATATGGATGC GTTGACTAG GATAGATTG TTGTAACATT TTAAAGCGCT 840
 CCTTACTTT CAAGATATTG ATTTAGTCCA CGTTGACGTA AATGACAAGC TGGACATTCA 900
 CCACAGCCAT CCCCAATGAT ACCG 924

15 (2) INFORMATION FOR SEQ ID NO: 746:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 971 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 746:

ATCTTTTGGG TATTTTCTAT ATTAGCTGTT GTTTTCTTAG TATTATTTAG TTTTGCTGTT 60
 GGTGCATCAA ATGTACCAAT GATGATTTTA ACATTTATAT TACTCGTTGC AACCTTTGGA 120
 30 ATTGGATTTA CTACAAAGAA AAAATATCGA GAAAACGATT GGCTATAAGT TATATCGACA 180
 ATAACACGCG TTGATTAAAG ACCACACATG TTCATGATGA ACATTTAACT GGTTTTATAT 240
 TGGCGCGTGT TTTTATTATT TAAGGTATAA AAGCGTAATA TGTTACCTAA TTGTTAGGAC 300
 35 TACATATTAC GCTTTTGGTA ATTATTTTGA TTTTCTTTTA AATCGTTGTT CATGTAATTG 360
 TATTAACGCA TCATGGTGGG GCTCTAAaTA ATCATtGCTA ATTcTTGATG GAGACcmCGA 420
 TtCCAACTA rGATkGAATT kGGACcmCTA AtkGTTAATG GTcctCCTAG TAAATTGGAA 480
 40 GCTTGTCAT TaCTTccATA TAAATAACCA ATCCGCCAGC AAAtCCCCAT GGTkGAAGTC 540
 TTGGCGTCAT ATATGCTTCT AAaTTACCTG TAGCAACTGA AACGATTTCA AGCGCTGCAC 600
 TACCATATGC CCTTGCACTT CTAGAATCAT TAACAATTTT TTTAAAGATT TCTCCTAAAA 660
 45 TTGGTTTAGT TAACCAGTTC GGATTGATCC CAATAATGCT TTGTCTTAGA TTAGAATCAT 720
 TCAATGGTTT CAAGGGTTGG CTACCACGAT ATGCACCTTC CCCTACTTTA GCATGATATA 780
 50 AGACATCAGC CATAACATCA TATACAAAAC CTGCATAAGG TTTACCATCG ATATAAATAC 840
 CAATTGAAAT TGCGAAATTT TCTTGTGAT GAACAAAATT CAATGTACCG TCTATTGGGT 900
 CAACAATCCA TACCGTACCT TTGGAAGTAT CGATGTCATG ACCATGCCCT TCTTCACCTA 960

55

(2) INFORMATION FOR SEQ ID NO: 747:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 865 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 747:

```

TTCTCGTCAC TGTACGTTGT ATTGCGGCAT CATCCATTAT GATACGTTCA GACATATTCT      60
CACCCTTTCA AAATTTAGAA ACAAAAAACC TCATGTCTTT AATATCGACA TGAGGTTTAG      120
TATATACAAT CAGCCTATTT ACATAGAAGA ATCGTAGTAT TACTACTTAT TCTTGCATAT      180
GTGAAAGTGT TATGATCATT TTAGTATATA TCTATATTTT CATGTCTTTG ACGTCTCTCT      240
GGACGCTCAA TTAAAGACTC TTTATGTTGT ACTTAAACTG TTATAAGGAT AATACTAATC      300
AGGCGCTTTG TCAATCTATA TTGAAAAATT AACCATACTC AGTAAAAATG TTAACCTTAA      360
TCTATTTTAT AATGTGTTTT CATTTAAAAA AAACGAAAAC GATAAACACT TATAGTTTAC      420
ATAATAATGT TATCGTAATT ATAGCAGGAC ACAATTATGC ATCTCTTTTT CGAATTGTAT      480
CTAATAAATC TTCAAAGTCT TGTGTAATT CAGCATGTCT TTCAATATAT TCACCTGTTA      540
CTGGATGTTT GAATCCAATA AGTCCAGCAT GTAGAGCTTG ACCACCAATA TCCAATGTCT      600
TTTTCGGTCC AACTTTTGGT TCACCAACTA ATGGGAAGCC AATATATTTT ATGTGAACAC      660
GGATTTGaTG CGTACGTCCT GTTTCAAGTT GACATTCAAC AAGCGTATAA TCTTTAAAT      720
GTTCTAGTAC GTTAAAATGT GTCACTGCTT CCTTACCATC ATCAnCAACA GCCATAGATT      780
GACGATCATT TTTGTTTCTA CCAATTGGCG CATCGATTGT ACCGTAATCA TGAGGAATAT      840
TCCCGTGGAA CTAAAnCGAT ATAGT      865

```

(2) INFORMATION FOR SEQ ID NO: 748:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 818 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 748:

```

ATAACATTGT GGAGCCTATG ACATTGATTT ATGTCTCGGT CTCAAATGTT CTTGTAAAAA      60
AACTAAGTAT AGTTGAATTT TAAGTTAAAA AAATTGTTTT TGCAAAAATA AAATAATCCA      120

```

5 TTTTAGGTTT TCGCCTCTA AAAATCCTTG CGaTATTAGA GCGATGTCTA ATTATCAATA 240
 TGATTGAAAC TAAGAAACTA ACGACTAATA AAATATAGTC TTGAATGATA AGCGAGCCAA 300
 10 TCACACAGCA AATTGCTGCA ACGATACTTG CTAAAGAAAC ATATTTAAAA ATCTTCAATA 360
 CAATAAAGAA GATAATTGCA AGTATTAGTA AAAGTATCGG ATTGACTCCC AAGACGACAC 420
 CTGCACTAGT TGCAACAGCT TTGCCACCTT GGAATTTTAA ATAAACAGGA TAAACGTGTC 480
 CAAGTATAGC GAATAAGCCA ACAATTAAAC CATTTGTAAA AAAAGTACTA ATAGGGCCAT 540
 CTGCGTGAAC TTGTAACCAT AAAGGGAAGA AAACAGTTAT GAACCCTTTG AAAATATCTA 600
 15 GAAATGTTAC CAAGAATCCT GCAGGACGAC CTAATACTCT AAAGCTATTA GTAGCGCCAG 660
 TATTACCACT ACCAAATTGT CTAATATCTT TTTTGAAAAA TAATTTTCCA ATTACGAATC 720
 CACTTGGGAA AGCGCCGATA AGATAACTTA GTAGTAACAT GACGATTATC ATCATAAATA 780
 20 TTACACATCC TTTAATATCT TAGGACTATT TTATCATA 818

(2) INFORMATION FOR SEQ ID NO: 749:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 971 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 749:

ATATCCAATa ACTGTTcATT GTCTTGATAT TCCTTAAATG TTTcATAATG CTCATTAGAA 60
 35 ATGACAATTT CAAGCAACCA ACCCGTGCCA CTTTcCTCTT TATTaATTAT TACGCCATCT 120
 TCTaATTCAT ATTCTATAAT TCTGCCGTGT TCATTcACAA TTTGAAACCT TACTGCTTTA 180
 AAAGTTTTCA TATTCCCACA TCCCATTAGT TTAATaTTTA TGATTTTTGA ATTATTAGA 240
 40 AGATTTAAGG TCATTTcATG ACTaATTTCT TTAAATATAA CAAATAACGA GATGCTATAG 300
 CTATTaATTA TATAAAATTT GcCATTTTga CCATTTTTAA GCTAGTAATT AAGATATAGG 360
 cATAGGAGGT GAATCACACA CTACATGCTA AATAAAATCG TAATTGTCGG GAGcTGACGA 420
 45 AAGACGCACA AATATTTGAA AAGGAGGATA GAAAAATTGC AACGTTTTGT GTTGCAACGC 480
 ACCGAAATTA TAAAGATGAA AATGGAGAAA TCGTCTGTGA TTACTTATTC TGTAAGCAT 540
 50 TTGGCAAGTT AGCTTCTAAT ATAGAAAAAT ATACTAATCA AGGTACATTG GTTGGTATAA 600
 CTGGTCAAAT GAGATCAAGA AAGTATGATA AAGACGGACA AACACACTTT GTCACTGAAT 660
 TATATGTTGA AACAATAAAA TTTATGTCCC CTAAATCCCA AAATAATGAA ATTCTCTCAG 720

TCCAATTATC CTAAACATCC TTAATATACA TTTAATCCAT CCTAAATTTT TTATAGATCT 840
 AATTAACCTG AAATATACTC ACTTAAAAAC GACTTACATG ACCTAGTACT GTTGGTAAGG 900
 5 TCGnCTTTTG GTGTAGTTTT CCTAATAGAn AATAGACnTA GGTGCCCCAA CCGATTACAT 960
 AAAAAATTGC C 971

(2) INFORMATION FOR SEQ ID NO: 750:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 832 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 750:

20 ATAGATTACT CTTTTGATCT GTAGATCCTA CTGTAACGAC ATTGTCCATA GATGCAGGAA 60
 CATCTTTCAC TTCGCCATTA CCTTGATATT CACGCTGTAA TTTTAGTTTC TGTGTGTCAT 120
 TGACATCAAT ACCATCATT A CCAGCTGCAG CAACAACGAT AGATTTTTTC TTCTTGGCGT 180
 25 AATTGATTGC TTTCTGTAAC GCATCGTATT CTACTTTTTT ATCTTTTCTA AATGTTTGAT 240
 GGTCATTTTT GTCCAAAATA ATATACTAC CAACACTAAT ATTAATGACT TGAATTCCAT 300
 CATTTGCAGC TTGAACAATC GCTTTTGATA CCCAAAGCAG TTCTGTTTTT TTACTIONAA 360
 30 ACACGCGATA CATTGTAAAT TTGTTATTCG GTGCAACACC TATTAACCTA CCATTAGCAC 420
 TCGTTTGACC CGACACCATC GTGCCATGTC CTTTCCTATC ATTGACATCG TGAACATCAC 480
 CTGTTTCCTC CGGTTTCAGTA CCTCTAAAAC CGTTTAAAGG TACTAAATTT TTAGAATCAG 540
 35 TCGAGAAATT ATTTTTCAAA TCGTCATGGT TTTTCATCAC ACCTGTATCT ATGATTGCTA 600
 TTTTGGTGTT AGCATGTTTT GGCAAATCAT CATACGATGC ACCATTATTG GTTATTTTAT 660
 40 TCATATCCCA TTGTCTTGAA AATAATGACT CATTGTATGT TCTGTCTATT GTTTTCTCGC 720
 TAGTAATACA AGTTGAACAT GTGGCATTGA TATATTAAAT ATCATTMTTA TAGTTTGCTA 780
 AAGCATTAGC ATGCATTTTC GTCATTTTAA TCTGTGCCAC ATGGAATTCC GG 832

(2) INFORMATION FOR SEQ ID NO: 751:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 475 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5
 10
 15

GTAATATCAA TCTCTTCATA AGCTGAATTA TTTTCATGCA CTTCTTGATG TGATGATTTG 60
 TCACGAACCG CTACAACTAA CATTTTATCG TCTAAAATAA GTTGTTTATA TTTTCTAAT 120
 TCATCAGGCG CTAAGTTGTA GCGTGATAAA ACTGCATGTT CACCATCTTC TCCTGTTAAC 180
 AGTTTAGTCA TTCTATCACT AAATGTTCCA CTTGTTGAGA TAAGGGAGAT TTCAGAGTCG 240
 TGTAAGTCAT TTAGGTGTAA TTTACTTTTA CTAATAATTG TTAGCTCTGA TTCTAAATAA 300
 CCTTCAGATT TCTTTTGATT GATTACGTTG TATAATTGCG CAGTGTCATT TACTACAGTA 360
 ATATCTGCCA TAGTTGTCGC CCCTTTAAAA ATTTGTTTAT tTAATCTTTT ACCCTTCTTA 420
 TnATAAAGTA AAACCCTTAC ATTATTAAGT nATAAGTCTT CATTCGCATT AAACG 475

(2) INFORMATION FOR SEQ ID NO: 752:

20

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1019 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 752:

30
 35
 40
 45
 50
 55

GCGACACTTG TGAGTTCTCC AGCAGCAAAC GCGTTATCTT CAAAGGCTAT GGACAATCAT 60
 CCACAACAAA CGCAGTCAAG CAAACAGCAA ACACCTAAGA TTCAAAAAGG CGGTAACCTT 120
 AAACCATTAG AACACGTGA ACACGCAAAT GTTATATTAC CAAATAACGA TCGTCACCAA 180
 ATCACAGATA CAACGAATGG TyATTATGCA CCCGTAACCT ATATTCAAGT TGAAGCACCT 240
 ACTGGTACAT TTATTGCTTC CGGTGTAGTT GTAGGTAAAG ATACTCTTTT AACAAATAAA 300
 CACGTCGTAG ATGCTACGCA CGGTGATCCT CATGCTTTAA AAGCATTCCT TTCTGCAATT 360
 AACCAAGACA ATTATCCAAA TGGTGTTTTC ACTGCTGAAC AATCACTAA ATATTCAGGC 420
 GAAGGTGATT TAGCAATAGT TAAATTCTCC CCTAATGAGC AAAACAAACA TATTGGTGAA 480
 GTAGTTAAAC CAGCAACAAT GAGTAATAAT GCTGAAACAC AAGTTAACCA AATATTACT 540
 GTAACAGGAT ATCCTGGTGA TAAACCTGTA GCAACAATGT GGGAAAGTAA AGGAAAAATC 600
 ACTTACCTCA AAGGCGAACT ATGCAATATG ATTAAAGTAC AACTGGTGGT AATTGAGTT 660
 CACCTGTATT TAATGAAAAA AATGAAGTGA TCGGaATTCA TTGGGGCGGT GTACCAAATG 720
 AATTTaATGG TCGGTATTIT ATTAATGAAA ATGTACGCAA CTTCTTAAAA CAAAATATTG 780
 AAGATATCCA TTTTGCCAAC GATGACCAAC CTAATAACCC AGATAATCCT GATAACCCTA 840
 ACAATCCTGA TAACCCTAAC AACCCAGATG AACCAAATAA CCCnGACAAC CCTAACAACC 900

CAATTAAGTT CGTCTAGATA AAATAGGAAG CAAAGTCTTA GCAACGTAAA ATATTTTGA

1019

(2) INFORMATION FOR SEQ ID NO: 753:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 533 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 753:

15	ATTCTTATCC CAGGTAGGTG nTTTACTTCT TATATTAGGT GCAATAACTC ATGCCACTCA	60
	TCTTGTGGCA CAACATGAaT CGCATTITCA AAAATTGCAT GTCCTGCTTG ATCAGTCAAA	120
	CCGGCAGCAT GATCAAAATG CATATGTGTC ATTAGCACAT AATCAATATC CTTTGGCGTT	180
20	AAATTATAAT TTGCCAAATC AGCAATTATA TGACTTTCTT CATCTACTCC AAAATTACGT	240
	AATTGCTTTT CAGATAATTT ACCATTACCA ATACCTGCAT CTATAATCAA ATTATATTGA	300
	GCCGTTTGAA TCAAAATTGG ATGTGTCGGT AAATTGATTT GATTTGCTTC ATTTGCATTG	360
25	TATTGCTTTG ACCACAACGG CTTGGAACA ACACCAACA TTGCACCGCC ATCCATTTTT	420
	GTATTGCCAC CATTAGaTA ATGAATAGaT ATATCCCCGA TTTTCATAAC ATCACCTATT	480
30	CTTCTGTAT TCGTTATTAC AATTAGCATG TGTATTATAC ATCATATCAC TAT	533

(2) INFORMATION FOR SEQ ID NO: 754:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1225 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 754:

40	ATCCCCTGCA AGTGATAAAA TTATTGAAAA AGGCGACATG ATTACATTAG ATTTTGGCGC	60
	GTATTATAAC GGCTATTGTT CAGATATTAC TAGAACATTT GCTATTGGAG AACCAGATCC	120
45	TAAACTGAAA GAnATATATC AnATAGTACT TGAATCTCAA ATGAAAGCAA TTAATGAGAt	180
	TAGACCTGGC ATGACTGGTG CAGAAGCTGA TGCCATTTCA AGAACTATt TAGAGTCAAA	240
	AGGGTATGGA AAAGAATTTG GACATTCACT AGGACATGGT ATTGGTTTAG AAATCCATGA	300
50	AGGGCCAATG CTGGCTCGTA CGATACaAGA TAAACTTCaA GTTAACAACT GTGTTACAGT	360
	AGAACCTGGT GTTTATATAG AAGGTTTGGG CGGTATAAGA ATAGAAGATG ATATTTTAAT	420

AGCGTGTAAT ATGAGGAGGA AACTGAATGA TTTCGGTTAA TGATTTTAAA ACAGGTTTAA 540
 CAATTTCTGT TGATAACGCT ATTTGGAAAG TTATAGACTT CCAACATGTA AAGCCTGGTA 600
 5 AAGGTTTCAGC ATTTCGTTCTG TCAAAATTAC GTAATTTAAG AACTGGTGCA ATTCAAGAGA 660
 AAACGTTTAG AGCTGGTGAA AAAGTTGAAC CAGCAATGAT TGAAAATCGT CGCATGCAAT 720
 ATTTATATGC TGACGGAGAT AATCATGTAT TTATGGATAA TGAAAGCTTT GAACAAACAG 780
 10 AACTTTCAAG TGATTACTTA AAAGAAGAAT TGAATTACTT AAAAGAAGGT ATGGAAGTAC 840
 AAATTCAAAC ATACGAAGGT GAAGTATCG GTGTTGAATT ACCTAAACT GTTGAATTAA 900
 CAGTAACTGA AACAGAACCT GGTATTAAAG GTGATACTGC AACTGGTGCC ACTAATCGG 960
 15 CAACTGTTGA AACTGGTTAT ACATTAAATG TACCTTTATT TGTAACGAA GGTGACGTTT 1020
 TAATTATCAA CACTGGTGAT GGAAGCTACA TTTCAAGAGG ATAATCTCTA ATTTGTTAAC 1080
 20 AAATAGCTTG TATTCATAT ACTGATTTAA CGTAAGAAAT TCTAAATAAG TCTCATAAAG 1140
 CTATTGCCTA AAATGATTAT AGGTTATATG CTGATATGAG GCTTTTTATT TTTAAATAA 1200
 TTTTAAAT ATAAACGAA TCGTC 1225

25 (2) INFORMATION FOR SEQ ID NO: 755:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1265 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 755:

35 ACAACAAAT CACTTAGTTA AAATTGACAG TGCATTATAT TCCGATAAAT TATTCAACAT 60
 TGTAAGAAAT AGAATTGATA CACCAGATAT TGGCTATAAT ACAGTGTTT TATCAGAAAA 120
 40 ATGAGTGTAG AAATTAAAGG GATACCTGAT GTGTTGAATA AATTAGAATC GGTATACGGT 180
 AAACAAGCAA TGCAGGCTAA GAGTGATAAA GCTTTAAATG AAGCATCTGA ATTTTTTATA 240
 AAGGCTTTAA AGAAAGAGTT CGAGAGCTTT AAAGATACGG GTGCCAGTAT AGAAGAAATG 300
 45 ACTAAATCTA AGCCTTATAC AAAAGTTGGT AGTCAAGAAA GGGCTGTTTT AATTGAATGG 360
 GTAGGTCCTA TGAATCGCAA AACATTATT CACTTGAATG AACATGGTTA TACAAGAGAT 420
 GGAAAAAAT ATACACCAAG AGGTTTTTGA GTTATTGCAA AACATTAGC TGCTAGCGAA 480
 50 CGTAAGTATA GAGAAATTAT AAAAAGGAG TTGGCCAGAT AAATGAATAT ATTAAACACC 540
 ATAAAGGAA TTTTATTATC TGATGCAGAG CTCAAAACAC ATATAAATTC TAGAATATAC 600

55

TATGATTTGC CTTCAGACTT TATGTCTGAT AAATATCTCA GTGAAGAATA CTTAATTCAA 720
 ATAGATGTAG AATCTTCAAA TAATCAGAAA ACAATTGATA TAACAAAACG AATAAGATAC 780
 5 CTGTTATATC AACAAAATTT AATTCAAGCA TCTAGTCAGT TAGATGCTTA TTTTGAAGAA 840
 ACTAAACGTT ATGTGATGTC GAGACGATAT CAAGGCATAC CCAAAAATAT ATATTATAAA 900
 AATCAGCGCA TCGAATAGGT GTGCTTTTTA ATTTTAAAGG AGGAAATAAG CAATGGCAGA 960
 10 AGGACAAGGT TCTTATAAAG TAGGTTTTAA AAGATTATAC GTTGGAGTTT TTAACCCAGA 1020
 AGCAACAAAA GTAGTTAAAC GCATGACATG GGAAGATGAA AAAGGTGGTA CAGTTGACCT 1080
 15 AAATATCACA GGTTTAGCAC CAGATTTAGT AGATATGTTT GCATCTAACA AACGTGTATG 1140
 GATGaAAAAA CAAGGTACTA ATGAaAGTTA AGTCcTGaCa TGaGTaTTTT CaATATTCCa 1200
 AGTGaTGaTT TAAaCACaGT TaTTGGaCGT ACTAAAGATA AAAATGGGTA CATCTTGGGT 1260
 20 AGGAG 1265

(2) INFORMATION FOR SEQ ID NO: 756:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 1111 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 756:

CATATACCCG GGTCCAAGA ATCCTCTATC CATATGGTCC CCCAATTTTG ATTCCACCAT 60
 35 ATTGTTAAAT ATCACTTTCC CAGCCTACTT TTTTATnCAG TTGGGCTCTC TAAaATTGTT 120
 TTGAAATCCA GGTGAACAT ATTTCAGTG TATGACCAAT CGAAACATTT GCACCAATAA 180
 GGCCGCCAAT TTTTCCTGTA TCATCACCAG TAACATTACC GTTGAATCCA TAAGTTAAAG 240
 40 TACTCATATA CTCTTTTGTA TCAATCGAAT TTCTTGATA GTAATCAGAT ATTTGAGCTA 300
 CTTCAATTATC AGGTAGTTGC AACTGTACCT TAAAGGCTGA AGGCCAGGCT AAACCACTTT 360
 TGTTAGCACC TTCTTCGCTA TAAACTCTAT ATTGACCAGC AATGGTACCT TTCGTTCTAA 420
 45 TAACTAGCAG TTTTTTATTA TGATTTTTAT CATCGATAAA ACTATAAAAT ACTTTTTTGT 480
 GCATGCCATT TTCTTTATCA TAAGTGAATA AATCACCTGT TTTTACTGTA GTATTGCTTC 540
 CAATATCTGT AGTACCGGTT TTAATATTAA TATCAGAATC TGCGGCATTA GCGACAGGAT 600
 50 TCATTAATAT GGAACCTAGC AATAGTGTTG TTGTTACTGA GCTGACTATA CGTGTTTTCA 660
 TTTTCATCAT CCTTCTATTT TTTAAACGA TTTGAGGAAA CAATAATCAA TATGTCAATT 720

55

GTAGCAACTG ATAAATTACT GAGTGATGAT GAGTGATTAT TTAAAGAATA TGTTTTTAAC 840
 TTTTATTTAA AATTTGAAAG GAAGCATTTT AATTTGAGG GTTAGTCAA GTTGAATAAA 900
 5 TTCTTTATGA AACAAGGAAA AGACATAGCT AATTTATTG ATTAATTTCT TTAAACTAA 960
 TGATTTGTTT GATTTAAAAA TGTAATCGAT TACAATATAA AAATACAAAT ATCTTAGAAT 1020
 TAAATCAATT AATTAATAT TAAATAAAAA TTAATATAT ATTAAGTAGT GTAAATTAAT 1080
 10 AAATAGAAAT AGAGAAAAAG GGTATTAATT A 1111

(2) INFORMATION FOR SEQ ID NO: 757:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 466 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 757:

TATCTATTGT ATGTAACAAT AGGTACATAA TATTTTATAGG GTGGGTATA TGAGCATAAT 60
 25 TACAAGATTG TTAAATAACA GTGATTTTGA AAAATTAAT CAACTATGTA AATTATATGA 120
 TGATCTAGGT TATCCAACAA ATGAGAATGA TTTAAAAAG AGACTAAAGA AAATAACGAA 180
 TCATGATGAT TACTTCTAC TGCTTTTGAT AAAAGAAAAT AAAATAATTG GTTTAAGTGG 240
 30 TATGTGTAAA ATGATGTTTT ACGAAWAAA TGCAGAGTAT ATGAGAATCC TTGCGTTTGT 300
 TATACATTCT GAATTTAGGA AAAAGGTTA TGGAAAGAGA TTATTAGCTG ATTCTGAAGA 360
 ATTTTCTAAA CGTTGAATT GTAAAGCMAT AACACTAAAT AGTGGTAATA GAGATGANAG 420
 35 ACATCTGCAC ATAACTATA TAGTGATANT GGGTATGTTA GCAATA 466

(2) INFORMATION FOR SEQ ID NO: 758:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 402 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 758:

CGTACAGCAT AACCAATAC ATATCTGCCA AAATCGCCAT ATGCCCCAGT TCCTCTTCCA 60
 50 TATTCGGTTG TCATTGAGTT TTAAATTTA TCGATAGAAA AATGTAAATG GCTATCTAAT 120
 TTGTTTTTTA AAGATTTTAT GTAATCTCTA TATTTAAATT CGAATATTCT TTTTATACC 180

AACTAATATA TwCAATTGTA TGCCTAGCAT AATCTCTCAT TAAAATATGA GGATATACCA 300
 TTTCTTTATC AAATATtTCT TCATAAATAT AATTAGCAAT CTCTATATGT ATTtCACAAA 360
 5 TGTCGATTG AAGTGTGCGC CCTAAAACAG rAGCATATAA TC 402

(2) INFORMATION FOR SEQ ID NO: 759:

(i) SEQUENCE CHARACTERISTICS:
 10 (A) LENGTH: 338 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 759:

AAGTTACTCA ACAGCCATTT GACACGTCTC GATTAGAGCG ATTAGGTATA ACCGAGCGCC 60
 20 AAATAAAGA CATGTATCGT TTATTAGGAC TAGCCAAATA TGAAGATCGA TTTGTTATTC 120
 CAACATCACA CAAAGAAACC TATTTAGATA CGTATCACGC ACAAGGTAGT ACAGGATACA 180
 ATTACGGCGG CGAGCATTTT GGAGATAACT GTGAAGGCTG TGGCGTTGCA GTAGGTTTCA 240
 25 GGAAACTGG TCAAGAAATT TATAATGGAG AATTCTATGG AGGGATTTC GTGGTTAATT 300
 TCGATAATTT AAAAAATATC CAAGAAAGTT TGGGTnnn 338

(2) INFORMATION FOR SEQ ID NO: 760:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 760:

CTCTATTATA TTTAGAGTTA TAGCTATTGC aGTTGTTTAT TTATTGAAAA TAAATGTACT 60
 40 ATCATTAGtT TTAGCAAGTG TATTAGGCAG TTTgGTATct AGGCTACTAT TATCTATTAT 120
 TTTAAATTTA CCTGTGTGGG TAGTGTGTGTT AAaCGCGATT CCAGGCGTAA TATTCACTTT 180
 45 AATTGTAGcT ATTCCTTTAT ATCTCACATT GAGAAAAAGA ATGGCAGTAT TACTAAGATA 240
 ATAAATCAAA ACACGGTCGT CACAATTACT GTTGGCGACC GTGTTTTACT AGCTATTTAT 300
 TGTTTTCAGT TTCTTTTGTA TCTAACAATT TCACTTTGTG ATTTTCCCAA TCAATTTTAT 360
 50 ATGTTGaTTT AAATGTTCTA GTTTTAAAGT TTTTATAAT 399

(2) INFORMATION FOR SEQ ID NO: 761:

55

(A) LENGTH: 639 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 761:

10 GCATATnCAA AACAAATAGCT GCTTTAGGTG CGGGGGCTAT TAGCGTTATT TCTGCACTAA 60
 CAGGAGGTTT AGTTGGCGCA GGTGTTGGTG GTTTCCTTGG ATCTATAGCT GCATCGAATA 120
 TTGATACTAG TAAGGGAATA TATATAAAAT TAAAACTAA AAAGTATGCA GCCGGGGAAT 180
 15 ACGTTCTGAC AGGAGAAAAA TGGGGATATC AGTAAGGGTG ATTTTATGGA TAGATTAAAA 240
 TATTCACCTA AAGTTGGAAT TTTAGCATT TTTATTATTTT GTACTTTAAA TTATTTAGTT 300
 CCAATGCAAA GCAATGCTTT TTCAATAATT ATATATTCGG CAATTTTTCG TGTGTTACTT 360
 20 ATGCTTTTATG TTTATATATT TTTAGGAATT TTAAGAAAT GACATGAAAC AAATTAGCAT 420
 TGGCTATGAA GAAATCTATG GGGATAGAAT TTTTCATAGC CATTTTTTAA AAGAGCATGA 480
 25 AGTAAAAAAT TGGAAAACCG TCAATCAAAT AATTGAAAAA GAACACTTGG ACAAAAATGA 540
 ATATGnAAGC GGAACCGCCG TTATTTTTCT TTTGCAGGAT CGTTAGATTA GATTnCGGTT 600
 GTATAGTATA ATGAAAAATA TTATAATAAA TTTATAAAA 639

30 (2) INFORMATION FOR SEQ ID NO: 762:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 852 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 762:

40 nTGTTTCGTTG TCTGGATCCA ACGTAAACAG TCGCTTCATC TAATACAATG ATAGGCGCAT 60
 CTTTTAATAT CATTCTTGCA ATAGTGAATC GTTGTTTTTC ACCACCAGAC AATTATCTC 120
 45 CCACTGTACC AACATTCGTA TCATATCCAT CTGGCAACTT TTCAATAAAT TCATGACATT 180
 GTGCTAACTT GGcAGCTTTT TCAACTGCCT CATCCGTAGC TTCTGGATTG CCAAGTTTAA 240
 TATTTTCTTT AAAAGTAAGA TTTAATAAAA AGTTATCTTG TCCAACAAA CCAACTAAAT 300
 50 CGTTAAGTTG TTTCGATTCA ATATCTTTTA TATTTATACC GCCAATCGTA ATTCACCTG 360
 AAGTCACATC CCAGTATCGT GATATAAGCT TGGCAATGGT TGACTTACCG CTACCAGaTG 420
 55 CCCCGACGAT AGCTGTGAAA TTATTTTCTG GTACTGTAAA TGATAAATGC TTAAAGACCA 480

GAGGCTTCTT AAATTTTGTT GACAATACTA ACTCTTCTAG ACTTAATATT TGATTCACCTT 600
 CAGTTAATGC GTATTGTATA GACTTTAAAT GATTTACATA ATTAGTAAAA TTCTTAATCG 660
 5 GTGCTACTAC ACCTAAAGAT AATACGATGC ATAGGAAAAA TTCCGCATAG TTTAATTGGT 720
 TGATAGATAT CAAATACATG CCGACCGGTA AAATCCCTAA AAATGTTGAA GGTAATACAC 780
 TAGCTCCTAA ATTCATATAC CCCCATGTAT TTTTAAACCA ATTCAAAGTG TGAATCTTAT 840
 10 AATTATCTAC TG 852

(2) INFORMATION FOR SEQ ID NO: 763:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 852 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 763:

CTCTTCATGC GTCATACGTT GTGTCGATTG ATCAGTTGTT TTGTCTAAAT CACTAGCTTT 60
 25 AAATTTAGAT TGATTTGATT GACGTGTCGT AAATnGTTGT TCCTTTTGTT GGCGTTTGCC 120
 TTTTTTCTT GATCTTATTA AAAATAAATT GATAACCCCA ACAATAATGA GCGCTAAAT 180
 30 AATGTAGCTA ATAATGAAGG TCGTAGTCAT TTAATGACCC CCTTAATTTT TATGGATTTT 240
 TACTTCAGCG TTCaTTCCaG GAACAACCTG TTTAGACGGT TcTGATTCTA GAGTGATTTT 300
 AACAGGTATT ACTTGAGAAA CTTTAGTGTA GTTACCATCA CTATTTGATG ATGGCATTAA 360
 35 TGAAAgtTG CAGCAGTTGC TTTTCCAATA CTATCAACTT TACCTTTAAT AGAAGCTTTT 420
 TGACCGTCAA TAGTCACATC AACATCTTTA CCTACTTCAA CATCTTTAAT ATCTTTTTCG 480
 TCAATATTTG CTGTTACATA TAAATCATCT AAATTGTATG CATAAGCGAT TGGGTTACCA 540
 40 GCTTGCACCA TTGAACCTTC CATACCATCT AATTTGGCAA TTGTACCTTT TTGAGGCATT 600
 TTAAGATCCA TATCTTTTCGT TTCGCCATCT TGACCTTGTA CAGTAACAAT TGCTACTTTG 660
 TCACCTTTAT CGAGTTTGTC ACCTTGTTTA ACATTAAGTG ATTTAATTTG TCCAGATGCA 720
 45 GGACTTGCTA TTTTAATTTG ATCGCCATTT ACTTTTGCAT TATCAGTTGT TACATAGCTT 780
 GTTGTTTTAT TCCAAAATA AAAGCCAGCA ATCCCAATGG CTAACnGTAC AACAAACCGTA 840
 50 ATGACATTnA AT 852

(2) INFORMATION FOR SEQ ID NO: 764:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 747 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 764:

	ACATATTGCA CATTCAATTT AAATTGCATA TAACTACATT AACCATTCA TGTGCAGCGC	60
10	CTGTTGTTTA GATGATCATT TATTGACTTT TTGCTCTTAT nCCAnGCAAT TTTTGTATT	120
	ATTGCTTATA ATACATTTTA AGACAAATGT TTGTTTCAAC ATCGCATGCG ACATCATTAA	180
	YTATTCATTT AAATTTATCC AGTTTCTATT TGTCTGCAAT TGCTGAAAAT TAAAGGAGCT	240
15	TATTTTTAAT AATGGAACAA ATTATCACTG AATTTATTAG CCGTTTGGG TATGCAGCCA	300
	TTTTTATATT AATTTkATkA GAAAATGTAC TACCTaTCGT aCCATCGGgA AATTATTCTG	360
	ACATTTGCTG GTCTTATGTC TGTAAATCA CATTATCAA TTTTAACTTT ATTTATTATT	420
20	GCAACTATCG CATCGTTTAT AGGGCTGTTA ATTTTATATT ATATCTGCCG TTTGATTTCa	480
	GAAGAACGTC TATATCGTTT TATTGATCGA CACGGTAAGT GGATTAAATT GAAAAGTAAG	540
25	GATTTAAAGC GAGCAAATGA TTGGTTTAAA AAGTATGGCG TATGGGCTGT ATTkATCTGT	600
	CGTTTCATAC CTGTATTACG TGTatTGATT ACCATTCCAG CTGGTGkaAA CCGCATGAAT	660
	GTTGTGACaT TtACCGTTAT TTCATTAAATA GGTACTACAA TCTGGAATTT CGGTTTAAATT	720
30	TTACTAGGAC GCACTTTGAG CGATAGC	747

(2) INFORMATION FOR SEQ ID NO: 765:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 820 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 765:

	TTTACCTTTT AACCAATCAT GCATTAATGC AATGTGTACT GTCTTAACAA ATGTTGCATA	60
45	TtCATCTTCT GCACTCCACT TTTTCAATTC CTGTTGATTa ATGGCATGTT TTTCTCCATT	120
	TTCCAACATA TTGACATACA ATTTATAATG TTTTCAACT GCCTGAATAA AAGCGAGTTG	180
	ATTTTTATTT AAATGTGATT TTAACTTTTT TAAATCTGAA TTAATAAAGT CTTCTATAGT	240
50	CGAATAAGAC ATATCTTGGT ATTCAACTAC TGCATTAATT TCATCTAATA ATTGCGATTG	300
	TGATTTTAGC GTTATATAAC TCTCGTCTCC ATAAGATATT TCTGTAGTAG CTGCTGTAGT	360
55	AGGATTTGGA GAACTGACTA AAATATGCTC GCCTAATAAT GCATTTATCa AACTACTTTT	420

TGTTCGTGG ATATCTTGCT TTGTACGCTT GAACAAAGGC ACATCCGAAA TTATATCAAG 540
 CGCTTTTGA ATATCTATAG ACATATTTGT TGTAGCTGTA TTTTGATTGA GCTGCTGATT 600
 5 ATCACGATTA TCCTGAGCAT TATCAGTAGC CACTTGATAT GTTGTCTCTT GTCGACCTAT 660
 TAATTTATCT AGAGATTCAT CTAAATGAAT ATAGTAGTGA CGATAATTCT TAGTCGTCAA 720
 TGA CTGACGC AGCTCATTTA ATTCTGTATA ACGTTGATAT TCTTTTAAAT CATCACTTTC 780
 10 TTCTGTTGGn AATTCATCTG CCTGCACATT TTCTATnATT 820

(2) INFORMATION FOR SEQ ID NO: 766:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 672 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 766:

TCAACATGAC ATTAATTTAA TAGCGATGCA TACAAATTTA GATGTAAATC CGTATGGTGT 60
 25 CAATATGATG TTGGCGAAGG CGATGGGTTT GAAGAACATT TCAATAATAA ATAATCAACA 120
 AGATGTATAC TATAAAGTTC AAACATATAT ACCTAAGGAT AATGTTGGAC CATTTAAAGA 180
 TAAGCTTAGT GAAAATGGAT TAGCGCAAGA AGGTAATTAT GAATATTGTT TCTTTGAAAG 240
 30 TGAAGGAAGA GGGCAATTCA AACCAGTTGG TGAAGCTAAT CCAACAATAG GACAAATTGA 300
 TAAAATTGAA TATGTAGATG AAGTTAAAAT TGAATTTATG ATAGATGCAT ATCAAAAGTC 360
 35 AAGGGcTGAG CaATTAATTA AACAATACCA TCCATATGAA ACACCGGTAT TTGATTTTAT 420
 TGAGATAAAA CAAACATCCC TTTATGGACT TGGCGTTATG GCAGAAGTGG ATAATCAAAT 480
 GACATTGGAA GATTTCCGAG CTGATATTAA ATCTAAATTA AATATCCCAA GTGTCCGTTT 540
 40 TGTGGTGAK TCTAATCAGA AAATTAAACG TATTGcAATT ATTGGTgGTT CAGGTATTGG 600
 ATATGA AKAT CAAGctGTCC AACAAGGcGC AGATGTCTgt TACGGTGATA TTAAACATCA 660
 GATGCCTAGA TG 672

(2) INFORMATION FOR SEQ ID NO: 767:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 403 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5 GAAAAAGTTA AACCTACTGT AACTACAACA AGcmAAGTTG AAGACmATCa CTCTACTAAA 60
 GTTGtMAGTA CTGaCACAAC AAAAGATCAA ACTAAAACaC AAAGTGTCTA TACAGTTAAA 120
 ACAGCACAAA CTGCTCAAGA ACAAATAAA GTTCAAACAC CTGTAAAGA TGTTGCAACA 180
 GCGAAATCTG AAAGCAACAA TCAAGCTGTA AGTGATAATA AATCACAACA AACTAACAAA 240
 10 GTTACAAAAC ATAACGAAAC GCCTAAACAA GCATCTAAAG CTAAAGAATT ACCAAAAACT 300
 GGTTTAACTT CAGTTGATAA CTTTATTAGC AcAGTtGCcT TCGCAACACT TGCCCTTTTA 360
 GGTTcATTAT CTTTATTACT TTTCAAAAGg AAGaaTCTAA ATA 403

15 (2) INFORMATION FOR SEQ ID NO: 768:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 768:

25 TTCAATGCTC GTTTTGCTTG CGTTAGTAAT TTCTAACATA TTCATTcACA TTTTGGAGCC 60
 TAAACTATCA ATCACCACtC AAATCATCAT CGTTTTGATT TTAATTGAAG CACTAATTGG 120
 30 ACTGCGTTTC TTGAAAGCGT ACGATGTtAA GCGTGGCAAA GATAAAGAAA ATAAGAAAAA 180
 TAGTAAGGAT TTCGTTAAAC TAAAATCAAT TTTAGTAGCA ATTTTATTtTA CATCATTTGGC 240
 GCTGACAGCA GGTACTGTAG CTGATATATA CGGTTTCACT GACTTAGGAA ATACTAGAAG 300
 35 TGATTTAATC GTTTGGAGCA TAGGTGGTAT TATATTTGGC CTCGTATGTT ACACAATGGA 360
 AGATAAAAGA TAACGATAAG GAGCTGGCGA TTATAAAGCT AGCTCCTTTT TTAACtTaTA 420
 TATGTAAAGA aCTaTCCTAA GGGkTTTTTAA TCATATGTCA ATAATTTCTA TAATACATTA 480
 40 TTA 483

(2) INFORMATION FOR SEQ ID NO: 769:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 769:

55 ATTATATGTC TGATGTATTT CATGTcGTTA AAAATAAAGG TACGGAAGAT TTTAAAGAAG 60

GGATGATGCT TGAAGGTGAA AAAATCAAAG CTTTTTATGA AGATATGCCA CCGTATCAGA 180
 CTGTCAAAAA AGGAACGATA CAAATTAAGC GTGATGGCAC ACCTATTATC CTATTAAATG 240
 5 TCATTATACG CTnGTAGCTA CCCGCAAATC GGTACAATCG CAGTTATCAT TTACGAATTA 300
 GCACAAAACC GCAGGATCAC GTTGAATTCC ATTATAGTAT TTACGGCTGA AGACTGTAAG 360
 10 nTAdACTGGT AACCATATCA TGGATGATTG ATCATAGATG 400

(2) INFORMATION FOR SEQ ID NO: 770:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 750 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 770:

ATCTCTTCTA CTTCTTTTAC AAATTGAGAA CGTCTTAAAA TGGCTTTCAC ACGGGCAACA 60
 ACTTCTCTAG GTkAAAAAGG CTTAGTCATA TAGTCATCGG CACCTAATTC TAAACCTAAT 120
 25 ACCCGATCAA ATTCAATCATT TTTCGCCGTT AACATTAATA TAGGGACTAA ATTTTTATTT 180
 GTTCTTACAG TCTTACATAC GTCAATGCCA TCTTTTTTATG GTAGCATAAC ATCTAAAATA 240
 ATTAAATCTG GCTGTTCACT TTCTACCTTT TCTAAAGCCT CATCACCATC AAATGCGACA 300
 30 ACAACTTCaT AACCAGCTGk TTCTArGyTA TATTTAAGTA ATGgTTaCGa TTGaATGtTC 360
 GTCATCTACT wCCaACACTT TTGCGACAT GgATGCCTC CCTAACTtAk AATTATATTT 420
 35 cATTATAACC GAACTATTTA TAAAAATAmC ATCCTACACA TTATCTTTAC ACATTTTTTA 480
 CATTACTTTA CATATAAATA AAATACTTCT TATATTTCCT TCTmTCATTT GCATGACTTA 540
 CTCTGGGACA ACGAAATAAA TTTTGTAATA ATAGCATTTT TATCCCACTA CCTATGCATG 600
 40 AGTTTTACTC ATTTATTCCT AAGCTTATGT ACATATTCGC TTTGTCTAAT GTGTAAGAAA 660
 CACTACATAA TCAATCATTG GTGACTCTTT ATTATTCTA TCCTGTGACC AaCTTCAATT 720
 45 CATTTAAAAA GGCGAACCTA GCAATTAAAG 750

(2) INFORMATION FOR SEQ ID NO: 771:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1462 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

	ATTTAATTAT TATTTTCCCA AATCAATACG ATATAATTTT ATGTTTTTAT CTTTTGGTGC	60
	TGAGGCAAAT TGATAAGtCG TTTTGCCTTT TTCAATATAT CCAAAATTAC CAGCAACTTT	120
5	GCTATATTCC GTATGTGTTA CTGCTTTGCC AACATAATTT TTAACAGCTT GATATGTTGG	180
	ACCATTAGCA TCATTATACA TCATTGAAAT ATGAGTTACT TTACCATTGT TTTTCTTACC	240
10	ATCAGTAGTT ACTAATAACA TACCTTTTTT AGTATGGAAT TCGTAATAAT GTTCAGTTCC	300
	ATCTTCATTG TAAGAATACA TCGGATTTTT GTACTTAGTT AAAACATCTT TAATAGATTG	360
	GCCAATTTTC ACATTTTCTA ATGACTGATC ACCCTTAATT AATTGTTTAA CAGTATCAAT	420
15	AGAGTTGCCA CTGCTGCTT CGGCACTAGT GTTCACTAAG CCTACTCCCA TTACAACAGA	480
	ACATGCAATT ATACTTGCCG TTAGTAGCTT TTTCATAAGC ACTCTCTCCT TTTATTTATA	540
	TCGTCTTG TG CTCACAACCA TTATACAACA ATCATTCTTA AAACAACAGT CATATTACAA	600
20	TTATATTACA AATAATAATT ACTACTTTTA TATTTCACTT ATCACTAAAA ATTAAACATG	660
	CTTTTCATCG TGATGTTTGT TGATAAATGC AGCTGTATCT TCGACGAATC TTTCTTGCTC	720
	TTcAACAAAT GGGAAATGCAC TTGATTCTTG ATATACTTCA AAGTCCGCAT TTTGGATTAA	780
25	ATCAGCTACT TCTTTAGCTT CTAATCTTGT TGTTCTTTCA CCGAATTCTC CTGCAATAAT	840
	TTTTGTGCGT ACTACGACAT TTCTATAAGT TTGAGAAATA TCAGCGTTTT TGAATACCTC	900
30	TTTAACTGCT TGAATTTCTT CTTTAGTTGA TATTTCAATT GTATCTACGA CATGTTTGAG	960
	GAATCGATTG ATTTTTCTCG GACGATAATA CTTACGCTTA TTTAAAAATT TAwCYTGtTT	1020
	TTCaGGATCC CaGTTtCGAA TAATATGGGC ATaTTTTCTA AATAAACGTT CTTCCGGTAA	1080
35	TTACCTTCA ATAGATGTTG GATTTACCAA CGTAAGTGAT GATGTAAATT CAGGATAACG	1140
	TACTGAAATA TCCGCACCAA TGATTCCACC CATtkCATGG cATACAAATG CAACTTCTTC	1200
	AATATATAAA TATTTAAGTA ATTCAACAAT GTCATCAGAA AAATCTTTTA TTTCAATGTG	1260
40	ACGAGGTTTA TCAGAATAGC CATGTCCACG TAAATCAATT AACACAACCT GAAATGATTT	1320
	TGCTAATTCT GCTGCTAATT TATTAAACAC AGAATAATTA TCAAGTACAG TATGAATCAA	1380
	TACGATAGGA TAGCCTTCAC CTAaNGTACT GTAATGTATC GATGTTCCAT CTTTTCTAGT	1440
45	AAATAGATCC ATAATTTTCT CC	1462

(2) INFORMATION FOR SEQ ID NO: 772:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 704 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 772:

5 ATTATTATAA GGAGTTATCT TACATGTTAA ATCTTGAAAA CAAAACATAT GTCATCATGG 60
 GAATCGCTAA TAAGCGTAGT ATTGCTTTTG GTGTCGCTAA AGTTTTAGAT CAATTmGGTG 120
 CTAAATTAGT ATTTACTTAC CGTAAAGAAC GTAGCCGTAA AGAGCTTGAA AAATTATTAG 180
 10 AACAAATTAAA TCAACCAGAA GCGCACTTAT ATCAAATTGA TGTTCAAAGC GATGAAGAGG 240
 TTATTAATGG TTTTGAGCAA ATTGGTAAAG ATGTTGGCAA TATTGATGGT GSTATCATTT 300
 CAATCGCATT TGCTAATATG GAAGACTTAC GCGGACGCTT TTCTGAAACT TCACGTGAAG 360
 15 GCTTCTTGTT AGCTCAAGAC ATTAGTTCTT ACTCATTAAC AATTGTGGCT CATGAAGCTA 420
 AAAAATTAAT GCCAGAAGGT GGTAGCATTG TTGCAACAAC ATATTTAGGT GCGGAATTCC 480
 CAGTTCAAAA CTATAATGTG ATGGGTGTTG CTAAAGCGAG CTTAGAAGCA AATGTTAAAT 540
 20 ATTTAGCATT AGACTTAGGT CCAGATAATA TTCGCGTTAA TGCAATTTC A GCTAGTCCAA 600
 TCCGTACATT AAGTGCAAAA GGTGTGGGTG GTTTCAATAC AATTCTTAAA AGAAATCGAA 660
 GAGCGTGCAC CTTTTAAAAC CGTAATGTTG ATCCAGTAGA AGTA 704
 25

(2) INFORMATION FOR SEQ ID NO: 773:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 427 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 773:

35 GACCACGACA CCACAAGCTA TGAAGAATGC CyTTGTTAAG CCGCcTTTGa ATTGCATAGA 60
 GATGAAAATA CCAATATTAA TAAAGAAGTT ACAGAAAATA CCTTTTGTA AAATATTCAA 120
 40 CCATGTTGAA TCAACAGTCT TTTTCTGAAC TAAAGCTGTT AAAGCTTGTG TCATTTTCAGG 180
 TGTCATGACA TGCGCAAATT TCATTAAGAA AAATAACACA AACCCACCTA AAATATTTCC 240
 45 TAAGAAACAA TATAATAAAA TCCAAGTCAT CTTCTTAACA GAAACGACTT TATAATACCA 300
 GCCTACTGTA AAGTACATGA AGTTACTTGT TAATAATTCA GAGTTAGTTA ATACAATAA 360
 AATCAAACCT AACTAAACG CAATGGCTCC CATTAAATTG ATAAGTCCAT CTACGTGAGT 420
 50 ACTCGCA 427

(2) INFORMATION FOR SEQ ID NO: 774:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 400 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 774:

	AGGTGAGCAT TATTAAATAT TGCCAAGAAA TCAAAACTAA AACTATGCA AAAGATGTAG	60
10	ATATATTATT TATTCTTGGC GGTGATGGCA CGGTTAACGA ACTTATAAAC GGTGTTATGA	120
	CACATGACCT TCAACTCCCT ATTGGCATT TACCAGGCGG TACTTTTAAT GATTTTACAA	180
	AAACGCTAAA TATAGCACCT GAATCATAAA CAAGCTAGTG AACAAATGGA TTTCTGCACA	240
15	AGTTGGTACA TACGATGTAA TTGAAAATGA AATGAATCAA TATGCACTCA ACTTTGTGGG	300
	CTTAGGTCTC ATGTTCAAAA CGCGGAAAAC GTACAAGAGG TTCAAAAGAT GTATTGGGTA	360
	AATGGAGTTA TATTGGGTCC ACCGTCCAAA ACnCGGcNAA	400

20

(2) INFORMATION FOR SEQ ID NO: 775:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 484 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 775:

	AATTATTCCT CAATAGGAGC ATTACCCGCT CTCTCACCAA TTCCCCAAAA TGTAGTTTCT	60
	ATTTGCTTTG CACCATTTAA AATTGCAGCT AATGTATTTG CAGTAGCCAA CCCTAGATCG	120
35	TTATGACAAT GAGCAGAAAa AATTATGTGA GAATATTTTT TTACAAAGTA ATTAAAAATA	180
	TCTCCGTATT CTAATGGTGT CGAACATCCT ACAGTGTCGG CAAATGTAAC TGTCTAACT	240
	TGATATTTTG AAATAATTTT CATGTATTCT TTCAATTTTT CTCTAGAAGT CCTTGTTCCT	300
40	TCTTCAAAAC AAATATCTAC TCCTTTTTTTA TCTTTCTTTA ATATATCCAA GCAGTCTTTG	360
	ATTTTCTGAA TATAATATTT ATTTGAAAAA TTAAGCTTTT CTTTATATG CAAGTCAGAT	420
45	ATTGGCAATA GTATTTtTAC CACTAAATTA TGAATTTtCA ATtTGACTAT TTTTgGTATG	480
	TCnT	484

50

(2) INFORMATION FOR SEQ ID NO: 776:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 788 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 776:

CGCAAGATGA AGTTAAAGAA CTTAATCGCT TATTAGGTAA AGTCATTCAT GCATTGATG 60
 5 AAACAAAGGA AAAATAATTA ACTTTTGTCA TGACAATTAA AGTAATGTTT AGAATTIATT 120
 AAGAATAGAA AAACAATTAG CACGCGTAAC TTGTTAGTAA AAAAAGTCTG TGAAAGGTTT 180
 CTTAGCCTAT CAAGCAGTTT TTTTATGCAT TATATTGAAT CATATTCCAT AAAGCACCTT 240
 10 GATTAGCTAG TAATTGTTGA TAGTTTCCCT TTTCAACTAT TTTACCATT ATCATCACAA 300
 TGATGGTCTC AAAACGTGAC AATAAAGTTA AATCGTGTGT AGCAACAATT AATGTTTCTG 360
 CATGTGCTTC AATTAAATCC ATAACTTTAA AACTATTTTG TTTATCTAAT GCAGTTGTTG 420
 15 GTTCATCTAA AATCCATGTT GATGCAGTAT CTTTTAATAA CATCCTCGTA ATCGCTAAAC 480
 GCTGAATTTT TCCGCCAGAT AATGTATGAC CATCTAAGTC AATTGACGT CTAGTGCCA 540
 20 AATGTTCTAA ATCTAATTGC TTAAATATTG CTGACCCG TCATCTTTT TCATCGGTAA 600
 ATAAATTTTG ACGTATTGTA CCATCAAATA ATTGTTGAGA TTGTAGCAAG ACATTTAACG 660
 ATTCAAACCT ATCTTTGTCA TCTATTTCAA ACATATCCAT ATTTTCGAAA CGAACAGAGC 720
 25 CACTATCTAA TTGATATAAC CCTGCCATAA TTTGTAGTAA TGTACTTTTT CCTGAACCAG 780
 AAGGACCC 788

(2) INFORMATION FOR SEQ ID NO: 777:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1145 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 777:

TTACAAAAGA AATGCAACAA AATTTTTGAA TCATTACATT TTTTATAAA AATTTCACTT 60
 40 TAGATTACAA ATAATTACTT ATTTTGTCAA TTTATTTAAT GTCAATATGT TGATTAATTA 120
 ATAGTGTGT CTAATGTATA TAATATTTAG GTCATCGTTA TAGTCAACAA TAATAAGGTA 180
 45 TTTGAGTTG AAATTTATCT TATTATTTTT CCACTTTTAC GTGCTATCCC ATTACACAAA 240
 AACAAATAAG TAAAGATATT AAGTACAAAA AAGTGAACA CCTGTAGAT GCTCCACCTC 300
 AATTATATTA AGTTATATTA TTTTGCTGCT TGGTATAATT CATCAACTTT TTTCCAGTTA 360
 50 ACAATGTTCC AAAATGCACT CATATAGTCT GGACGTTTAT TTTGATATTT CAGATAGTAG 420
 GCATGCTCCC AAACATCAAA TAGTAAGATT GGTGTTTTGC CTCTGTAA TGGATTATCT 480

55

GATCCAAATA ATGTTGTTGC TTTATTTGCA AATTCATTTT TAAATTCATC TAAAGTGCCC 600
 CACTGCGCTT TGATGTCATC TATTACGCCA CCTTTTTCTT CAGAATTAGG TGATAGTATT 660
 5 TCCCAGAATA ATGAATGGTT AAAATGACCA CCGCCATTAT TACGGACTGA CATCCTCATC 720
 GCTTCCGGTA CCTTGTCTAA GTnAGCAATC ATATCCGCTA GTGATTGATG CTCTAACTCT 780
 GTTCCTTCAA CTGTTGCGTT TAATTTTCGTC ACGTACGTAT TGTGATGTTT GTCGTGATGA 840
 10 AACTCCATTG TTCTTTGATC TATATATGGT TCCAATGCAT CATATGCATA TGGTAAATTT 900
 GGTAATTTAA ATGCCATAAG TATATTCCTC CTTTTATGAA TATACTTTTA TAATAATTAA 960
 TTTTGGGTGT GTTTTGCAAT AATTATTATT TATTTGTTAT ATCCAATTAT ATAAAATTTA 1020
 15 AATATTCGAA ATTGATATAA TATTTAATAT TGAATGCAAA AAGGCATTTA ACTGCTTTTG 1080
 TTTCCGCTT TTAAAGAGAA TTAAAAAGA CAACTTCCAT TTTTCAATAA GAAATTGCCT 1140
 20 TTTCT 1145

(2) INFORMATION FOR SEQ ID NO: 778:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 902 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 778:

AAAAGAGGnC CAATATTATC AATATACGGC CGGAAGATTG CCTTTTTCTT CTCCACTTAC 60
 TAAGCCCATC AATGGTGTAT GTTATCaATT GGTGCTATCC GCATTAAATA ATTATCTTCA 120
 35 ATTGTGGTAT CAAAGCTATT CGCTTTTTCT CCAATAATGT CTTTACCATA TTTATTGAAT 180
 GCACCTACTA AAATCAACAT AAATGGAATT AACAAGGTTA CTACAATTGC CATTCTTATA 240
 40 AATTGTGGTA TATAAGAAGG TCGCCAATT AAAGCTTTAA TACCTTCTTW TCTAAATGGA 300
 GTTCTATAA AGCGATATGA TAGCTCTGCA AATATAATTG TTAAACTTAT ATCTATAAAG 360
 TACACATATA CAGGTATCTG TCCGTCTACA TAGTAATAT GTACGAACT AATTACTGCA 420
 45 AAATGCCATA AATATAAACT ATAAGACCTT TTCCCGATAA ATACTAACAC TGGATTTGAA 480
 AATATCTTCG CtATCCATGT AGATGGATGA ACGACACTAG CAATAATAAA TAACGTTAAT 540
 ATGGATATTA AATAGAAACC ACCATCATAT ATCCAATTCG TCTCATCATT AATAATGAAA 600
 50 AATAATAATA TAAGTACTAT AAATGATAAA CTACCTATGC TATCAATAAC ATATTTTACA 660
 ACTTTAGGTG GATCATTTTT CAATTTAAAC GGTGGCCATA AAAAAGCTAA AATAACACCC 720

ATACTATAGA TAAACATCAT TAACCCTAAA GAAATTATTG ATACTCCCCA AAAAATAAAT 840
 CCTATTTTGT ATCGCTTTTT AATTGTTAAC AATAATGTAA CCAAATAAC TGGGAAAAAT 900
 5 AT 902

(2) INFORMATION FOR SEQ ID NO: 779:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 779:

GGAATTAGTT TCCTAATGAT CAAACCATAA CCATTAGCTT TTAATGATGT TACGCCTTTA 60
 AAGATTAGAG GATGATGACT AATAATAGTA TTATAACCTT TTTCGATTGC TTCATTTACT 120
 ACTTCCAACG TACAGTCTAA TGCTGTTAAA ACACCAGTAA CTTCAACATC TTCATCACCT 180
 ATTAACAATC CTACATTATC CCAAGATTCA GCAGTACTAA ATGGCACATG ATGATCTAAC 240
 AATGTCATTA AATCAGCTAT TTTCATaCTA TAACACCCTT TCAATTACaG CAATTTGTCGTC 300
 aTTAATTTGa GCTAAAcGTT GakGakGTTG TTCAGTawTG AGtTTCGaTT TAAtATGATA 360
 AAgTGCyTct AACTctCTTT GcCATTtTT TATAAAATAT 400

(2) INFORMATION FOR SEQ ID NO: 780:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 557 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 780:

GTTCTGTTT TAACATCAAT ATCTATCTTT TTAATATCTT GAACATAAAG CAGATTCAGT 60
 TGGCTAATAT TTGAATTAT GTTAATCACT TTTTTCGTTT CATTTTTAAA ATGTACTGTG 120
 TAAGTnGCTT TTTTTCGATA TTTTAATTCA CGTTCCCCGA TACCTCTATC ATGCTTCAAT 180
 ACTGaTTTAA CTTTTTTAGA GATATCTTTG TAGCCTACAC GTGGGTCACC TGTAATTTT 240
 AAATCTGaTA AAATTGGTGT TGATGTGCCA TTTACTGCAA TTGTGTATGG TACATATCTG 300
 TCTGCTTTCG CTTTAGTTCC TGTTTTAACA GTAATTTCaA TTCTCTTAAC ATCTTTCGCA 360
 TGAACtaAGT TTGCTGTGTA ATTTTTTGAA TTCAACTGTA AGATTCTTTT ACCACCATT 420

CTATCATGCT TCAGTACTGA TTTTATTTGA CTAGTTAAAT TTTTGTAATT TGTCCAAGGT 540

TTATnTGAAA ATGAAAG 557

5 (2) INFORMATION FOR SEQ ID NO: 781:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 475 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 781:

TTCTGATAAC ACAACTTTCT TGGACCAAGT GACAATGGTA AAGTTGATAT TCTCAAAAAA 60
 TATATTTTAT TGAAAAGGTA ACTATATCGA CCGCTCTGTG CGACTTGCTG TG TAGAAACA 120
 20 TCAAGCCCAG CTTGAAACTT CATATAAACT AATAACTTTT GATAAATGCT ATTCCCAATT
 TGTTCATACT GCTCCAAGTT ATCATTAAAT CTAAATTTAA TACTTACTTC ATTTTAACT 240
 GGAATAAAAT GTACATCACT CGCTTTCATT TCTATCGCTT TATTAATTAT TTCTTGAAAT 300
 25 AGAATCTTCA AAAAAACACC TCCTACATAT AATCACGTAG GAGGTGTTTT TATTACTTCA
 ATTTAACCGT GTAAAAATGG ATTTAATTGT TCATCATCAA CCGTCGTATA TGGACCATGT 420
 CCAGGGAATA AAGGTAAATC GCCTTCTAAT TCAAATATTT TATCTTGAAT AGAAT 475

30 (2) INFORMATION FOR SEQ ID NO: 782:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 782:

ATGAATATCG GTGCAGGACG TATCGTTTAT CAAAGTTTAA CTCGAATCAA TAAATCAATT 60
 GAAGACGGTG ATTTCTTTGA AAATGATGTT TTAATAATG CAATTGCACA CGTGAATTCA 120
 45 CATGATTCAG CGTTACACAT CTTTGGTTTA TTGCTGACG GTGGTGTACA CAGTCATTAC
 AAACATTTAT TTGCTTTGTT AGAACTTGCT AAAAAACAAG GnGTTGAAAA AGTTTACGTA 240
 CACGCATTTT TAGATGGCCG TGACGTAGAT CAAAAATCCG CTTTGAAATA CATCGAAGAG 300
 50 ACTGAAGCTA AATCCATGAA TnAGGGCATn GGCCATTTGC ACCGGGGCCT GGCCGTAATA
 AGCCAAngGA CCGGGCCAAC CGTTGGGACC GGAAGAAAA 400

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 810 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 783:

10 GATTAATGTA CTCGTGTACA TAACGTTTGA AGTATGATAA ATAGATATAG AAATAAAGTC 60
 ATATGTGACA TCAATTAAAT GATGTTCAAA TGACAAGATA CAATAGAGGA ATGTTTGTAA 120
 15 TAAAAACGCT TCATATAAAG GTCGAGTCAA TATATGATAC GACTTTATAT GAAGCGTTTT 180
 ATTTGCTATG AGCTAGTATA TTTTATAATA ATTTTCTAT TTCTCTTTCG ATTTGAACAG 240
 GTTTTTTTTG AGGTGCAAAT CGTTTAACAA CGTTACCyTC GCGATCCACT AAAAAGTTAG 300
 20 TGAAATTCCA TTTGATTTTC TCATTAAAGA ATCCGTGTTG TGCCGCAGTC AAATATCTAA 360
 ATAAAGGTAA TTGATGTTCC CCTTTTACGT CTATTTTTTG ATGCATAGGG AAGGTAACAC 420
 CATAGTTTAA TTTACAGTTT TGAGCTGCTT CTTGCGCTGA ACCAGGTTCT TGGCCACCAA 480
 25 ATTGATTACA AGGGAAACCT AGAATTACAA ACCCTTGATC TTTGTATTTT TCGTATAATG 540
 ATTGCAAACC TTCAAATTGT GAAGTAAAGC CACATTCGCT AGCTGTATTA ACAATTAGCA 600
 30 TAACGTCACC CTTATATGCA TCTAATTTGT AAGTAACACC TTTATTTGTT TCTACTACAA 660
 AATCATAAAT TGTCTCCATT GTATCATCCT TTCGATTTAC TTAAaATGTA CCaAAAATC 720
 GTGTAAtAGT CTkTACTAa ACTCTATGAT AGAATACTTT GAGTAGGATT TTATTAAGGA 780
 35 GATGTATAAC ATGGnTCAGC AACAAATTCA 810

(2) INFORMATION FOR SEQ ID NO: 784:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 788 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 784:

AAATTCATTG CAGAAAACAT AGATAAATAA ACAAATTGAC TTAAACGAG CGTTGCAACA 60
 50 TATCTCGAAT TGTAAGGAG CTTGAAAATG AATAAAAATA TAGTCATTAA AAGCATGGCA 120
 GCATTAGCCA TTCTAACCTC AGTAACTGGA ATAAATGCTG CAGTCGTTGA AGAGACACAA 180
 CAAATAGCAA ATGCAGAGAA GAATGTTACG CAAGTTAAAG ATACAAATAT TTTTCCATAT 240

5 ACCAATAAAC ATGTATCAAA AGATTATAAA GTTGGCGATA GAATTACTGC CCATCCAAAC 360
 GGTGACAAAG GAAATGGTGG TATATATAAA ATTAAAAGCA TTTCTGATTA TCCGGGTGAT 420
 GAAGACATCT CTGTCATGAA TATTGAAGAA CAAGCAGTCG AACGTGGACC AAAAGGCTTT 480
 AATTTTAATG AAAATGTCCA AGCATTCAAT TTTGCGAAAG ATGCTAAAGT TGATGACAAA 540
 10 ATTAAAGTTA TTGTTTACCC ATTACCTGCT CAAAATAGTT TTAAACAGTT TGAATCTACA 600
 GGAActATAA AAAGAATsAA AGACAATATT TTAAATTTTG GATGCCATAC ATTGGAACCC 660
 GGGGAATTCA GGGATCACCA GTTCTAAATT CTAACAATGA GGTCAATAGT GTGGTGGTAT 720
 15 GGGCGGGTAT TGGGAAAAAT TGGGTTCTGG AATAATAATG GGTGGCCGTA TACTTTACGC 780
 CTCCAdAT 788

(2) INFORMATION FOR SEQ ID NO: 785:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1023 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 785:

30 CAATATTTTA GAAACAaTAA TTAActTGCG ATGAACAAAC TATTAACAAT AATCTTGAGT 60
 ATTATATTTA TCTTAATTAA TAATATATTT ATTAAAGATT ATGTTACTTC CAACTTTCAA 120
 AGTAGAAAAA CGGTATAATT TGTTGATGGG TGTTTATTGA TAACTGCAA AAAATACAGC 180
 35 CTACAACCAC ATAGATTGTA GACTATATTT AAAATAATAG GTATTTATCA TATCTCGTAA 240
 ACTTAGTGCC AATTTTTATA CTCGTGGTGC TGTAAGCTA CCCTTAAATT CAGGAACGTA 300
 GTGTGTaGGG CTATAActTG GAACAGCATA TTGATAAATT ACATTTTTGA TATTTAATGA 360
 40 TGGTTTCCCA ATTTTATAAC CATTTGATTG TGAAAATGAG AAATAkTTCT TCACACCTTT 420
 AACTACTTTA TAAGAATAGA AGTATTTATA GTCATATGCT TTATTTACTT TAGCATTTTG 480
 45 ATGTGTTGCT GTTGTGTTAT TTTGGAAACT TGGTACATGC ATACGATGTG AATTGTGACC 540
 ATATGGTGGA ATTACCTTGA AACTATTTAT TTGTGGCACA ACACAAAAGT GATTAATTTT 600
 AATGCTAGCA TGCCCTGGTG TAACAAATTT ATGCGCGTGA TATCCAGGAA CTGCAAAATG 660
 50 ATGCTTGATA ATTAAAGATT gAGATGGATG TGTATATCTA GGCGATTCTG ATGGTTTAAC 720
 AATAAAGTGT TTATTAATAG AATCCTTTGC ATGATTTACA TGTTTATGTA CATGTGTTGA 780
 TTTGTATGAA GTAATAACTT TCTTGTAGTG GGTTCGCGTA GTAATGAAGT GGTGGTTTAC 840

TTGTTGTGAA GCTGGATTGT TTGTTGCTAC ATTCACGCGA TTATTGATTT CTTTGTACTC 960
 TGGCACAATG TTTCCTAACT TTGATTCTGG nACGACAAAG TTTTATCTA CAATTTTACT 1020
 5 AGC 1023

(2) INFORMATION FOR SEQ ID NO: 786:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 487 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 786:

TCAAGTTTAT TCGCTTGTG GAAAGCTTTA ACTTGATTTT CTAAAGCTTT ATCAAATTGC 60
 20 GTTGATTCAT TATCAACTTT ATAACCTAAA GCTGATAAAC CAATTTTAAT AGTTTAAATA 120
 TTTTATCAT CGTCTCCAAC TTTAAATGTT TTCGTATTAG GAATGACATT TAAAGATTGA 180
 TATTTAGGTG TGTCAATAGT AACGTCTGGT TTAATGCCTT TACCGTGAAT ATAATGACCA 240
 25 TCTGGCGTTA ACCATTTTAT TTCAGTATAT TTTAACAATG AACCATCCTT AAACCTCTCTT 300
 GTAGTTTGTA CGACACCTTT GCCGAATGTT TTTGACCCAT AAACCTTTAGC TTTATTATAG 360
 TCTTTTAGCG CACCAGTAAA CACTTCAGAA GCGCTAGCTG AACCTTCATT CACTAAGATG 420
 30 GATATATCCA TGTCTTTCGC TTCTTTTAAC GCATCATTAG AAGTtGAATT GCTCAGTACT 480
 TTACCTT 487

(2) INFORMATION FOR SEQ ID NO: 787:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 803 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 787:

GnAAACnATA TTTAnAATTA AGTGTTCGAT TTGCAGTAAA TTCATACAGA AAAGCTAAAA 60
 TTAAATGAT ACaGTGGAAT AATGTCACAT TCTCCATAAA TAATTGACCG AAGAAGCATA 120
 50 AAACATAGAA TAGTACAGTA ACACTTACCG GCTGCTTTTC TTTATATATA ATTGCATTAA 180
 TACAATAATA AATAATAAAG AGTGAAATTA GTGTTGATGT CGCATAATTA TAAATCCTG 240
 CAAACCAGCC ATATGTATCT GCATAAATAG CACTTGGTAA AATTAACATT AAAGAAAATG 300

TGCCCATACT AATGAGGCCA TATGAAAGCC AACGTAACCA GCTTACATGT ACAGCTATAA 420
 TTTCAAAGAT ATTTCCGATA TAGCGACCAT TAAGTGATGC AAATCCTACT TTTAAAATAT 480
 5 CAGTATTATA ATTGCTAAAC CATTGTAAAT CATCGTGCAT GAGTGGTAGT AAGATACCCA 540
 TAAAAGTATA AAACAATAAT ATCGCAATTA ATATCAAAGT TGTCTTGTGT AATTGAATTG 600
 TTTTCACTTT GCTAATCCTC AAATCTAGTT AAATTTTCCT CAACTTGTAG GTCGAAAAAT 660
 10 TAATTCAATA TTTTAAATGT ATTTCTAATT TTCACCTATG CATGTTTCCT CAATCAAATT 720
 AGATAAACAA GGTATTTAAT ATTACTTTCA ACAATTTATC TAAATCGCCC CTCGTCTTTT 780
 TCTATGaCGA ATGATTACAC TTG 803
 15

(2) INFORMATION FOR SEQ ID NO: 788:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 788:

GAAAAGTTGT ATTATGAATA AAGTTAATCA AGGTGCTCAG GAAGAGGCAA TGGAAGAGTT 60
 ATTAGTGACT TTTCAAAAAT TGATTAAAGA CTAAAGGAGT TTAAGATGAT ACATCAAAAT 120
 30 ACGATTTACA CAGCGGGAAT TGAAACAGAA GAACAAGTAA GTCAATTGAC AGAACGCATT 180
 TCAAATATGA TAGGTGTTCA TCAAGTGAAT ATTAATATAA TAGATGGTCA AGTAACTGTA 240
 TCGTATGAGA CACCAGCAAA TTTGAATAGT ATTGAAAAAG AAATCTATGA TGAAGGATAC 300
 35 AAAATTGTAT TTTAGGGTAT AATGTAGAGT GCGCTATAGA TTTTAATTTT GAAAATAAAT 360
 TAAAAATTTT GTAAATGATG TAGTAAAGGT ATGTCGAATA 400

(2) INFORMATION FOR SEQ ID NO: 789:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 789:

GTTTAAAGAT GGAGAGATTA TTGATTTTTC AGCTGGAAAA GGTGAAGCGG TATTGAAAGA 60
 TTTAATCAAT ACTGATGAAG GTTCAAGAAG ATTAGGTGAA GTAGCATTAG TACCTGATGA 120

ATGTCATTTA GCGATTGGAT CTGCTTACGC CTTTAATATT CAAGGTGGAA CGGAAATGAC 240
 TGTGAGGAA AAGATTGCAA GCGGATTAAA TGATTCAAAT GTACATGTCG ATTTTCATGAT 300
 5 TGGTAGTAGT GATTTGACTA TTTATGGCAT ATTCGAAGAT GGTTCAAAAG AACTAGTATT 360
 TGAAAATGGA AATTGGGCAT CAACATTTTA ATAAATGTTA TTTTGAGGTG CTGAGTAGGA 420
 AATGAAACAT GTATTTAAAG GTCAAATACG TGATTAAAGT ATAGATTGGG AGATAAAATA 480
 10 ATGACAAATC AGGACAGACC AATGAAATCT ATGTCAGAAT CAAAATGTTA TAAAAATAGA 540
 CAAGTTTTCC CTCAAGATAC GAATCACCAT CATACAATGT TTGGTGGTAC ATTGATGGCT 600
 15 AATATTGATG AAATTGCAGC AATCACAGCT ATGAAACATG CTGGTGGCA AGTAGTTACC 660
 GCATCTACAG ACTCAGTAGA TTTCTTAAAG CCGATTAAAA CAGGGGACAT ATTACAATAC 720
 GTAGCGATGG TTTCATACGC TGGGACTAGT TCAATGGAAG TG 762

20 (2) INFORMATION FOR SEQ ID NO: 790:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 790:

TTTAATTTTA TCTTTTGGAT ATCACATAGA TGCTACTTTT CTATACTGTT CTAAATCTAT 60
 TATTTTCGTCA CATTGATTG CAATAGCCAT ATCATGTGTT GCTAAAATCA GCACTTTATT 120
 35 TTCATCGACC AAACCAAATA ATGATTGAAT AATCATCTGT CCTGTTTTAG GATCTAACGC 180
 ACCCGTTGGT TCATCAGCTA ACATAACAAT CGGATCTTTT AACATCATT TAATTAAAGC 240
 GACACGTTGT TGTTACCTC CACTTAGCGT ATGAACTTTT CTTTTTAAAC TGTTTGACAG 300
 40 ACCAACTGT TCTATATAAC GATCTCTAT TTGTCTTTT TCTTCTTAC TTATTTTTTT 360
 ATATGCTAAT CCAATATCTA AATTTTCATT TACTGnCAAA 400

45 (2) INFORMATION FOR SEQ ID NO: 791:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 791:

AACGTAAAGC GGAAGTTGGG ATGGTTGAAG ATGCTGAATT TCATGCGCAA CTACATCGTT 120
 ATAATGCTTT TCTAGAGCAA CATCAAGATG ATAAAGTGTT GTATTTnGAA ATTGGAATTG 180
 5 GTTATACTAC ACCACAATTT GTGAAGCATC CTTTTAGCGG TTGACACGTA AAAATGAAAA 240
 TGCCCTTTAT ATGACGATGA ATAAAAAGGC TTCGCATTCC GAATTCAATT CAGGACGTAC 300
 CATACTTTAA CTGAGATTCT CAACTTGTTT AGCAGACTCC GGACGCAGCA CACGAAATAC 360
 10 ACATGGGGGC AGAGTnACTT ATGGACCGTT GAATGnGATA 400

(2) INFORMATION FOR SEQ ID NO: 792:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 678 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 792:

GTTCCGTTTG ATACAGAGAA GCGAATGTAA GCATAATCTT TAACAGTATC GTATGATATA 60
 25 AATTAAATTGG CAACTTTTTG TCACCTTCAT AACTTCAAA TTTTCTCAA AATTGACCTG 120
 ATTGTAATCC TAATTCAATT TCTGGTTTTG AATCAGTGAA AATAACTCTA GCAGGTTTAA 180
 30 CAGAACTTGC ATAATGATAA AACTGTTGAG TTCCATCTTT CTTTTTCATT TCAAAATCAA 240
 TTGGACGAGA GTTTGGTGCG CTATGATCTT TGTCTTTTAT TGCAGGGTTT TTAATCGCTT 300
 CTCTAAGTTC CTGaTTCAAA ATAGGATATG TATTGTTAgT GGCyTTTGCT GCTGGTTTAA 360
 35 CTTCTTTTGT TTCCTTAGGG GcTTTAACTT CTTTAACTTC TTTAGCTTCT TTTGTTTCAG 420
 AAGTAGGGGC CTCAACTTCT TTATTAGATA CTGAGACAGC ATTAGCTACT GGTTTAGTTT 480
 CTGGAGCTTT TTCAGATGTT GTTGTTGGAC TTGCAACTGC TTCAGTTTTT GGTTGTGCTT 540
 40 CTGTATTTGT ACCACCTGTT TCTTCAGCTG CTGCTTGTGT TCGCCATTG ACATTAATAA 600
 TAAAAGTGTA CTAATTGCTA CAGATGCAAC GCCTAGTGAT GACTTTCTAA TTGAATAAAA 660
 TGATTTAAAT TCTTTTTG 678

45

(2) INFORMATION FOR SEQ ID NO: 793:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 351 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

ACTGGCGGAT CCACGATGGC ATGTGTTAGT GAAGCAATTC ATTTATTACC ATATAATGTA 60
 TTCTTCGTAC CAGCCAGAGG TGGACTAGGC GAAAnTGTTG TCTTTCAGGC AAACACAATT 120
 5 GCAGCCAGTA TGGCACAACA AGCTGGCGGT TATTATACGA CGATGTATGT ACCTGGATAA 180
 TGTCAGTGAA ACAACATATA ATACATTGTT GTTAnGAGCC ATCAGTCATT AAACACTTTA 240
 GGACAAATTA AAACCAAGCA AACGTTAATA TTACACGGGC CTTTGGTGAT GCGCTGGAAG 300
 10 ATnGGCGCCT CGGACGTCCA ATCACCTGGA AAAGGTCCTT GGAACCACTT C 351

(2) INFORMATION FOR SEQ ID NO: 794:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 794:

GGGCGAGGCG TTCGCGCACC GGCCGAGCAG CAAATAGGCC GGGATGTCGC GTCTTGCTGC 60
 25 AGGTCCGAAA AATATCAACT TGAAGAGGTA ACACCGnTAG ATATGTTTCC ACAAACAACA 120
 CATGTTGAGA CAGTGGCATT ATTCAATTTG AAATAGCGCA CATGATTGA GATAAAATTT 180
 30 CGCCTTAATA AGTGAAGAAA GAAGTCGTAT ATTAAAATTT TATTACAGCC AATTTACTTT 240
 CTTGATATGA ACTTTTTTAAA TTAAATGTGA CATTGTATAC TATATTTAAA GAAGAATAAG 300
 AATGTCATGA TCGGGAGGGT TGGAATGCA TAAATTTGAT TTAACGACAA ATAATTTTCA 360
 35 AATGCGAGGT TTATTATCCT GGCAACCTGG TTATTGCGTT 400

(2) INFORMATION FOR SEQ ID NO: 795:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 407 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 795:

ATGTTCCAGG AACGTATTTT TACAACCGAC AATTAGCATA TGTCATAATG AGTTTTATAA 60
 50 TTGTATTTT TATTGCATTT TTAATGAATG TTAAATTACT GAGTAATATT AAAGTGCAAA 120
 AAGGTATGAT TATAACTATC GTCTCACTAT TATTACTGAC GTTAGTAATA GGTAAGATA 180
 TTAATGGTTC TAAAAGTTGG ATAACTTAG GATTTATGAA CTTACAGGCA TCTGAGTTAT 240

55

TAAGTAAACC AAAATTAATc TTAAGTCCTA TTGTATTAGC ATTAGGTTGT ACGTTTTTAG 360
 TTTTCCTACA AAAAGACGTG GGCAACATTA CTAATATTAA TnATnTT 407

(2) INFORMATION FOR SEQ ID NO: 796:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1053 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 796:

CAGTAGGTAC ACTTGCAGCA TTTGTTGGAT ACTTAGAGTT ATTGTTTCGGG CCTTTACGTC 60
 GTTTAGTCGC ATCATTTTACA ACTTTAACGC AAAGTTTTGC TTCAATGGAC CGTGTATTCC 120
 AATTAAATTGA TGAAGATTAT GACATCAAAA ATGGTGTGG TGCTCAACCT ATTGAAATTA 180
 AACAAGGTCG TATTGATATT GATCATGTTA GTTTTCAATA TAACGATAAC GAAGCTCCAA 240
 TTTTAAAAGA TATTAAATTTG AGTATTGAAA AAGGAGAAAC AGTTGCTTTC GTAGGTATGA 300
 GTGGTGGTGG TAAATCAACA TTAATTAAC TAAATACCGAG ATTTTACGAT GTAACCTCTG 360
 GGCAATTTT AATAGaTGGT CACAACmTTA AAGaTTTTTT AACGGGAAGT TTAAGAAATC 420
 AAATAGGATT GGTGCAaCAG GATAATATTT TATTCTCTGA TACAGTTAAG GAAAATATTT 480
 TaCTTGCTCg TCCaACAGCa ACAGATGAAG AAGTAGTTGA AGCGGCgAAA ATGGCTAATG 540
 CACATGACTT TATTATGAAC TTGCCACAGG GATATGACAC TGAAGTAGGT GAACGAGGTG 600
 TTAAATTATC AGGTGGTCAA AAACAAAGAT TATCGATTGC TAGAATATTT TTAAATAATC 660
 CGCCAATTCT TATCTTGGAT GAAGCAACAA GTGCACTTGA TTTAGAAAGT GAATCCATTA 720
 TTCAAGAAGC ATTAGATGTG TTGAGTAAAG ATCGAACGAC ACTTATCGTA GCGCATCGCT 780
 TGTCCACTAT TACACATGCT GACAAAATTG TCGTAATTGA AAATGGACAT ATTGTTGAAA 840
 CAGGTACGCA TCGTGAATTG ATTGCAAAAC AAGGTGCTTA CGAGCATTTA TATAGCATTG 900
 AAAACTTATA AAGTATTAGT TGTTTGACTT CAGTACAATC TTGAAGAGAA AATTTGTAAC 960
 AGGATGGTGG GGTGATAACA TAGAAAAGC AGTAAGAGAT TTTCTTAGTT GAAAATAATC 1020
 TThCTGCTTT TTAAATTTA ATTTGnGAT TCA 1053

(2) INFORMATION FOR SEQ ID NO: 797:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 544 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 797:

5 AAGAGTGACT CCTGAAGTGA AAGAAGGGGA CCGTGTCTGT TTCCAACAAT ATGCTGGTAC 60
 AGAAGTTAAA CGAGATAATG AAACATATCT GGTATTAAAT GAAGAAGATA TTTTAGCAGT 120
 TATTGAATAA TACAGAACTT AATTCATAAA TAAATTAAAT AGAACGAAAA TGAAACACAA 180
 10 CTAAACAAAT GGAGGTTTAT CATTTATGGT TAAACAATTG AAATTCTCTG AAGATGCACG 240
 TCAAGCAATG TTACGTGGTG TTGACCAACT TGCAATGCA GTTAAAGTAA CGATTGGTCC 300
 TAAAGGACGT AATGTTGTAT TAGATAAAGA GTTTACAGCA CCTTTAATTA CGAATGATGG 360
 15 TGTGACGATT GCyAAAGAAA TCGAATTAGA AGATCCATAT GAAAATATGG GGGCTAAACT 420
 AGTTCAAGAA GTCGCAaATA AGACAAATGA AATTGCTGGT GACGGTACGA CAACTGCAAC 480
 AGTATTAGCT CAAGCAATGA TTCAAGAAGG CTTGAAAAAT GTTACAAGTG GTGCGAACCC 540
 20 AGTT 544

(2) INFORMATION FOR SEQ ID NO: 798:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 798:

CAATTGCATA GATATTGCTG ATAGACGTAC GGCTTTGTTT ATCAACTTCT AATAATCCAC 60
 35 GGTCAGCGAA TTTAACACCT AATTCGTCT GTGTTTGGAC GACGACCTAC AGTTACTAAT 120
 ACATAATCAG CTTGATTGT TTTCTCTTCG CCTTTAGCTT CATAAGTAAC TTAACTCCG 180
 TTATCTGTTT CTTGAGCTGA TTTAGCCATA GCTTCAGTAA CGATTTCAAC ACCTTTTTCT 240
 40 TTCATACCTT TTTTAACAGG TTGTGTGCAT TTGnTTTTTC GAGGCCACCT AAGATATCTT 300
 TAGCACCTTC AAGGGATGGG GTACTTTCTG nAACCAAAGT TAGCAAATGC TGTACCTAAT 360
 45 nCTGGTCCAA TG 372

(2) INFORMATION FOR SEQ ID NO: 799:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1182 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 799:

5 AAACGACATT TGTTTAATTG GAAGAATTaG CACCATTTTT TGAGGCGGGT ATAGATTCAT 60
 TTAAATCGA TGGTATTCTA CAAACGGAAG AATATATTAA TGTGGTAACG GAACAGTATC 120
 GTCAAGCGAT AGATTTGTAC AATGAAGATC CTGAAATCTA TGAAGATGAG AAATTTATGT 180
 10 TGATGGATCC AATTGAAGAA ATTCAACCTg ATCATCGACC ATTTGACGAA GGTTTCTTAT 240
 ATAAACAAAC AGTATATTAA GGAGGTTAAT CATGAAGACA ATAGAAGAGA TTAAATCAAC 300
 TCCTAAACA GTTATGAAGA AACCAGAATT ATTAGCACCT GCTGGAAACT TAGAAAAGTT 360
 15 AAAAATAGCA GTACATTATG GCGCTGATGC CGTATTTTTA GGTGGTCAAG AATATGGATT 420
 ACgTTCaAAT GCTGaTaATT TCaCGATGGA AGAAATAGCT GAAGGTGTTG AATTTGCGAA 480
 CCGTTACGGT GCCAAAATTT ATGTTACGAC AAATATTATT GCACATGATG AGAATATTGA 540
 20 AGGTCTAGAA TCATATTTGC GTAATTTGGA AAAGACTGGT GCGACAGGTA TCATTGTTGC 600
 AGATCCTTTA ATTATCGAAA CATGTAAAGA AGTTGCGCCA AAACCTGAAA TTCATTTATC 660
 TACTCAACAA TCACTTTCTA ATTACAAAGC TGTAGAATAT TGGAAAGAAG AAGGATTGGA 720
 25 TCGTGTGTGA TTAGCACGTG AGACCGGCGC GATGGAAATG CGTGAAATGA AGGAAAAAGT 780
 AGATATTGAA ATCGAAGCAT TTATTCATGG TGCTATGTGT ATCGCCTATT CAGGTAGATG 840
 30 TACATTAAGT AATCATATGA CTGCAAGGGA TTCCAACAGA GCGGTTGCT GTCAAAGTTG 900
 CCGTTGGGAT TATGAATTAT TAGAAGTTGA TGATAATGGT GAACTTGATG TTTTTTATAA 960
 TCAAgGTGAA GTTACACCGT TTGCGATGAG TCCTAAAGAT TTAAAATTAA TCGAATCAAT 1020
 35 TCCTCAAATG ATGGATATTG GTGTGGACTC ATTAAAAATT GAAGGACGTA TGAAGTCAAT 1080
 TCATTATATT GCAACAGTTG TCTCAGTATA TCGTAAAGTC ATTGATGCGT ATGCCGCAGA 1140
 TCCTGACACT TTAAGATTAA TCCGGAATGG TTAATAGAGT TA 1182

(2) INFORMATION FOR SEQ ID NO: 800:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 989 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 800:

50 GTATTCTTCC AATAACACCT TTCATTTGGA AACCGTTTAG AAAGAATACA GTTACTTCAG 60
 TTTGGTTTGC TTTAAAATTC TCTaGTGCTT TGTCTTGGAT GTTTTCGtTG CAATCATCTG 120

	TTTTCTTTAT CTAACCAATG AACACTCATT TTATTCTTGA ACCATGTCAT TTGTCGTTTT	240
	GCATATTGGC GTGAATGTTG CTTTAAATCA TTGACAGCAT CTTCATAAAT CATTGTGCCG	300
5	TTAATCACAG GTATTAAATC TTTATATCCA ATAGCTTGCA TACTTTGGCA AGATTTCATAG	360
	CCTTGTTCAA CAAGTTGTTG CACTTCTCTA AATAATCCGT GATCCAACAT AATATCAACA	420
	CGTTTATTTA TTCTTGAATA TAATGTTTTA CGCGACATTT CAATCCCTAA TAATAATGTA	480
10	TCATAATTTT CAGTAAATTG TTGCACTTTC TTGCGATTAC TCAAAAGTTT TTTTGTTTTT	540
	AAATAATATT CAATAGCGCG CAACACTCTT TGGcGGTTGT TAGGGTGAAT ATTTTCTGaC	600
15	AGAAACCGaC ATCAAATTGA GCTAAATAAT CGTGTAGTTG CTGATTATCT AAATGTTCTA	660
	ATGCAGATAA CTTTTGTTTA ACTATGGATA ATTGTGCAGG TGTAcTGTTC CATCTTCTAA	720
	TTCATAATTA TATATTAATG ATTGAATATA TAAGCCTGTT CCACCTGCTA TGATTGGaAC	780
20	TTTACCTCTA TTCGTTATAT CAGTAATTAA ATCTTCTGCT AATCGCTTGA ATTTCATATGC	840
	TGAAATGTA TCATCAGGAT TCAAGATATC AATTAAATGA TGTGGAATAC CATCCATTTC	900
	TTCAGGTGTh ACTTTTGCAG TTCCAATATT CATATGTCTG TAGACTTGCA TAGAGTCACC	960
25	GCTTATGATT TCACCATTGA TACGCTTCG	989

(2) INFORMATION FOR SEQ ID NO: 801:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 543 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 801:

	CGTTAGATAT AGAATTAAAA GATAATATTA AATTTCAAAA CGGTCAAAAA TTGACTGCAG	60
40	AAAAAGTGAA ATCTAGCCTT GAAAATAGCA TGAAAAAAG CGACTTGGTC AAATATTCAT	120
	TACCAATATC ATCAATTACC GCTAAAGGTC AAAAAGTAC AATTAAACC AACTCCGCTT	180
	ACCCTGaACT TGTATCTGAA TTAGCTAATC CTTTTATGGC AATTTATGAT ACAGATGCTA	240
45	AATCAGATGT TAATCAAACCT CCTGTTGGTA CAGGCCCTTA CCAAATAAAA GATTATAAGC	300
	AATCTCGAAA AATATCATTG TCGAATTTTA AGGACTATTG GCAAGGTAAA CCGAAACTTG	360
	ATCATATTAC TGTGACGTAC CAAGAAGACG GCAATAATCG CGTCAGAAAT TTAGAATCTC	420
50	AAAAAGATGA TTTAATAACT GATGTCCCAG TTAATAAAGT TCAAGACATA GAAAATAATC	480
	AAAATTTAAA AGTGTCAAAA GAATCTGGAT TTAGAACTTC TTTACTTATG TATAATCATA	540

55

(2) INFORMATION FOR SEQ ID NO: 802:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1139 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 802:

	GCGAATTCGT ACACACATAT ACACAAAGAT AATCATAGTT TTACATTGAA GCACATGAAA	60
15	GATAATTCAT TTAAAGGAAG GTATTATCAA TGACTAAAAA AATGGGATTA TTAGTTATGG	120
	CTTATGGCAC ACCTTATAAA GAAAGTGACA TAGAGCCATA TTATACAGAT ATTAGACATG	180
	GTAAACGTCC ATCTGAAGAA GAACTTCAAG ATTTGAAAGA TAGATATGAA TTTATAGGTG	240
20	GTTTATCACC ATTAGCAGGT ACAACAGATG ACCAGGCTGA TCGCTAGTT TCAGCATTAA	300
	ATAAAGCATA TGCAGATGTT GAATTTAAAC TATACTTAGG ATTAAAACAC ATTTCAACCAT	360
	TTATCGAAGA TCGGGTTGAA CAAATGCACA ATGATGGCAT TACTGAAGCA ATCACGGTAG	420
25	TACTAGCACC ACATTATTCT TCATTTTCAG TAGGATCATA TGACAAACGT GCTGATGAAG	480
	AAGCTGCAAA ATATGGTATT CAACTTACAC ATGTGAAACA TTATTATGAA CAACCTAAAT	540
30	TTATTGAATA TTGGACGAAT AAAGTCAACG AAACATTAGC TCAAATACCG GAAGAGGAAC	600
	ATAAAGACAC GGTATTAGTT GTTTCGGCAC ATAGTTTGCC AAAAGGTTTA ATCGAAAAGA	660
	ATAATGATCC ATATCCACAA GAACTAGAAC ATACTGCGCT TTTAATTAA GAACAATCTA	720
35	ATATTGAACA TATCGCGATT GGTGGCAAT CTGAAGGTAA TACAGGTACA CCTTGGTTAG	780
	GGCCAGATGT ACAAGATTTA ACACGTGATT TATATGAAAA ACATCAGTAT AAAAAGTTTA	840
	TATATACGCC AGTAGGTTTT GTATGTGAGC ATTTAGAGGT GCTTTATGAC AATGATTATG	900
40	AATGTAAAGT AGTTTGCGAT GATATTGGTG CGAATTATTA TCGTCCAAA ATGCCGAATA	960
	CACATCCATT ATTTATCGGT GCAATTATTG ATGAAATCAA GTCTATATTT TAATGACGAA	1020
	AGAAGCGTGA AACGTTGTGA CTAAATCAGT GGCTATTATA GGAGCGGGGA TAACAGGTTT	1080
45	ATCAAGTGCA TATTTTTTAA AACAGCAAGA TCCTAATATT GATGTAACCA TCTTTGAAG	1139

(2) INFORMATION FOR SEQ ID NO: 803:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 989 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 803:

5 TATCCnAAAT AnAAATGTGC GCATTTATCA AAACCTTACA TACAATAGAG TTTTCCCTAA 60
 CAGTAAATTA GATATTATTA CACCTGTTGA TATGTCTTCT AATGCCAAAC TGCCAGTTAT 120
 TTTTGGATG CACGGTGGTG GTTATATTGC GGGTGATAAG CAGTATAAAA ACCCATTATT 180
 10 AGCGAAAATT GCTGAACAAG GGTACATTGT TGTGAATGTA AATTATGCAT TGGcGCCACA 240
 ATATAAATAT CCCACACCAT TAATTCAAAT GAATCAAGCa ACTCAATTCA TTAAAGAAAA 300
 TAAAATGAAT TTACCTATTG ATTTTAATCA AGTAATTATT GCGGGTGAYT CTGCAGGTGC 360
 15 TCAATTAGCT AGCCAATTTA CGGCAATACA GACGAATGAT CGCTTAAGAG AAGCCATGAA 420
 ATTTGATCAG TCATTCAAAC CATCGCAAAT TAAAGGTGCT ATACTATTTG GTGGTTTTTA 480
 TAATATGCAA ACAGTTAGAG AAAGTGGT TCCAAGAATA CAGTTATTTA TGAAAAGTTA 540
 20 TACTGGCGAA GAAGATTGGG AAAAGAGTTT TAAAAACATT TCACAAATGT CGACAGTAAA 600
 ACAATCGACA AAAAATTATC CaCCAACATT TTTATCTGTT GGAGATAGCG ATCCATTCTGA 660
 AAGTCAAAAT ATAGAATTCA GTAAGAAATT ACAAGAATTG AATGTACCAG TAGATACTTT 720
 25 GTTTTATGAT GGTACGCATC ATTTACATCA TCAGTATCAA TTTCACCTTA ATAAACCTGA 780
 ATCGATAGAT AATATCAAAA AAGTGTTACT TTTCTTAAGT CGTAATACAT CCTCTAGTGG 840
 TATTCAAAC GAAGAGAAAC CACAAATAGA AAATCCGAGT AATGAATTAC CGTTAAATCC 900
 30 TTTAACTAA TGATAACAG TAGTAATTTA TTAAGTAAAG AACATTTAAG ATTTTCAAAT 960
 TAAAAACGAG AATTTAAAAC ATGTGGTGC 989

(2) INFORMATION FOR SEQ ID NO: 804:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 711 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 804:

45 TAAnCCCTGG TTTAATGATT TTGATTACGT GTTTTTATAA TAAAAACATA TCGAACATTG 60
 ACTACGTTAT TAAGCTGCTT TTTGTACAC TTTATAACCA ATAGCTTAAG ATTTAAACT 120
 50 AATCGGAAAG AACAATGATT CACCAAAAA ATATTTATGT TGCTATTAAA AATCAGTTAA 180
 TACGAATGTT AAAATACGTT TGATTTTCAT TAATAATGAT TCAAGTTTAT TTAAATGAGC 240
 GTTAATGTCA GTCTGTTTTG ATGCACCTTA TAATAAGAC AGATAGTTCA AATTACGTAA 300

AAAGCAACAT TAGCATTAGG AATATTA ACT ACAGGTGTGT TTACAGCAGA AAGTCAA ACT 420
 GGTACGCGA AAGTAGA ACT TGATGAGACA CAACGCAAAT ATTATATCAA TATGCTACAT 480
 5 CAATACTATT CTGAAGAAAG TTTTGAACCA ACAAACATTA GTGTTAAAAG TGAAGATTAC 540
 TATGGCTCTA ACGTTTTTAA CTTTAAACAA CGAAATAAAG CTTTTAAAGT ATTTTTACTT 600
 GGTGACGATA AAAATAAATA TAAAGAAAAA ACACATGGCC TTGATGTCTT TGCAGTACCT 660
 10 GAATTAATAG ATATAAAAGG TGGCATATAT AGCGTTGGCG GTATAACAAA G 711

(2) INFORMATION FOR SEQ ID NO: 805:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 680 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 805:

ATCAGGAAAA ACAGaATAAT GTAAATCAAG CTGTTCA GCC TCAAAATAAT ACTAATGAAA 60
 25 CATCAAAAGT ACCGGCTAAT TTTGTCAAAT TGAATGATAT TAAACCAGGT GATACTTCTA 120
 TACAAGGAAC AACTTTACCA AATCAATTTA TACTATTAAC TATTGATAAA AAAGATGTGA 180
 GCTCAGTTGA AGATTCTGAC AGCAGCTTTG TTATGTCTGA TAAAGATGGG AATTTTAAGT 240
 30 ATGACTTAAA TGGTCGCAAA ATTGTTTATA ATCAAGAAAT TGAAGTGTCT TCATCAGATC 300
 CCTATTTAGG TGACGATGAA GAAGATGAAG AAGTAGAAGA AACTTCAACT GAAGAAGTTG 360
 GTGCTGAGGA AGAAAGTACA GAAGCTAAAG CTACATATAC AACACGCGA TATGAAAAAG 420
 35 CGTATGAAAT ACCGAAAGAA CAGCTAAAAG AAAAAGATGG ACATCACCAA GTTTTTATCG 480
 AACCTATTAC TGAAGGTTCA GGTATTATTA AAGGCCATAC CTCTGTAAAA GGTAAAGTTG 540
 40 CTCTATCTAT TAATaATAAA TTWATTA ACT TTGAGaCAAA TGCTAtGGTG GtCCaAATaA 600
 AGAAGaAGCG AAATCTGGAT CAGAAGGAAT CTGGATGCCT ATTGATGACC AAGGATACTT 660
 TAATTTTGAC TTCCAAACGA 680

(2) INFORMATION FOR SEQ ID NO: 806:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 950 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

GTATTTTCATT AGCCATTGGC AATTGACATC TGACAATGAG GAATGGCTTT TGCATTGGAA 60
 TTCACAGTTG GAGTGCATTT ACTGGTTATT TTGGAGGTCA TTGTTGATAC ATTATTAATG 120
 5 GAGAATAACC GAATTTGTTA TGTTATTTCC ATTTTAAATA TTTGCAATTG TATTAAATGC 180
 TGCACCTTGA GATAAAATTA AAAATCCTTA TGGATCTGCC ATAATTCTTG TTCTAGTTAT 240
 TATCGTATTA AGTTGGGGAG GTATTGACAA GACTTGTTTCG TGGTAAAGTA CTTCAAGAAA 300
 10 AAGAAAATGA ATACTTTTTG GCAGCAAAAT CAATTGGTAC ACCCACATAT AAAATTATTT 360
 TGAAACATCT TTTGCCGAAT ATATTAAGTG TAGTTATCGT ACAAGCAACA TTGTTATTTG 420
 CCGGTATGAT TGTAGTGGAA TCAGGrTTGA GCTTTTTAGG ATTCGGAATT AGTAAAGCAA 480
 15 TACCATCTTG GGGTAATATG TTGAGTGATG CTCAAGAAGG GGATGTTATA AGTGGTAAAC 540
 CGTGGATATG GATGCCACCT GCTATAATGA TTACATTAAC TATATTAAGT ATAAACTTTG 600
 TAGGGGAAGG GATTAAAGAT GCTTTTAATC CTAGAGGTAG ACGTTAAATA ATAAAAGAGG 660
 20 CACTAGTTAA TTCTAGTACC TCTTTATTTT ATCTCTTACG TCCTAAACCC ATCGCTTTTT 720
 CCATTTTTTT GACAGTTTTA AATGAACTT TGTGTGCTyT ATCTCTACCT TGATCTAAAA 780
 25 TATCAyCAAG TTTATcTGAG TTATAGAAAC TTTCGTATTT TTCTTGAAT TCTACTAAAA 840
 ATGCTTTAAC TATTTcAGCA AGGTCACCTT TAAATTTACC ATAACCTTCG CCCTCATATT 900
 TTGCCTCAAT ATCTTTAATT GGCATGTCTG TTAATCCAGC GTATATTGAA 950

(2) INFORMATION FOR SEQ ID NO: 807:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 418 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 807:

GTTACCAGTC GAAGCTTTAC GACCTGTTAC GAAACATTGT TTACCCATAT GAGTACCTCC 60
 TTTAATAAAAT ATAAATACAC ATAACACTT ATATACTTAA TTAAGATAGC ATAGTTTCAT 120
 45 TTGAAAAACA ATGAATAATT TTCACATAAA AGTCAAAAAT ACTGAGTTTT GTGATATAAT 180
 TGTAGACTGT GAAGTTATGT AGTATGATAT TTTAGAGAAA TAAGTGAATG AAACATTTAA 240
 AATTTTTATT TtATGATACT ACATCTATTA AAAGAAAACG CCTATAGATA ATCGTCaTGA 300
 50 GTGTAAATTC AGTttACTCa AcAtATAaTA GGkGaTTTAA gCgTAAGtCa TTCTGATAAG 360
 TTGTACCATT GATATTTTAT GAAATAnTCn ATGAACTTAA AGCGTTTATG CTACACTA 418

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 842 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 808:

10 AATGGCnATT AATCTTTAAT ACGATGCTTG AGGATTTTTC CTAATAAAAC CTTGATTTCm 60
 AAAAGGGTTT AAATCmAATG AAACAATAAT AAaAAATGw CGCAATATAA TAATAAGTAC 120
 15 AAATTTAATT AAGAAATTAA ATTGATTGTA TATGTATATT TTGGTAACGT AAAAGAGAnA 180
 TATACAAAAT AATTAATTAT TTATATGAAA AGAGAATATA AATGAAGTAT AAAACAGAGA 240
 GACGTGAAGC GATGGGATAT TTAaMAAGGT TTGCAITGTA CATAAGCGTT ATGAYTTTAA 300
 20 TATTTGCGAT AGCAGGTTGT GGCAAGGTA ATGAAACAAA AGAAGATTCA AAGGAAGAAC 360
 AAATCAAAAA GAGCTTTGCG AAAACATTAG ATATGTATCC AATTAAGAAT CTCGAGGACT 420
 TATACGACAA AGAAGGATAC CGAGATGGCG AATTTAAAAA GGGTGATAAA GGGATGTGGA 480
 25 CGATATATAC AGATTTTCGCC AAAAGTAATA AACAAGGTGG ATTGAGTAAT GAAGGTATGG 540
 TCTTATACTT AGATAGAAAT ACACGGACTG CAAAGGGACA TTATTTTGTT AAGACATTCT 600
 ATAATAAGGG CAAATCCCA GATAGAAAAA ATTATAAAGT TGAATGAAA AATAATAAAA 660
 30 TTATCTTATT AGATAAAGTA GAAGATACAA ATCTAAAAA GAGAATAGAA AACTTTAAAT 720
 TTTTGGACA ATATGCAAAC CTTAAGAAT TGAAAACTA CAACAATGGT GaTGTCyCAA 780
 TTAATGAGAA TGTTCCAAGT TATGACGCAA AATTTAAAT GAGCAATAAA GATGAAAATG 840
 35 TT 842

(2) INFORMATION FOR SEQ ID NO: 809:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 809:

TAAGACTATG CCATCTTGGC AACGCGTTGT CGCATATTCA TTATTAGTAA GTATATGCAA 60
 50 GCATGATTTT GCCATAActT TACcCtTTC TATATTAAAG TACCACTTTT ATCATTCCCT 120
 ATAATATTTA ACTTATTTAA TTAAACGTAA ATTTAACACA ATACGAATAC CAAATCAAAA 180

AAAAATAGAT GTAGTCAGTT TAATTAACT ATCCAATTGA AACACACTA CTTTTTAGTA 300
 TTTTCAAAAT AATTTTAAAT GACCACATCT ACAACGTATT ACTATTATCT TTGTTAGTT 360
 5 ATATATTTCAT TTTCAATTTA TAAATAAACA CTTAATTAAA TCACATTCTA CTTTTGAGAT 420
 GAGCTCAAAC TAACATTAAA TTGTTTATAT T 451

(2) INFORMATION FOR SEQ ID NO: 810:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1499 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 810:

ACATTTTGTG CTGTTGTCC AATCCAATT CTACTTGTG GAACAGTATG AGTTGCTTGA 60
 CCACGCATGT TACTCAGCAT TTGTCCAACA TGAGTATTTT GAGCAAAGTT ACGAGATTGA 120
 TTCAACACAC TGTTTTTCAT CATTTCCATT TTAGGTGAAA TACGACTAAC TGCTTGTTC 180
 25 ATTGTCTTTT GCAATCCTGT TTTAACTGCA AATTTTCTC CGCCTTTGAA GCCTAATTTT 240
 ATTAACTTTT TCCCAGCACC TTTGAGGAAG CCCATACCTG GCAATGGAAT AAGCGAAAGT 300
 CCTCCATGA TTCGTTCTTC TTTAGATAGC TTTCTCCAG TTACAATATT TTTTCTGTG 360
 30 GCTGCATTG CTGCTGAATA AGCGGAATAA GCACCAATAG CTACTGCTCC AGCAACAGGA 420
 TTAATACAAG ATAAACTAC CATACCAGCT AGTGCTGCAA TCTCTATAGC TTCTTCTTTT 480
 TTCTTTTGCT CCTCTAGCTC TTCCGCTTCA GCTACTGTCA TGAATCACA AGAAGCGCCC 540
 35 TGAGTCATCA TTTTTCAAA ATCACTTTTA GATAATTCT TATCTTCTT TTTCAATTCT 600
 TTGTAATGCT GATCAATTAC TTTTACAAAT TCTTGATCCG ATTTAATAT TTCATCGATG 660
 CTATAACAT CGCTGTCACT TTTCTGGCA TCTCCCTTG CAATTTTCCC AGAAGCCCCA 720
 40 TAGGTATCTC TACCATGTCC ACCACCATT TTAATTAGTA TATTTTCCA TTTAACATT 780
 GTTATACCTT TGAATGCTT ATCTATATCT TCTATAAATG GCGTGTCAAT AATTTCTTTC 840
 45 ATTTTCGAAT CTAATTGAGA TGCATAATCA TATAACTCaT TTTGATAATA CCCCATTAA 900
 TTAGATCCTT GTGGACCTAC TACACTACTT GTTAATGATA TTGCGCCATC TTCGTCTAAA 960
 GAACTGACAT CGCTTACCGC ATCATCCATT GTTGATCTA CATCTAAGAA GTCTTGACCC 1020
 50 TcTATTGCAT CTATTACACT TTGGATAGAT GACTTTTCAT TGTCAATATC AGCTGTTAGA 1080
 TATTCAATAT CTTTTGTCAT GTCATGCACC TATCCCTCAA TATTATAGTA AAGCTTGGCA 1140

CCACCGAGTG TACTCAATGG ACCTTTGAGA CCACTAATGA CTGTTGCAGC ACTACTATAG 1260
 CCATTTTCCA TAGAATGTGC AGCTTTTTCa CTACTGTAGT AATTtAATTG CGCTATTTTT 1320
 5 CCAAAGTTTT GAGCTATTTT ATCATACTTT TGACTTATGC TTTCCAATTC TTTTACAACA 1380
 TGTGCAATCG TTTcAGCTTT AACACTAATT TTTCCACTCA ACGTCATTAC TCCTCtGcTt 1440
 TATtAATATG ATTTTcATCA GTATCGAATC CAAATAATTC TCGTGATAAA nGTATTGAA 1499

10 (2) INFORMATION FOR SEQ ID NO: 811:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 620 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 811:

CnCGCTGCAC CGAAGAATGC TTTTGGTTTG TGTAAGATG CAGGATCTAA ACCACCTGAT 60
 AATGTACGAC CACTTGGTGG AATAACyAAG TTATAAGCGC GTGCTAATCT CGTTATAGAA 120
 25 TCCATTAAAA TAATGACATC TTCCCCAATT TCTACTAAAC GCTTTGCACG TTCAAGTAAT 180
 AATTcAGCTA CTtTAACATG GTGTTCTGGT GGTTCGTCAA ACGTTGAATG AACGACTTCA 240
 GCAGCTTCTA CTGAGCGTTC TAAATCTGTT ACCTCTTCAG GACGCTCGCC AACTAACAAA 300
 30 ATAAATAGCT TTGCATCTGG TTTGTTcGTA CTGATTGCAT TCGCTATTTT TTTTAATAAC 360
 GATGTTTTAC CTGCTTTAGG TGGCGCCACT ATTAAACCAC GTTGACCTAA ACCAATCGGT 420
 GTTACTAAAT CCATGATGCG CGTTGAATAA TTTTGTATTT CTGTCTCTAA TTTAATACGC 480
 35 TCATCTGGAT AAAGTGGTGT CAAAGCTTGG AAATGCGGAC GTTTCTTCAC TTCTTCTGCG 540
 TTATGGTCAT GACAAAGTCA ACTTGTAATA AGCCATAATA TTTTTCGTTA TCttTAGGTT 600
 TTCTAACTTT CCCAGTTACT 620

40 (2) INFORMATION FOR SEQ ID NO: 812:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1094 base pairs
 (B) TYPE: nucleic acid
 45 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 812:

CCGGCTGAAA TTAAGTTCTT GTTATATTTT GCAAAAATGA CAAAGAAAGA TAAAAATAAA 60

TGAAC TATTT GCAATTAGAG AATTGTTAGA GAAAGGTTTG ATTGGTGATT GTATTCAACC 180
 TATAATTGAA CCAATTAAAT ATACAACCAC ATTTAAAAAT ACTTTGCAAT ACTGTGGTGA 240
 5 AAAAGCATT C TCTATAAATT TAGTAGTAAA TTCAAAGTTA ACTGAAGAAG AGATTAGTAA 300
 CGAAaCTGt GCACATTTAA CTGAAATAAT AACAAAAAAC AAAAGTGTTA TTCAAAAAGC 360
 10 TTACTTGGGT CCTTCTGATG AAGGCAATGA TAGGTTGAAA CAGCAATTTT CAAGTAATAG 420
 TTTAGCTATT TTAACAAGTG TAGATGATTG GGAAATGTTT GGAGATAAAA ATAAACTTGA 480
 AATGGTTTTT GTACCAGATG ATAGACACAT TAAACGTAAA TTGCGTAATA TTCCAAACAA 540
 15 AGGCATsATT ATGGATCCTT TTAATAAACT AAGTCGTAAT GTTGATTATT TAGATAATGA 600
 TGACGAGTTT TATAGCGACG ATCACCTTTA TTATAAGGAA GATGGATACG TAGCATT TTC 660
 AGACTATTCT GTTATAGGTG GAGAATATGT AGACGGTGGC TTTTCGCCAT TAGCtATTGc 720
 20 GtTACATATT GTCTATTTTG ATGAGGCTAA TGAGCTAAGA GTTAAGCATT TtGTCTCTGa 780
 TTCTAATAAT GATAGATCAA ATCCAGGTAA AaNGTTTTTT GAGGCTGTAG ATAAATTAGT 840
 AACATGGTCA AAAAaCTTAG ATATTAAAAA TAGATCTTAT GCGCTTGGAC AATTTGAAGA 900
 25 ATTAAATGAA AATAATAAGT ATCCAGGATT AGGTTTAATT AAaNGTTATC TATCAGCATC 960
 ACCTAGAAAT TATGAATAGA TACTGGGTCT CAAAGAAATA GGAAAAGTTT ATATCGAACT 1020
 30 GAACTGCAAG AATGCATTGA CATGGATGAA CACACGGGTA AAATGATCGC AGCTGAAATG 1080
 ACTATGGTAT TAGC 1094

(2) INFORMATION FOR SEQ ID NO: 813:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 930 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 813:

45 CACCTATCAT CGTCTCTTTA ATTCTATCTA TCTATTTACC CCAATCGTAA ATATGATGGA 60
 CGTTATAGAA TACCAAGAGT TGCAGGTATn TCTATTATTT ATCTAGCTGT AGTAGGTGTT 120
 ATTACGTTAA TTGTTAATTT ATTGATACCT ATTATTGGTT CGCAAGTAGA TAGTTTAGTT 180
 50 AAAAATTCAC CGCAATATCT AGAAAAATTA ATTAATTCTA TTGATAAAAT AGCAAAATAT 240
 ACGTTTTTCT CTTCGTATTA TAGTCAAATT AATGATTGGT TAAATTCTTT ACCTAAGAAA 300
 ATACCATCTA TGTTAAGTGA ATTTACAGAT GGCTTTGGGT CTAAAATTGC AACGTTTGCA 360

CTTAAAGATG GACATCACTT CAAAGAATTT TCAACGAATA TTATGCCACC GAAATTCCGA 480
 AAAGATTTTC ATGATCTACT TGAAAAAATG AGTGTTC AAG TTGGTTCATA CATTCAAGGA 540
 CAAATTATCG TTTCATTCTG TATCGGTATA CTGTGTGTTA TCGGTATTTC GGTATATCGGG 600
 TTGAAATATA GCTTAGTATT AGCTAGTATT GCGGCAGTTA CAAGTGTGT ACCATATTTA 660
 GGGCCTACTA TAGCGATTTT TcCAGCTATT GTAATnGCTG CTATAACaYc GCCGTGGATG 720
 CTCTTAAAT TAGCAGTAGT ATGGACTTTA GTACAATTTG TTGAAGGGCA CTTCAATTTCA 780
 CCAAATATCA TGGGTAAAC ACTTAAGATT CATCCACTTA CAATCATTTT CATTTTACTG 840
 TGTGCAGGCA AATTGCTTGG TATTGTAGGC GTTATTTTAG GTATTCCGGG ATATGCTATT 900
 TTAAAAGTAT TlAGTTACTC ATTTATTCCA 930

(2) INFORMATION FOR SEQ ID NO: 814:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 461 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 814:

GTATGATTGC TGTTTTAATA CCAGATGATG GCAGTGGCAA ATCTTATGAC TATATGCTTG 60
 TGAACCCAAA AATTGTAAGT CATAGCGTTC AAGAAGCTTA TTTACCAACT GGTGAAGGTT 120
 GCCTTAGTGT CGATGATAAT GTTGCTGGTC TAGTTCACCG TCATAATAGA ATTACAATTA 180
 AAGCCAAAGA CATCGAAGGT AATGATATAC AATTACGACT AAmAGGATAT CCAGCAATTG 240
 TTTTCCAACA TGAAATTGAC CATTTAAATG GTGTAATGTT CTATGATCAC ATTGACAAAG 300
 ATCACCATT ACAACCACAT ACAGATGCAG TAGAAGTTaA aACACATTTT cTAATTATCA 360
 AAgCTTAGGA TAAtATGATG tCCTAAGCTT TcCTTTACAA CTTTTCGGAT AACCAACAGT 420
 TAATATATCA CCTTCTAACC AAACTTTTAA TCCCTCATTA A 461

(2) INFORMATION FOR SEQ ID NO: 815:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 573 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 815:

AACGATATAA TAGAATTGAA TAAAGTGGGT GATAATGTGA CTAAAAATGA GATTAGAAAA 120
 TACATTTTAC ATAAAAATGAA GAATTTTAAAT AAAGCTGAAA AGCGAAAAGC AGACACATGG 180
 5 TTAAGAAATC AATTTTTTGC AACTGAAGAA TACAAAGAAG CAAACGCAAT TCGCTAGTT 240
 CTTTCTTTTA ATCATGAAGT AGATACTTTT TCTATTATTG AACAAGCCTT AATGGATCAT 300
 AAACGTATTT TTGTACCGAA AATGGATTAT TTAAATCATC AAATGACTTT TAAAGAGATA 360
 10 TTTAATCTCA AAGATATTGA TGTCGATAAT AAGGGGATTT ACTATCCAAC TTCAAAGGT 420
 GAAACAACGA ATAACCTAGA TTTAATTGTT GTTCCTGGTG TTGGATTTC AAGACGATGGA 480
 TATAGAATTG GGTATGGTGG TGGCTATTAC GACAGGTTTT TAGCTAATTA TCAGACAAAG 540
 15 ACAATAAGCT TATTATACGA TTTTCAAATA ACA 573

(2) INFORMATION FOR SEQ ID NO: 816:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3989 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 816:

TAGTTTAACA ATGTCTATTC TCATAGAATT TGTCCAATA AATGTAGACG ATTCACTAAA 60
 30 CTTCTAAAAA AATAAACCCC AAATTATAAT AGTCTCAATA TTAATATATT ACAATTTATT 120
 CATAGGATTT TTATCATTTT AATTAATAGG TTGTTATCAG GGTAAATAA ACATTTTTTG 180
 35 TAATAGTCCT TTTTACGTT CTTTAAATAA CTCAATTCTA TTCATTTGAT TATTCATTTT 240
 ATTGTCTATA GATTTTAACA ATGCACTTAC TTTATCTGT TCAGTGAGAC AAGGTATTTT 300
 AAGATTIATA TTGCATAAGT TCGCATGAGT TAATTTAGCT GGTGCATTAC CTGTGACAAA 360
 40 TGCTCGTAGT TCTTTAAAT TTAAATAATA ATTCATAAAA AACAAATTAT GATCATTACT 420
 TTTAACTACA TCGCATGAT TATTTACCCA GTATTGCCCA TTAGCAATAA AGCTACTCGT 480
 CTCAAAGTGC CCCCATTG CACCATCTTC TCCTATTAGT AATCGTTCTT CATTATTGAA 540
 45 TAAATAATCT TTTACGTAAT CAATAATTCC AGTTGCACCA TAGTAAGGGT ATAACCCCTT 600
 TTCTCTTAAT GAAGAAGTAA TTGGTTTTCT TCTATTATT TCAAAGATAA AGATGTCTTT 660
 TATGAATTTG TTTTCCCACT CTGGATATTC TTCACCATTC TCATCTTTGA ATCGCAATTC 720
 50 TTGTGTGAAG ATTTTCTGCA TATAGCCTTT TTTCTGTTGT TGAAGCAATT CAAGTTTTTG 780
 TTCTTCTAAT TCAATTTGTC GGTGAGTTT GCTGAAGAAC TTGCCTATTT TTTGCTGTTT 840

55

	TCGACTACCT	CCACTTTGTG	CAAGGAAAAT	TTTCCTTTTA	CCTTTTCTTG	ATAATAGATA	960
	CTGTCCAAAA	AAATTATAAT	AATACTCTTT	TTTCAATCTA	ATAATACATA	CATGTTGATT	1020
5	TAAATTAGCA	TGCGTTTCAA	CTATCGAATT	AATGGCTGTT	CTACCTATTG	ATGCTCCTGT	1080
	AATATTTAAA	AGAACATCAC	CATAGTACGT	TCTACTATTT	TTCATCTCAT	CATCTATATC	1140
	TTTACTAATA	TAAACTAAGT	CATTAAGATT	TAATTTACCA	TTTCTAATAT	TTTGACTCCT	1200
10	TAAAAATGGT	ATGCCTTTGT	TTGTATAGTT	TTCACTTCCA	CCTTTGGGAG	TCTTTCCACT	1260
	ACCTATTTTG	GTAGTAAGAT	TCCCTAACTT	CTTCTCTTCC	CATTGCGCTT	CAAATCCTGG	1320
	GAATCTCAAC	TCTGGCACAT	TTTTCGTTTG	TGTATTACTC	ATCTTTCAAC	ACCCCAAGTT	1380
15	CTTTTCAGGTA	TGCATTGATT	TCTTGTTCAA	TTTCTGCGAT	TTCTTTGTGCG	ATATTTTTTCA	1440
	AATCTTGTTG	GACTTGATCT	AAATCAATTG	GTGCTTCTTC	TTCGAATGTA	TCAACATATC	1500
20	GCGGTATGTT	TAAGTTGTAA	TCGTTATCGG	CGATCTCTTG	TAATGTCGCG	CTGTAGCTAT	1560
	ATTTATCAAT	CGTTGCTTTA	CGCTTATATG	TGTCTATAAT	ACGTTGCGACT	TGGGCATCGC	1620
	TTAAATGGTT	TTGATTTTTT	CCTTTTTTCAA	AATCATTGGA	TGCATCGATA	AATAGTACGT	1680
25	TGTCGTCTTG	TTGGCGACAT	TTTTTAAATA	CTAAAATACA	TGTTGGAATA	CTTGTCCCAT	1740
	AGAAAATATT	GGCTGGTAAC	CCAATCACGG	CTTCTAAGTA	GTTCTTTTCT	TCTATTAAAT	1800
	AGCGACGAAT	CACACCTTCT	GCGGCACCAC	GGAATAATAC	ACCATGTGGG	aGTACGACTG	1860
30	CCATGGTACC	TTCATCGTCT	AGGTAATGTA	CCATGTGTTG	AATAAAGGCA	AAGTCTGCTT	1920
	TGGaTTTTTG	CGCAAgCTTG	CCGTAACCAC	TGAATCGTTC	ATCATTTTCA	AATTTTGAAT	1980
35	CTGCTGTCCA	TTTCGCACTG	TATGGTGGGT	TCGCAATAAC	CGCATCAAAT	GTATTGCCTA	2040
	AAAAGGCTGG	ATTTTCCAAT	GTGTCATCAT	TACGGATCTC	GAAGTTCTCA	TAACGCACAT	2100
	CATGTAATAA	CATATTCATG	CGTGCTAAGT	TGTATGTAGT	ATTGTTACGT	TCTTGTCCGA	2160
40	AATAACGATA	CACCTGCGTT	TCTTTACCAA	CACGTAACAA	CAATGAACCG	GAACCACATG	2220
	TTGGGTCGTA	CACGTGACGT	AATTTATCTT	TACCGTCTGT	GACAATCTTC	GCCAGTATCT	2280
	TAGATACTTG	TTGTGGTGTA	TAGAACTCGC	CTGCTTTTTT	ACCCGcTGTC	GCCGCAAAGC	2340
45	GCCCCGATTAG	GAATTCATAT	GCATCACCTA	ACATATCAAT	TTCCATGTCA	CTGTGAACGA	2400
	ATGGTAAGTC	GTCAAGATTA	ACCATGACTT	TAGAGATTAA	AGCAGTACGT	TCTTTGACAT	2460
50	TGTTACCTAG	TCGCGTTGAA	CTCAAATCCA	TATCGCTGAA	CAGCCCGATA	AAGTCATTTT	2520
	CACTTTCTTC	ACCTAATGTA	GATGTTTCAA	CTTTGCGAAT	CGCCGTCGCT	AGATGTTCTA	2580
	TATCGAAATC	TTGCGTTTCA	ATTTCACGAA	TCATCGCACT	GAATAAATCT	TGTGGCTCAA	2640

55

	CCCATGCTTC TTGATACGTG ATGTCTTCAC CTGACAAGGC ATCTGCATAT TCTTGTTCCG	2760
	CTTTTTCAGA TAAGAAGCGA TAGAAAATCA AGCCTAAAAT GTAATTACGG AATTCACCTG	2820
5	CATCCATGTT CCCTCTTAAA TCATTGCGAA TCGACCATAA TTTTATATGT AATTCAGCTT	2880
	GTTGCTGACG TTGTTTTTCA GTAATAGACA TGTGATTCCT CCGCCTTTGC ATAAGTAATT	2940
	TATCTCTTTG TGTAATAGAT TTATTATAAC ATTTGGTTAT GTTGCATGT TGATAATTTG	3000
10	GATGTTGGTG GTGGAAATTT TGAGTTTTAG TGGCGCAATT GTTATTGAAA AATTTTATAG	3060
	AAATGTTGTA GCTTTCAAAT GCTTTCAAAA TCATTTATAT TCTTAATGAT GTCAAAAAGT	3120
15	TGTTCAATTCA TACATAAATA AAACCAATCA ACAATTGAGT TGGTGAAAAT CAATCGTTGA	3180
	TTGGCTTTGA TGCATATTAA ATAATGCAAT ATATATTAAT AAATGTTAGT TATAGTATAT	3240
	TTTGTGAGGA TTGGGTGAAT GTCTAAGTTT TAATTATTTA TCTAAATTAT CTGCAATGAA	3300
20	TTTCTTAATT TCAGGAGAGA AATAACAGC AAATCCTCTT GTGCTTTCAC CTGATGGCTT	3360
	ATTACCGGCA TAGATTACAC CAATAGCTTC GTGTTTACTA TTAAATATAG GTGAACCAGA	3420
	GCTACCAGGC TGAATAATTG CATCCGATGA CACTATATTC CCATTCACTG ATAATACTTT	3480
25	ACCAGTTGAT TCATACATTT GTAGTTTATT TCCATTAGGA TTTGGATAAC CAATGACTGA	3540
	TATAGGTTCA TTTTCTTTAG CTTCTGATGC TATATTAAAT TTACTAGTGA AATCTTTGAA	3600
	TTTTCTACCT TTTGGTTGTG TTGATTTTTT TCAACTTGT ACAACCGCAA TATCTTCTTT	3660
30	ACCAGGATAA TCTACAATCT TAGTAACTTT ATAAAGTCCA CCACCGTTAT TATAAAAACC	3720
	ATTAGGATGT GCTTTGATTT CATCACCAGC TTTCATGTGA TAGGTAACAT GTTTATTGGT	3780
35	AATGATTGTA TGATTTCCAA CTACAAATCC TGTTCAGCG CCCATCCATG TAACACCACT	3840
	GATGGTGCA ACATTTGTAT TTGTAATTTG TTAAACAGTA TTTTCGGCTT TGGCTGTTTG	3900
	TTGAATACCT TCAACCATTG TTGTCCGAC ACCAGTTATT GATGTTAAAA TCGTCAATGC	3960
40	TGCAATACTT TTGATGATTA TATTTTTAT	3989

(2) INFORMATION FOR SEQ ID NO: 817:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 904 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 817:

CGTTGAGCAC AGTTTTAGAT AATAGATAAT CTGCTCTAG TTGTTGCAAT GTCTGCGATT 60

TATGCGTGTA ACGTTCAGTT ATCGTTTCTA AGGTATACTG AGATGTATTT AATTGATATC 180
 CTTTAGATTT AATACTTACG ATGATATCAT CCATAAAATT GCTATTGATA ACATGGATAT 240
 5 CATTGCCGAAC TGTGCGGTTT GAAACATTGA CATGTTTAGC AATTTCAATTA GAACTAATGT 300
 GCTTTGATGG ATTTTAAATA AAAAAGTGA GTAGTTTAA GTGTCTATCA AGCATTAAAC 360
 ATGTACCTCC TTTCTAAATT TTTGCTGTAA GCGTTTAA GGGTGTATTA ATATTATTAA 420
 10 ACATGAGAGC TTATACATAC GTCAATGACA TTAAAGCGAA CTTTATATG ATTTTAACAG 480
 AGTGCGAATT ATGCAAATAA AGAACAGCAG TAAGATATTT CAAATAGAAA AATATCTCAC 540
 TGCTGTTTTT CTGAATTTAT GCATCTGGTA CTTGTGGACG TATCAGGCAA ATGATTAATT 600
 15 TTTAGGTGAT TGTGCTTGAG GTGTTTGTGTT AGAAGGCGTA TTGGTATTAT TTGATTTATT 660
 AGCAGGTGGT GTGTTCTGTT GTTGATTATT TTGGTGTGTT GTTGACGATT GACTGCCACC 720
 20 ACCGTGAGTA TTATTTTAT TTGAATTTGA GTCTTTGGC TGAGCTGGTT GCGTATTGCA 780
 TGGTACTGTT GAATGGTTGT TAGGAACATT TGACGGTGAA GATGGTTGAT TGTCATTAGT 840
 AACGTTATTG CCATTATTTT GATTCTTATC TGTTTGTCT GCACTATCAT CTTGTTGATC 900
 25 ATTT 904

(2) INFORMATION FOR SEQ ID NO: 818:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1104 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 818:

TTTAACatGa TAATAAAAAA TCITTTGTGA TATCATTAGG AATATTTGAT GAACTTGATA 60
 40 TTAAGGTTAC ATTTTGAGAA ATGGACTTAG GGATATTTCC ACTATTATTT AGTAAGAAAT 120
 CTTTAGCTAA TAAAGATTTT CCTACACCAT TTTACCAAC AATGTGATTG ATCTGACCAA 180
 GATAGAAATT TAAATCACAA TTGTCAACTA GTTGTGTTGTT TTTAACTTTT AAAGAATAGT 240
 45 TATTTAGTTT CATGTATACA ACTCCTATGT ATAAAGGGAT TTATTACACC GATATTTAAT 300
 TGTATTTTAA AAAATTCtTT CACATTATGT ATAGAwGTTA TAAATTAGTA TATCaCACTA 360
 TATTTtGkCT AAATGATAAA TATATCGTTA TATTTTACA ATATTCTGAA ATTTATGTTT 420
 50 GCCTCTGAAT GCTATATCCA GTGTAATGTG TTTTGCATAT ATGAAAGCAA TTTCAAAATG 480
 TGAATATAGG TTCATTGTGG TATGACAAAC TTCATTGCTT GTCATGAGAT GGATATAATG 540

GACGGGACGT TTTTAGATTG AAAAAAGACA TATGATAAAC TTAGATTTGA AGCGATTTTT 660
 ACTGAACTTA GAAATAGAGA TATTACATTT ATTGCTGCGA GTGGCAATCA ATATGCGAAG 720
 5 TTGAAGTCTA TTTTCGGGGA TAGAGATATG TATTTTATTT CTGAAAATGG TGCAGTTATT 780
 TATAATGGCA ATGAGTTATA TAATTATAAA AGCTTTAATC GTCAGGTGTT TCAACAGGTT 840
 GTCGATTACT TAAATATGAA GCAAAGTATT GATCAACTCG TCATCTGTGG TTTGAAAAGC 900
 10 GCGTATATTT TAAAACATAC TTCTGAAGCG TTTAAAGAAG ATACGAGATT TTATTATCAT 960
 CAGTTAAAAG AAATTGACAG TCTACAGCAA TTACCTGAGG ATGATTATGT CAAAATAGCA 1020
 15 TTTAATATTA ATCGTGAGAC GCATCCGAAT GTTGACGAAG AAGTAGCAAC GCAATTCAGC 1080
 AATGATATTA AACTTGTCTC AAGT 1104

(2) INFORMATION FOR SEQ ID NO: 819:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1425 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 819:

CCCnTTTTAC GGATTAAngG CTTTTTCCTA TTTAAAACCT ACGGCATTTT CTTTTCAACC 60
 30 ACGGCGGCTC CCATATGGAT GGTATnGGGA TTGGGGTTTA TAAAATGGGG ATTGGAAATG 120
 GTCCATCCTA AATACTCAGT TTAGTGCTTA TTTCTTTAG TGCTGACGAA TAAATATGAT 180
 35 CTAACCTAC GAAACCTAAT AACGTATCAA ATGCTTCATC ACCATTTTCA AATTTACTTA 240
 AACTATTTTT AAAATCATGT CGCAAATCCT CTAAATATAG TGATGGTTTT CGATCTACAT 300
 ACATTTTTAT ACTAACCTCC GATATATAAT CATCTTTATT GTACCTAACA TTTTATTAAG 360
 40 ATAATACTAA TACACTATCG AACTTTGGGC TGATACTGAA ACAGCACAAA GAAACTCAAA 420
 CAATTCGAAT TATGTATCAA AACCTTCAAT TACAGGAAGT GATTTTAAAA TGATTTAACG 480
 CAAAAAACA CCTGTTACCG TTATATAGGT ACAAGTGCTT AATTGATAGA GATGTTATAC 540
 45 GTCTTTAAAT GATTCCACAA CTTTTGGATG TGGACCATCC ATAAGCGGTT CTCTTTGTCG 600
 AACGCCACCA CTTTGGTTGC CAATTGATTC GCTATCAAAG TACGATTTAT CTTGATTTGA 660
 TTGTTCTTGA ATATGTTCTT CATTATCAGT TGTTCATAT TCACTATCCA CCTCTGTTTT 720
 50 TTCCATTGTT GCTGTATGGA ATTGTACGAA GTTCTCTTCT TCTAATGCTT TGATTTCTTC 780
 TTTTGATAAC GCTCGATACC AATCTTTTCG TTTTTCGCC GCAaTAGGAa CAaCAcCTTT 840

55

TACTTCAAAT TCTTTCCATT TTTCGGGGTA ACCTTTCATT GTAAAGGGCA TTCCCTTAAC 960
 CTCCAATATG TTATAATTCA TGTTATATAC CACCATAATT ATAAAATGAA ACATATTAAA 1020
 5 CACAATTTTT TAGACTCCCG TCATAATAAC TTCAGACAAA CGTTAGTCTG ACCTGATAAA 1080
 TATTTAATTT TAAACGTTCT TGTCTCTCTC ACGAAAACCA TTGTGTTGAA TATTCATCAT 1140
 TGCTGATTGA ATATAATTAT ACTGTGGTAA ATCTGGTAAT ATCTTTATTT CTAATTCTGT 1200
 10 ATTTAATTCG AAATGCTTAG CGATGTTTTT AAATATTGCT AAATACTCGC CCATTAATTG 1260
 TTCATTTATT GTAAGTCTAT CTTCATTAGC CATGGCTCTA TTAAACATAA AGCTAATTTT 1320
 15 TTCTAATGCG AATAAGCTAG GATAATAATT TTGAATCAAT GTCTTATCAC TAAATAATTC 1380
 ACCATTAGCT GCATTATAAA CTTGCGTGCA TGTTATTTAA TTTGC 1425

(2) INFORMATION FOR SEQ ID NO: 820:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1165 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 820:

nCGnCCCTTAT ATATGTTTTT CATGTCCTAC AAAAAACGAA ATATTCCAAT TGCCTATATA 60
 30 CAGATATTCA TTAATGACAC ACTCAATAGC AACCATHACA ATATTATGTT ACCTATTTAA 120
 ATGAAATAGT TCTTTTAAAG AAATACATTT TTCACATATT AATCTATAAT CAAAATCAAC 180
 35 TGACCGATAT TCTATAATTT ATGATTAAAA TAAGTTATAA TATAATAGTA AAGATAAAGA 240
 TAGAGGTGGC TATAATGTGT GGACTTAGAA GTATAACATT AGGTACaACA AATATAGAAC 300
 AGACaAAACa TTTTCATGGTT GACATATTAG GATTAAATTA TGAAGAACTT CTTGAAAAC 360
 40 CAATTCGTTT CGGCGATGCA GATATAAGCC CAGGAACAAG ACTTCAATTT ATACAAGTTC 420
 CAAGTGAGCA ATTAGAAGAA TCTCACTTTG TGGGTATTGG ATTACGTACA CCAACTGACT 480
 CAGGTTTAGA GGAGTATGCG GAAATATTAT CGAATAAGGA TATTCCATTT ACAACAGTTA 540
 45 AAGAATTAAA TGGCAATAAA TATTTTCAGTC TCGAAGATAA CAATGGTCAT ATTTTCTCAA 600
 TATATTCAAA CGAGAATAAT TATGGCGTTG GTTTAGGTAT GCCTTCTTyT GAGAGTGCGG 660
 TCAATCCGTT ACATCAAGTG CAAGGTTTAG GACCAGTGAT TCTTAAAGTG AATCATGTAG 720
 50 ATATTACAGG TCAAATTTTA ACAAATATAT TCGGACTTGA AGTATTTGCA GAATACCAAC 780
 CCTTCGACAA TGCTGACTAT CATGTCCAAG TATTCAAAGT TGAACGGGT GGTCTAGGTG 840

55

ATCAAGTTGA GTTTGAAACG AAAGATGCAG ATTTCTTTAA TCAAGCGAAA TCGCGCTTAG 960
 ATGAAGTGGG AATACCATAT CAAACGCTTG AGCAAGATGA TATTGAATCA ATTAGAATTA 1020
 5 CTGAAAACAG TGGATTATCG TTTATATTCA CTTTACAAAA ATAATTTTTT TACGATAGCG 1080
 AGGACAAAAT TTATGTTACA TGAACTTGG AAAGAACGTA CACCAATCAA GAAAGTAGAA 1140
 GTCATTAATA CAGATGCAAA GAAAT 1165

(2) INFORMATION FOR SEQ ID NO: 821:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 821:

TCGCCCAATT ATTATGAAAT ATTCAATACA GTTTATTGAT CAAAAACAA AAATCCATTA 60
 TGAACCTTGT CATCACGAAT ATTAAGTAC GTTAGATGGA CCTCTTTAAT GATAGAACAA 120
 25 TTAATAAATA ATGCACTTAA GTATGCGAGA GGTAAAGATA TATGGATTGA ATTTGATGAG 180
 CAATCCAATC AATTACACGT AAAAGATAAT GGTATCGGTA TTAGTGAAGG NACTTGCCTA 240
 AAATATTTGA TAAGGCTAT TCAGTTTATA ATGGCCAGCG CCAAAGTAAC TCAAGTGGGA 300
 30 TTGGTTTATT TATCGTAAAA CCAATTTTCA ACACACACAA ACCATCCGTT TCCGTCGTAT 360
 CTAACAAAT GAGGGTACAA CCAATTACGAA TACCATTTC 400

(2) INFORMATION FOR SEQ ID NO: 822:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 760 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 822:

TGATATAATA CTTTGTGAAA GAAAGCATG TGTGGGAGGT ATGACCTGTA TGTCGAACGA 60
 AATACTTATC GTAGATGATG AGGATAGAAT CAGAAGATTA CTTAAATGT ATTTAGAAAG 120
 AGAATCTTTT GAAATCCATG AAGCAAGTAA TGGCCAAGAG GCTTATGAAC TTGCAATGGA 180
 50 GAATAATTAT GCTTGCATAC TACTAGATTT AATGTTGCCT GAAATGGATG GTATCCAGGT 240
 GGCAACTAAA TTGCGTGAAC ATAAACAAAC ACCGATTATT ATGTTGACTG CTAAAGGTGA 300

TTCACCAAGA GAAGTAGTCT TAAGAGTTAA AGCACTTCTA AGAAGAACGC AATCTACAAC 420
 TGTAGAACAA AGCGAACCTC ACGCACGTGA TGTGATTGAA TTTAAACATT TAGAAATAGA 480
 5 TAATGATGCA CATCGCGTAC TTGCTGATAA TCAAGAAGTT AATTTGACTC CTAAAGAGTA 540
 CGAATTATTA ATATATTTAG CTAAAACACC AAATAAAGTA TTTGACCGTG AACAATTATT 600
 AAAAGAAGTT TGGCATTATG AATTCTATGG TGATTTAAGA ACAGTTGATA CTCATGTAA 660
 10 ACGcTTAGAG AAAAGTTAAA TCGTGTGTCT AGTGAAGCTG CGCATATGAT TCAAACAGTC 720
 TGGGGCGTTG GGTATAAATT TGAGGTTAAA TCTAATGATG 760

(2) INFORMATION FOR SEQ ID NO: 823:

15

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 465 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 823:

25 CAAAGGAAAT TGCACAATTA GAAGACCGAT TACGTTACG CTTTGAATGG GGGCTAATTG 60
 TTGATATTAC GCCACCAGAT TATGAACTC GAATGGCAAT TTTGCAGAAG AAAATTGAAG 120
 AAGAAAAATT AGATATTCCA CCAGAAGCTT TAAATTATAT AGCAAATCAA ATTCAATCTA 180
 30 ATATTCGTGA ATTAGAAGGT GCATTAACAC GTTTACTTGC ATATTCACAA TTATTAGGAA 240
 AACCAATTAC AACTGAATTA ACTGCTGAAG CTTTAAAAGA TATCATTCAA GCACCAAAAT 300
 CTAAAAAGAT TACCATCCAA GATATtCAA AAATTGTAGG CCAGTACTAT aATGTTAGAA 360
 35 TTGAaGATT CAGTGcAAAA mAACGTACAA AGTCAATTGC ATATCCGCGT CAAATAGCTA 420
 TGTcTTGTcy AGaGAGCTTA CAGATTTCTC ATTACCTAAA AATTG 465

40

(2) INFORMATION FOR SEQ ID NO: 824:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 355 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 824:

50

CACCGCGGTG GCGGACGCTC TAGAACTAGT GGATCCCCCG GGCTGCAGGA ATTCGGCACG 60
 AGGTAAGGAG GTCTCTGTAC CATGGCTCGT ACAAAGCAGA CTGCCCCGAA ATCGACCGGT 120

55

GGAGGGGTGA AGAAACCTCA TCGTTACAGG CCTGCTACTG TGGCGCTCCG TGAAATTAGA 240
 CGTTATCAGA AGTCCACTGA ACTTCTGATT CGCAAACCTC CCTTCCArCG TCTnGTGCGA 300
 5 GAAATTGCTC AgGACTTTaa AACAGATCTG CGCTTCCAnA GCGCACTATC GGTGG 355

(2) INFORMATION FOR SEQ ID NO: 825:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 1099 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 825:

TTCACCTGGC TTGTTGactG ACTTGTAatAT GATGATGTGC TTTGTGAATC GGATTcGCTC 60
 20 GTGCTTGAC TTGTTGAGTT TGAGGCACTT TGGCTTGCTG AGTTTGAGTC TACTCCGCTT 120
 TGATTcATTG AGGCACTTAG TGACAATGAT GTACTCGTTG AGTCAGACAA ACTTGTACTC 180
 GTTGACGTAC TTGTACTTCC TGATGTTGAT TGAGACATAC TTATGCTCAT TGATGTTGAA 240
 25 TCGGATTTAC TTTCACCTGA TGATGTTGAG TCGGATTCAC TTTCACCTGT AGAACCCTT 300
 AATGATGTGG ATGTACTAAT GGAATCAGAT TTACTTGATC TGATTGAATC ACTTGTCGAC 360
 ATTGATGTAC TTAATGAATC AGACTTACTA TCACTTGTTG AATCACTTAA TGATGTTGAC 420
 30 AAACCTGTAG AGTCAGACAA ACTTGTACTC GTTGACATAC TCAGTGATGT TGAAACACTC 480
 TCGCTCTTAA ACGTTGACGT TGATTCACCTG ATACTTGTCG ATGTTGAAAT GGACGTACTA 540
 CCACTTGTTG AATTACTTAA TGATGTTGAT GTGCTACCAG ATTCTGATGT ACTGTCTGAT 600
 35 AATGACGTAC TCTCACTTGT CGAACTACTC ACTGACTCTG ATGTTGATTC AGACGTACTT 660
 TCACTTAATG ATTCACCTAA AAAGGCAGAT GCACCTTGTG ATTCTGAATC GCTAGTACTA 720
 TTTGATTCAC TTAATGATAT AAACGTGCTC TCTGAAGCAG ATATTGCTTC ACTTATAGAG 780
 40 TCGCTCGTTG ACGTTGATTC ACTTATTGAA TCAGACTCTG ATGTACTTAA GCTTGTTGAA 840
 TCACTCaTAG ATGTTGaTGT ACGTTCTGAA TTACTTGATC TTAACGATGT CGAAGTACTT 900
 45 ACCGAACCAG ATGTGCTCGT AGAAGCACTT TgTGATATTG ATTCACCTGA TGCAGTTGAT 960
 GCTGATTTGC TATCACTCGT TGAATCACTA AACGACGTTG ATATGCTCaT TGAATCGGAT 1020
 TGACTTGACAC TCantGAACC AGACGTACTT TGTGATTCCG AAGTACGTAC TGAAGCACTT 1080
 50 GTCGACGTTG ATGTACTTG 1099

(2) INFORMATION FOR SEQ ID NO: 826:

55

(A) LENGTH: 338 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 826:

10 TATGCGTATC CTGATAGTCA CTTTGATTTT GATATGGAAT TAGCGAAAGA GCAATCTCAA 60
 GACAATCCAG TTTACTATGC TCAATATGCA CATGCGCGTA TTTGTTCAT TTTAAAACAA 120
 GCGAAAGAGC AAGGTATTGA AGTGACTGCT GCGAATGATT TTACAACGAT TACTAATGAA 180
 15 AAAGCGATTG AATTGTTGAA AAAAGTAGCT GATTCGGANC CTACAATTGA AAGTGCTGCT 240
 GAGCATAGAT CGGCACATAG AATTACTAAT TATATCCAAG ATTTAGCCTT CTCATTCCAT 300
 AAATTCTATA ATGCTGAAAA GTGTACCAGT TGGTATTG 338

20

(2) INFORMATION FOR SEQ ID NO: 827:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 838 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 827:

30

GTATTTTCTT TCGCGAATGA TTTTATTACT TGAATACCAC GAATAACCTC CAATACCTTT 60
 TCCACTAATT GGTTTTGTAC ATTATGATAC GCTGGCGCAT TTTGTCGACT CTTTCTTTCT 120
 35 AATAATTGAA TCGCAAAAAA TGATAGTAAT ACGCCAATGC ATGCTAATAA TGATACTTGC 180
 CACGAACTA CAAGTAGAGA CAATATGAGT ACTGTAATTA ATATGTATCC ATTAACAACA 240
 ACGTCCACCA TTTTCATAGC AAAGTTTCT AAAAAGGTTA AATCTGTTGT TACTATTGTT 300
 40 GTTAACTCAT TTGAATGATG CGAATTAAAA TAACCTAACC TTACATTTT CAATTATCC 360
 CCTATATCTA AACGTTCTTT CGCACTCATT TCATAAGCGA TGCTCTCATG GCTTTTGCTT 420
 TTGAAATATG CTGTAATAAA TCGTCCAATC ACTAATAAAA CCATGATAAT TACAACATTC 480
 45 AATATATCTT TCATATAAAT AGGTTTATGA GATAGCACAT TATTAAATAT TTTTGCAGCT 540
 AAAAAGATAG GTAACGCAAT AAAAATAGCA TTTAAAAATG ACATGCTAAA TCCTAAAATC 600
 50 ATTCTTGCTT TATATGGTCT TATCCAGTTT AAAATTTTAA ATGTAATTTG AAACATATTT 660
 GACCTCCTTT TCTCATTAGT TTAATTATGT CCAGTATTGA TTCCCAATC TTTTGTATGC 720
 ATGTGCGTAT CCCACATTTT CTTATAATTA CCGTTTAATT TTAGCAATAA GTGATGTGAC 780

55

(2) INFORMATION FOR SEQ ID NO: 828:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 898 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 828:

TCCnAAAGAA GAAATAaCAT TATCATaAAA TCCATTGAAT ATAAATTCAG TCATCCCCTC 60
 TGAACTtCTC CATAAATATA ATGGTGAATA AGTATTTGTA ATGCAATTAG TAGTTGTTTG 120
 ACTTATTAAA TAAGCTTTTA TAATTAAATT TTTAAATCCA TCAGTCTTAT AACCATTTAA 180
 TCGAACTCTA TCTCTAATTA TTTCATATT GTAATCACTA GGCAACTTAA CATGATATTG 240
 CATTGCATGC ATTTAGCATA CCCCCTTTTA TAAAAAGGAT AGCAATAATA AGTAAAATCT 300
 CATATTATCC AATTGTGATA TAGTTATCAT AAAAAGTGAT AGGTGATTAA ATTGAACTTT 360
 AATGATTTGG AAATTTTAT AACTGTATGT GAAGAAGCAT CTATCAATAA AGCTGCAATT 420
 AAACCTAGAT ATGCACAATC TAATATATCT CAAAGAATTA GCAAGCTTGA AAATGAATTA 480
 GGTGTAGTTT TGCTTTTTAG AAATCAAAAA GGTGCTAAGG CAACTAAAGC AGGCGAAGAA 540
 TTCTTAGCGT ATAGCAAAAA AGTATTAAGA GATACAGAGA CTATAAAAAA TAAATGAAA 600
 AATAATACTA TGTCTATTTT ATGCTCAGAA CTGTTATTTA ATTATTTATC TGAGAGCGAA 660
 GAAATTATGA TGTCAATAA CTCAATTAAT TTTATTTCTA GTGGAAATAT TAGAAAAGCT 720
 ATAGAAAAAA ATAATTATGA TAAGGTTATT TCATTCTATAA AAaTTAACGA CTCAAATTAT 780
 AGACTTAGTA ATGTTGATAC TATGaaAGTA ACGCTTTACA GTAATGGAAG TAATTATGAT 840
 AAAGAGGCTT TACTAATAAA TAAAGATGAG TTTGGTCCTT TAAGGnAAAT AACTTTAG 898

(2) INFORMATION FOR SEQ ID NO: 829:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 829:

AGTAGGAGTC ATAAATCGn ATTTCAAAAG CAATTAATGC AAAATACTGA GTCTACAGTG 60
 CTATTTGCCA AAGCGTccAT ACGAAgTTGA GcAATGCTAA TAATAATGGT CTATCAAAAA 120

AATATGTTGA TAGTTACACT GATTCTCTTA CTGGAGTAAC AACTTCTGCT TTTTAAATA 240
 AAGATACAGG CAAAGTAACT CTCGGGATGA CTGGGACTAA TTTACAAGAC GAAGCCTTTA 300
 5 AAAAGTTAAA AGAAGGTGAA TTTTCAAGAC AAAATGTTAC CAATGCTTTG GAAACAGTTA 360
 AAGATGGATA TGCAGATCTT AAAATATTAT ATTCTCCTGC ATCTGATCAA AACTATAGAT 420
 ATGCGAATAC ACAAGAATTT ATAAATAAAA TAAAAAGTAA GTATGACATT GATTTTATTA 480
 10 CTGGACATTC ACTAGGTGGA AGAGATGCGG TAGTTCTAGG AATGAGTAAT GGTATTCCGA 540
 ACATTGTGGT TTATAATCCA GCTCCTATTT CTATAACTAG TTTGAATCCT AATTCCCCAG 600
 ATGGAACG TTTATTAGAA TTATATAAAA ATTATAAAGG TAATATTACT AGGTTTGTG 660
 15 CAGAAAATGA TGCATTGACA GAAATCTGA AGAAATATAA GCATTATGTT TTTTTCGGTA 720
 ATGATAAAGT CTTTAAAAAT GGTAAAGGTC ATGAAATGKA AGGCTTCTG ACCGAAGAAG 780
 20 AACAAAAAGC tATAAAAn 798

(2) INFORMATION FOR SEQ ID NO: 830:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 830:

AATATCAATC TCTTCATAAG CTGAATTATT TTCATGCACT TCTTGATGTG ATGATTTGTC 60
 35 ACGAACnGCT ACAACTAACA TTTTATCGTC TAAATAAGT TGTTTATATT TTTCTAATTC 120
 ATCAGGCGCT AAGTTGTAGC GTGATAAAAC TGCATGTTCA CCATCTTCTC CTGTAAACAG 180
 TTAGTCATT CTATCACTAA ATGTTCCACT TGTTGAGATA AGGGGAGATT TCAGAGTCGT 240
 40 GTAAGTCATT AGGTGTAATT TACTTTTACT AATAATTGTT AGCCTGATCT AAATAACTTC 300
 AGATTCTTTG ATTGATAnGT GnATAATCGC AGTGCATTAC ACAGAATACT GCCAAGTGCG 360
 CCCTTAAAT TGTTATTAnC TTACCTCTAT ATAAGAACCC 400

45 (2) INFORMATION FOR SEQ ID NO: 831:

(i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 475 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

AAGTCTGCAT GACTTTTATT TGCAAGTTTCG ACTGCTTGAT CAAAAGCGCG ATGATCTTCA 60
 GGATAGCAC TTTGAACTGA GCTGAAATTC GGATCAGGTT TACATTGTGC TTCTACAAGA 120
 5 TTA AATTGAT TGAAATTCAA AGATTGTAAT AACTCAGGTA CAATCGGAAC ACTTGTACCA 180
 TGCAAACTAG TGAACACAAC TTGTAAATCA GATTTAGGAA TATAGCCAAT CATATTTTGA 240
 ATGTGTTTCA TATAGTCATC AGTTACAGAT TTCGGAAAAG GCTTGaTATA AGATGTATTT 300
 10 TGTTTAGAAA LAGGTATATC AATCTGTAAT GGATCGCCAA CTTCTTCGAT ATAACGACTT 360
 GCAAGCTCAG ATGCATCAGT CGATAATTGC GCACCATCAG AACCATATAC TTTGATGCCG 420
 15 TTATAGTCTT TCGGATTATG ACTTGCTGTA ATCATAATGC CAGCAGTAGT ATTAA 475

(2) INFORMATION FOR SEQ ID NO: 832:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 724 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 832:

CTTGAATTGA AGGATACGCA AATGATGCAT AATCAGCTTC TGTATAAAGC TGTAATGTAA 60
 TGTCATCAAG GTCACCTTTT CTAACAAGCA CCTTATTAAT AGAAGTATGA TTCGCTTGCC 120
 30 AAGTACCTTG ATTATTTTGT TCTAAATGAA TGACTTCGCC TAACGATTTC AACGTAATAT 180
 CTGCACGCTC GTCCTCGCTA ATAGTATATG TCTTACCATC TCGCAAATTG AGCATCTTCA 240
 35 ATTGTTTGTT ATATTTTATA ATCAATTTAT GCATTGTCTT TGCCTCAGTC CTATACTATT 300
 TTTTCTTTC AGCTTCTTGG CGTTTTCTT TATCTTTTG TGCTTGTTCT TTTTGTCTT 360
 TTTGTTCTC TTCTTGTTGC TTTAATTTCT CATCTTTCGC TTTTGCTTTC TCTTCTTCAG 420
 40 ATTTGCTTTT TTCATCTTTA ACTTGTTTTT CTTTGCTTAA AATATCTTGC AATTTATCGT 480
 TATACTTTTT CGTTTCTTCA GAACGTTTAT CATTCGATAA ATCTCCGTTA TTTTAAATCT 540
 CATTTAATTT ATTAATCAAC GCTAACTTTG TAATATCGTT ATCATCTAAA TAAGTGGCAA 600
 45 TATTAATCGC TTCATCAAGA TGTCCTTGTC CTAATCCAT CCAATATAAT AAGTAGTCTT 660
 TGTTTGAATT TGGTGTCACT TTATTAAGTA AATTTTCTTT CGTATCTGTh TCTAAACChT 720
 50 GTTh 724

(2) INFORMATION FOR SEQ ID NO: 833:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 994 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 833:

	AGCAGTTTGG CTCGTAGAAA TCTCCCGTCT CTATTCTATG TATTAAC TTT TATTATATTT	60
10	GTTATGATAC TATCGATATG AAAGCGTTGT CAATGGTTTT TGTA AAAAATT TTGTCAAATT	120
	TAATTTTCA AGTCTTTATA AAAGTAGTTT AATTTGAAAA CTAGAAATAC CAATCCTAAC	180
	TAATTTATTT ATGACGCTCT TTTGTTTAAT TGACATTCTT CAGCCATATT TTTTATGACT	240
15	AGTTGCATTT TTTACTAATA AACACCTCTA AAAC TTTAAT GATTTTAATC GTTTTAGAGG	300
	TGCTAATTAT TTTATTTGAT TATTTTTTGT TTGATACCTA CTGCATATCC CATATGAAAA	360
20	CGGCTTTTTT ATTATGTTAT ATGACTAAAT CTCGTGAAAA ATGAAATTTT TGCAGACTTA	420
	TGATTTACCA AAGTTTATCA TAACTAGTAG TTACATATTT CGCTCCATTT TTAACAGCTT	480
	CATTGACTTC ATCTATTGTA TTAATTAGGC CACCTGCAAT GACTTGTGTG TTTGTTTCTT	540
25	TCTGAATATG ATGAATCGCT TTA CTGCAA CACCTGGAAG TACTTCAACA AAATCAGGTT	600
	CAACTTTTTT TATCAAATCT ATACTGCGTT TCAATGCTTG ACTATCAATA ATAAATACTC	660
	TAAAAATCGT TAAAGTATTT AATGATTTAG CTTTTTTTAT TACTTTAGAT TTAGTCGATA	720
30	CGATACCTTT TGGCTTGAT TGCTGAATAA TAAATTC ACT TGCAAATTCA TCGTGGCTTA	780
	AACCTTTTAT CAAATCTATA TGAATAAAAC ACTCTATATG ATTTTGCTTC AGCAATTCCA	840
	TAATACTTTT TATATGTCCT ATATGCATAT CTAGAAGCAC ACACATTTTA TAGTCTGTAT	900
35	TAATCCAGTT TCTCTAAATC NTTAATGTTT CCTATAAGCA GGCAATATGT TGTTAATTCA	960
	CTnGATCCAT CCTCTCTACA TCACACGCTT AAAT	994

40

(2) INFORMATION FOR SEQ ID NO: 834:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 783 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 834:

50

	TCnACAAAGT CGGTATTAGG CTATGGGCCA TTTnACTATT ATA ACTAATC GGAAATATAC	60
	CACATAACAT CATTATTGAG TTGATTCTAT CATTTGGCTT ATTAGGGTTT TTTATCATAA	120
55	TGATTGTCAT TTTGCTACTA GTTTATAAAA TGATTAGGAA CTATGATCCA AACACTATAG	180

5 ATTTAGTTGT AAGTGAATTT TGGTTTGTGT TGTTCATTT TATTACAAAA GGACGGCGTC 300
 ATCATGGCTA AGAAAGTTTT TATTATGGAT AGCGTAAAGA CAATAATTGG TACGTTGCTT 360
 10 ATAGCTTTAG GATTACAATT TTTAGCTTAT CCAATTATTA ATCAACGAGT AGGTAATGAA 420
 GCGTTCGGTT CTATTTTAAC GATTTATACA ATAATAACAA TCACGAGTGT TGTATTAGGC 480
 AATACGCTTA ACAATATACG mTTGATTAAAT ATGAATCTAT ACAAATCCAA TCATTACTAC 540
 15 TGGAAATTTG CATCGATACT TTTAATCTCA ATTCTGATTG AGAGTATAGC TTTAATTATT 600
 GTATTTCTTT ACTTTTTTAA TTTGAACATC ATCGATATTA TCTTTTTAAT TCTACTTAAT 660
 ATTTTAATGT GTTTAAGGAT TTATCTGAAT GTATTTTTTA GGATGACTTT AAAATATAAT 720
 CAGATTTTGT ATATTGCTCT TATTCAATTT TTAGGTTTGC TGATAGGACT ATTTCTATAT 780
 NAT 783

20 (2) INFORMATION FOR SEQ ID NO: 835:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 794 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 835:

30 TTACCTAATT TTTCAAtCaT AAGATTCCCC CTATTGTTTA AACaTAAAAA TATTATaCGA 60
 TTAAGCACGA GmCACTTCAA TATATTTTTT AGAATATTCT TATAAATATT AATATGATCA 120
 35 TTTCACTATT TAAACACGTT TAATATTAAA ATAAGTTATT CATATAAACT GGTGCTGTTT 180
 GATCCAATTG CAGACTTACG AGTCATTGAA AAGTCTCACA AACTATTGmA AGTAAALATC 240
 TTAAAATAGA AAGTGAATGG TTAATTTAAG TATATTTmAA AAATATTAAC CTTTTTAAGC 300
 40 ACTGCTATTT AGGATATACT AAATAATAAC TAAGTTTAGA AAAATAGGAG GAACATCGTT 360
 TATGTTAAAC AAGGTTTGGT TCCGAAC TGG CATCGCTCTG ATTATGCTGT TCATTCTCAT 420
 CAAACTATTT ATGGAAGTGC ACGAAGTATT TACTCCAATA GCTACAATCA TCGGTTCTGT 480
 45 ATTTCTTCCA TTTTAAATTA GTGGTTTTTT ATTTTATATC TGTCTACCTT TTCAAAACTT 540
 ACTTGAAAAA GTCGGTTTTT CAAGATGGGC TAGTATAACA ACAATCATGT TAGCTTTGTT 600
 TGCTATTATC GGCTTAATTG TTGCGTTTGT TGCACCAATC ATCATTTCaA ATATCAACAA 660
 50 TTTAATTAGT CAGACACCTG gCCCTACAAA AAGGAAGCAG AGCAAATTAT TAAATTCGCA 720
 CTGGCTCAAA TGGnTAAATT ACCTGGAGGA TGTAACCAnG TGGATTACCA ATATGGTAAA 780

55

(2) INFORMATION FOR SEQ ID NO: 836:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 836:

TGTTTATACT GTGCCTGAAT TAGAAGAGGT TTAAACACCT ATGAGACAAG ATGGAACCTCG 60
 TGATATTTAT GTTAATTTAG AAAATGTGAG TtATATGGAT TCGACAGGTT TAGGTTTATT 120
 CGTAGGTACA TTAAAAGCAT TAAACCAAAA TGATAAAGAA CTATACATTT TAGGTGTGTC 180
 AGATCGTATC GGTAGACTAT TTGAAATTAC TGGTCTTAAG GATTTAATGC ATGTTAATGA 240
 AGGAACGGAG GTCGAATAAC ATGCAATCTA AAGAAGATTT TATCGAAATG CGCGTGCCaG 300
 CATCGGCAGA GTATGTAAGT TtAATTCsTT TAACACTTTC tGGCGTTTTT TCGAGACTGG 360
 TGCCACATAT GATGATATTG AAGATGCCAA GATTGCAGnT 400

(2) INFORMATION FOR SEQ ID NO: 837:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 433 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 837:

ATTGATCCAT TTTGCGGTTT GGGTACAATA GCTATAGAAG CTTGTTTAAT TGCTCAAAAT 60
 ATCGCACCTG GTTTTAATCG CGAGTTCGTA TCAGAGCAAT GGAACATCAT GCCAGCAAAT 120
 ATTTATGATG ATTACCGTGA TGAAGCGGAT AAGATGGCTG ATTATGATAA AGAAATCGAA 180
 GTATATGCTT CTGATATCGA TCCAGAAATG GTAGAGATTG CTAAGCGTAA CGCTGAAGAA 240
 GTTGGGTTGT CTGATATTAT TAAATTTAGT GTAAAAGATG TCAATACATT AACAATTGAT 300
 ACAGAAGAAC CGGTGGCGTT AATTGGAAAT CCTCCATATG GTGAACGTAT TGGTgaTCGT 360
 GAAGAAgtG AAGAAaTGTA CCGTTATATT GGTAAACTAA TGAAACAACA TCCATTTTTTA 420
 TCTACATACA TTT 433

(2) INFORMATION FOR SEQ ID NO: 838:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 926 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 838:

	ATTGGGCCAA CTAATCCTGT ATATCCAAAT CCAGCAGACA ATGGTGTTC TTTAACCTGA	60
10	AGAACGTAAG CAATGATTCC AGTAATTATT CCATTTATAG TCAATGGTAT CGAAATAATT	120
	AAATTTTTC AGTAACTGG GATCATCATT TTTGCAGCTC CTATGAGTAA AACTGCATTT	180
	ACACCAATAG AATTGACACG CAATGAGCCA AATAAGAAAG TAACACAGGC AGCCACTATA	240
15	CCTAGGTTTG CTGCTCCACT TCCTAGTCCG TTTAACTAA TCGCAGTTGC AATCGCTACT	300
	AACGATATTG GTGTTACCAT TAATAATGAA AATGCCACAC TAATAAGTAT AGACATTAAC	360
	AACGGATTTA AGTCTGTAAA AGAATGAATT ACATTTCCAA TTGCTTGAGT AATTTTTCGA	420
20	ATGTAAGGTA ATGTGATTAG ACCGATACCC CCACTAACGA TAGGTACTAA AACTGGTAAT	480
	ATAATTAATT CAAAAGATCC AAGTTTGTTC TGTAATACCA TATATATAAG ACATGCAATA	540
25	ATAACAACCA AACTCGTATT TATAATGtCa CCTATACCTT TTAACATAAA ACTATTATTG	600
	CTATATACAA CAGCACCTGA ACCAATCATA GCTGATGTAC CTACTATAGC AGCACCTGCA	660
	CCATTAAATT TAAATTGATG AGCAGCTAAA ACCCCAATAA TAAATGCCAT AAATGATTGA	720
30	ATTAGTATCA CTAAGTATA CGTTAATTCT AAAATTTTCAT TACCACTTTT AAATATTTTt	780
	AATACTTCAC CTAATAAAGC ATTCGGAACA AGTGCAATAA CAACACCAGC ACCAATAGAA	840
	TTTAAAATCT TACTGAAAA CTGTTTATTA TCAGCATTAT TTGCGTTACT CATAAACGAC	900
35	CTCCAATTTG AACATTAnAC TCATCA	926

(2) INFORMATION FOR SEQ ID NO: 839:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 400 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 839:

	CTTCGAAACG TACTTGGTCA TTCCCTTTAC CAGTACAACC ATGTGCAATA CCTACTGAAT	60
50	TTGTTTTCTC AGCAATCTCT ACTAATTTTT TAGCGATTAA TGGTCTTGAT AAAGCTGAAA	120
	CTAATGGATA TGCATTTTCA TACATTAAAT TTCCTTTGAT TGCATAACTT ACATACTCAT	180
55	CACTAAATTC TTTTGTGCA TCAATAATAT GACATTCAAC TGCTCCATA TCTAAAGCTT	240

CGTATCctTTT GTCGATAAGC CATTGAACGG CCACACTTGT atCTAGtCct CckGAATATG 360
 CTAACAAT TTTCTCTTTC ATAAAATTCA CCTCATTGTA 400

(2) INFORMATION FOR SEQ ID NO: 840:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1043 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 840:

GGATCGCGGT GTTGTGCTTG ATATACTTTG TAGCGATAAC GTTTACCTAC ACAATCATAA 60
 CGACAATGAA AATCGTCATC GACTGTAACT ACATTGTTGA CATAAATATC ATCAGGTAAC 120
 GTTCGGTTCA TTGCATATTG CCATTGTGAC ATAGGTATAT TCAGCTCTGT GTCGAAATGA 180
 AAGTATTGCT GTATCGCATG TACACCTCTA TCAGTCCTAC TTGAAGGATG GATTCTCACA 240
 TGTCTTTTAT GCATGCGTTG TAATAGCTTT TCAAATTGTT GCTGTACCGT TCGACCATT 300
 TGTGAATTT GAAAACCTAG AAAATTATTT CCTTGATACG CAATTCTAC TAATATACGC 360
 ATGAATTTAC ACTCCTGAAT ATTTAATAC GAATAAAATA ATTGCAATTG GGATAATCAT 420
 GGTTAAAGAT ATCGTATCTC TCAATTGCCA TTTAAGCTGT CTGTAGCTCG TTCTCTTAAC 480
 ATTGGCATCA TAACCCCTAA CTTCCATTGC GACCGCTAAT TCTTCGGCGC GTTGGAAGC 540
 TGAGATGAAT AGTGGCACTA GTAATGGAAT AAATGATTTA ATACGTGTTG CAATGTTCCC 600
 TGAACCTTATT TCAGAACCAC GCGACTTTTG CGCCAAAATG ATTTTATCTA ACTCATCCAT 660
 TAACGTCGGG ATGAATCGTA ACGCAATGGA CATTATCATA CTTAATTGAT GAACTGGTAA 720
 TTTAAACATC tTTAGTGGTG CAAGTAATCT TTCAAACGCA TCTGTTAAAT CAATTGGACT 780
 TGTAGATaGT GTCaTAATTG TTGCAATCAT TAaATCCCA ATTAAACGCA GTGrTATATA 840
 TAGCCCTTCT AAAATACCAT TAGTTTCAAT CGTGATGCCA TGCCATTCAA CTAATACATA 900
 TCCACCTTTA GTTAAAAATA TATGCATCAT TAATGTGAAG ATTAAAAAGA AAAATATTGG 960
 TGTTAAACCT TTGATTAGGA ACCATAATTG AATTTTGTCT AATCTCATAA TGAAnAnGAT 1020
 AAGTGCAAAC ACCCAAAGAT ATG 1043

(2) INFORMATION FOR SEQ ID NO: 841:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 343 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 841:

5 TCGAAATCAA ATTATAATAG ACAATTTTAG GAGGTGGACT TTCGATGACC AAAGGAATCT 60
 TAGGAAGAAA AATTGGGATG ACACAAGTAT TCGGAGAAAA CGGTGAATTA ATCCCTGTAA 120
 CAGTAGTAGA AGCTAAAGAA AATGTTGTAT TACAAAAGAA AACTGTAGAA GTTGATGGAT 180
 10 ACAACGCAAT CCAAGTTGGA TTTGAAGACA AAAAAGCATA CAAAAAGAT GCAAAATCTA 240
 ATAAATATGC TAATAAACCA GCTGAAGGTC ACGCTAAAAA AGCTGACGCh GCACTGAGAG 300
 nTCCCCTCAT AATTTCCCCA AnCGTAACCA TGTGTGAATA AAT 343

(2) INFORMATION FOR SEQ ID NO: 842:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 430 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 842:

ATGAGTATTT TAATGATGCG CTACGAGCGT ATGGTCTTAC TGTGAnAACCA GGTGAATTTG 60
 GAACACACAT GAATGTTAGC ATAAATAATG wtGGwCagTC TCTGGTAGTT TAGAATTAGA 120
 30 AGAATTAAAG CAATTTAGAC AATGGGGTTC TAAAACACCA GGTCATCCTG AATACAGACA 180
 TACAGATGGT GTAGAAGTTA CTACCGGACC ACTTGGACAA GGTTTTGCTA TGTCmGTAGG 240
 ATTMGCTTTA GCAGrAGATC ACCTAGCAGG GAAATTTAAT AAAGAAGGAT ATAATGTTGT 300
 35 AGATCATTAC ACATATGTAT TAGCTTCTGA CGGTGATTTA ATGGAAGGTA TATCGCATGA 360
 AGCAGCTTCA TTTGCTGGAC ATAATAAATT AAGTAAATTA GTTGTTTTTAT ACGATTCAAA 420
 40 TGGATATTTT 430

(2) INFORMATION FOR SEQ ID NO: 843:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 329 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 843:

GGACTGCCCCG ATGATTTGAC AAATGAATTG CTGATTTGAT TTATATATTA TCTGnAATTA 60

55

TCATCTATCC AACAAATATGT CTTATCAATG GTATAGTCTT TGCACACCAA TGGAGGdAAA 180
TAAATCTCAA CCTTACTATA TTAATATATA ATCAAATCTT AGATTAACTA GTGTAATGAT 240
5 ACAGATGGAT AATTGAGTAC AAATTTAAAA CCCTGAGATT TTCGCTTTAA TTTGAAAACC 300
TCAGGGnTTA TTTGATTTTT ATATAATGA 329

(2) INFORMATION FOR SEQ ID NO: 844:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 844:

20 AGATGAAGAA AGTGTTCAT CGATTATCTT AACAGTTGTT GCTTTATATA CTACanTTCT 60
TCCTAGATCT TGTACCGTTG AAAAGAGACC TGGTTGTAAA ATCTTAATTG ACATTTTCAA 120
TCACCACCCA GTCATCAACA TTAAAGTTGC CATCTGATAT ATCTCTTTCG ATTTGTATAA 180
25 ATTTCTGTTC ATCTATTGCA TAAAATTGTA TCCATTCTCC TGCTTCGTAC ATTGACATTG 240
GTTACAGCTC GCTGCTAAAT ACTTTTAAACG GTGTGCGTCC AATAATTTGC CATCCGCCAG 300
GGAAATCTGA TGGGATATAG TCCTGTTTGA TTATTGCGAA TACCTACAGA ACCTGGCATG 360
30 AATTTTTTAA CCTTGGGcng ATTACGTCTA nGnGTATGT 400

(2) INFORMATION FOR SEQ ID NO: 845:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 535 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 845:

45 CATAAGTCCA GTAACCTTCA ACTGTGTCCT CATCTTTAGT TCCAAATATA TCAACGTATT 60
TATTTCTTAA CTGATTAATG TTTCCCAAC TCTCGGATCC AAACACTTGA ATATGACTAT 120
ACCAAACCCA CGTTTGCAAT GTTGCATGAA TGCTACCATT GGGCTTTTGC CATAACCATT 180
50 TTCCAGATAA TGAAAAATGC GGCTGAGTGT AATATTTGAT TAGTTCATTG ATATTAGTCT 240
CGTTTTCACT GATATTATAG GCTTTTGCTT CAGATGAAAA ACTGATTGGT GTTTTAGGAA 300
GTTGTGTTGA TGTAGTTTCT AAAAGTAACA ATGCTGTTGA TAAACTAAT TTATTCATGA 360

ACCGATTCCG AATTAAATAA AAGCTAAAAC TATGTTAAAT AAACCTAAAC AGTTAGTAGT 480
 GTTATTTAAG CAAAACCTTAT CATTTTTAAG TTGGACAGAA CAGATCAATA AGAGC 535

5 (2) INFORMATION FOR SEQ ID NO: 846:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 509 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 846:

TCAATCGATC AAATGGTGAT TGACCTTTTA AATTTGTATG ATCAAAATAA TGATTGTTGG 60
 CCATATCTTT ACTATGCTTA CGTGCAGTTT CAGAATTCTG TTTAGAATAC TTCAATGTAG 120
 20 ATAATTGATG TTGTTTTCTT TCAGCATTAA CTAAATCAAA ATTTTGTAGT TCAAAACTAT 180
 CTGCAAGCGA TTTCGATGGT GCTCCATATT GTTCTTTTAA TCTATTTTCC ATAGCATCAC 240
 TTACTTGTA AAGAGCTGTT ACATTATTAC GTCGATGCTT ATCATAAAAT ACCGTCGTAT 300
 25 AAATGTGATT TTTATGGAAA ACATCATATT CTTTATTATT TTGTTGTA CGCACTCTAC 360
 CTTTAACAAT CTCTGTTTCT GGTTCGCCA ATCTTTGCCT TACAACCGAT TTAGGTGTAT 420
 TGTATTTAAT TTTTGATTTT GAAGTGATTA TATTTTGATT TGTATATAAC GCATTAACCT 480
 30 TATCTTTCAT GTAACCTATC ATTATAAAA 509

(2) INFORMATION FOR SEQ ID NO: 847:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 847:

CGTATTGATC TATAAATAGT GTTTAGATGC TATAGTCGGA TGCTTAAGTA ATTTAAAGAA 60
 45 AGTATCTTTA ACATCGATGT GTGTATAATC ATTTTTAGAA GTATTATAAT CTTTTCTTTC 120
 TCCCTTCTAA AATATATACA GGTGCTTCAT CAGCTAGTGG TTCAACTGGA ATGTCAGCAT 180
 AAACCTCGCA TCATATGTTA AACAAAAACG ATTGTATCTG TACTTCACCT ATAACAGCAC 240
 50 TATCCAATTC GTGCTTATCA AATAAACTAA GATTTTGTG AGTACCTTTT CACAACTAGT 300
 ACATACGTCT TGAGTTCTGA AGCATCATTT CATAAGGAGA ATACTGGCCA CGTGTGGACT 360

55

(2) INFORMATION FOR SEQ ID NO: 848:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 848:

TACCACTTGA ATACATTGCA TATAAAATAC GACGTTGTAC TGGTTTTTAAA CCATCACGAA 60
 CATCTGGCAA TGCACGCTCT TGAATAATAT ATTTACTATA TCTTCCAAAG CGATCACCTA 120
 AAACATCTTC AAGTGATAAA TCTTGAATTA TTTCACAC TAGATTTTCCT CCTCATCAA 180
 TTGATCATTT TCAAGCACTT GTACTTCAGA ATTATCTAAA ATACTTTGGT CCTCTTGCAT 240
 ACCAAACTCA ACATGCTTTT CAATCCATTC ACGTCTAGGT TGTACTTTGT GCACCCATTG 300
 AATGTTGTTA CACGTTTAGA TGGAAACGCAC TTGCATCTTG CAACTTGTGA CACGnATTGA 360
 AAGTTnCGTG TTTGCGGGGG TtnCAGCGTC GTTTGCCCAT 400

(2) INFORMATION FOR SEQ ID NO: 849:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 849:

TTTATATGAT TTATTGAGG AGTTAAGAGA TTTATTTAAA GAAGAAGATT TAGAACCATG 60
 GACATCATGC GAATTGATT TTACAAGAGA AGGTGAATTA AAAGTTTCAT TTGATTATAT 120
 TGATTGGATA AATTCAGAAT TTGGTCAAAT AGGTCGACAA AATTACTATA AGTATAGAAA 180
 AATTGGAATT TTACCAGAAA CGGAATATGA AATTAATAAA GTTAAAGAAA TCGAGCAATA 240
 TATTAAAGAG CTAGAAGAAT AACTATCTT AATGTAAGAC TAAACAATAA AGCTTTGTTT 300
 AGTCTTTTTTA GCGTTTAAGT AAAAAGCaAT AGATACCGTA AAGTTGATGC TCATCAAATA 360
 ATAATATAaA GATAATTTTA GGTTTTTTAAA CTTTTAATCG 400

(2) INFORMATION FOR SEQ ID NO: 850:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 923 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 850:

5 CAGGCCATAC TTACAATTTT GGCCATCGTC CGGATGTATT nGGAATCGTT GGATAATAAG 60
 CATCAAGGTC GTGACTATTT CGTAAAATTC CAACTGTCCC AGAATTTACA TCATTATGTC 120
 CAATTACTGG GCAACCaGAT TTTGCCAACT ATCTATATtT CaTATATTCC aAATGTTAAA 180
 10 ATGGTTGAAT CAAAATCTTT GAAATTATAC TTATTTAGTT TCAGAAATCa CGGTGATTTT 240
 CACGAAGATT GTATGAATAT TATTATGAAT GATTTGATAG AGCTTATGGA CCCACATTAT 300
 ATTGAAGTCT GGGGCAAGTT CACACACGTG GTGGAATTTT TATTGATCCT TATACAACT 360
 15 ATGGACGTCC AAATTCTAAA TATGAAAAAA TGGCTGAGCA TCGTTTGATG AATCATGATT 420
 TATATCCCGA AAAAATAGAT AATCGTTAAA TGTATCATTT AATAAACACA CCAATAAGTT 480
 20 GATTTTCCTA ACTTATTGGT GTGTTTTTCA TTTAGCATAC ATAATAGGTT ACATTAAAT 540
 AACATTTTAT ACCAAAGTAC ACCAAAAGAA TATTAGTACA CGAATTAAAC AACATTTTTA 600
 TAGAAACCTA TTGCACTTTA ACGTCAATAA GTATATTTTT ATATTATCTC TAATTAATTG 660
 25 TGCGCGCTTA ATAACAGAAT ATTCTCAATA TTTTATTTT TTTGTGATTT GTTGAATAT 720
 TTAGTTGATA AGGCACAATC AAATTTACTT AACTATTGT ATTAGGGGAA GAAAGGATGG 780
 GATGTATACA TGACACAACA AAACCTCCAT GGAAATCAAA TTCAAGACAT ACCTCAAACA 840
 30 GGATTTTTTCG GGcATCCTCG AGGACTAGGC GTACTCTTCT TTGtAGAGTT CTGGGrAAGG 900
 KTTAGTTATT ATGGGCATGC GTG 923

(2) INFORMATION FOR SEQ ID NO: 851:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1004 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 851:

45 TTAnTTGCAT CTATTTCAST TGGTGTAAG AAAGCGGAAT TTGATTTTAT TGAAAAGTTA 60
 GCTCAAGAAA AATTAATCCC CgAATATATT ACAATAGATA TTGCGCATGG TcACTcAGAT 120
 TCaGTGATAA ACATGATTAA ACATATAAAA ACCCATATAC CTGATAGTTT TGTTATTGCT 180
 50 GGTAATGTTG GTACGCCAGA AGGTGTTAGA GAATTAGAAA ATGCTGGTGC TGATGCTACC 240
 AAAGTCGGTA TAGGTCCTGG TAGAGTTTGT ATTACAAAGA TTAAACAGG TTTTGGTACT 300

GCCGATGGTG GTATAAGAAC GCATGGCGAC ATTGCTAAAT CAATTAGATT TGGTGCATCA 420
 ATGGTCATGA TTGGTTCATT WTTTGGCGCA CACGAAGAAT CACCTGGTGA AACTGTAGAA 480
 5 CTTGATGGTA AACAGTATAA AGAATATTTT GGTAGTGCAT CTGAATTTCA AAAAGGCGAA 540
 CATAAAATG TAGAAGGTAA AAAAATGTTT GTAGAACATA AGGGTTCATT AATGGATACC 600
 TTAAGAGAAA TGCAACAAGA TTTACAAAGC TCAATTTTCAT ATGCCGGTGG AAAAGACTTG 660
 10 AAATCATTAC GTACTGTAGA TTATGTTATT GTTAGAACT CTATTTTCAA CGGTGATAGA 720
 GATTAATATT TATAGTAGGT GAKGTAAATT AAAAAATTCA TAGTAACTGT TGTTCGTTTT 780
 TTATCAATTA TTATCATTGC GCCANTAAAC GAATTTAAAC CATTCAATTC TTTACAAAAT 840
 15 GAAGTAAGAC AATATATTGA CATTACATC AATAAGAAA CAATTTCTGC GGAAAATAAA 900
 TTGGATACAC CGAAGAAACA ACAATTTGCC TTTAATAATA TACAAATGAA CATGTCGAAA 960
 20 TCAGATGTTG AGAAAACATT AAATAAACCA AAAAGAGTGA CATn 1004

(2) INFORMATION FOR SEQ ID NO: 852:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 852:

GTTCGCTCCA ATGTACGCAG TTGTATATGC ATAGACACCA AAATCGAACC ATTCATTGC 60
 ATTACCGATA CCGGTTGCAA CAACGGTTTT TTTAGCTTTC TTTGCATCCA CCATGTTAAT 120
 35 ATTCTCTTTA TTAAATCCA TGAATTCATA CACTCCCTCT TGTATGTTTC ATAATTATAC 180
 GGAGTTAATT AGTTTTGTCA AATTTATAAG AATTTAACT TTTATATAAA TACCAAAGTA 240
 40 TGTGTTGTTCA AAAATTTTCG TTTATCGnAT TAAATTCATG GTTTTATACC TCATACATTG 300
 AATTAGACATA ATAATCTAAA CACGACGTTA CTTTGGATAT ACTnAATGGT nnCACACTCA 360
 GTAGGAAGCA CGCATTGCT CGATAATGTA ACCATACTCA 400

(2) INFORMATION FOR SEQ ID NO: 853:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 361 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

GGTGAAATTG TAACTATATA ATCGATATTA CCTTCAGTGT TCAAACTGG AAAATAATAA 60
 TTACCATCAC TCTTTTTGTT GAATTTATAA ATTTTAAAAG GTTCACCAAG TGTATACGAG 120
 5 GCTTTCTCTT TATTATAGAT TTTATCAAGT GAAGTAACAT AAGAAAGATA GTCTTTTTCG 180
 GCTAAATTAC GTACGTCAGT AGGAACACTC TTGTCCTCGA CATTAACTTG CACATGCTTC 240
 TTATCTTTTG CTTTGATATT TGAATTGCTC TCAGCGTTTG CAATAGGGGT AACACTTAAA 300
 10 GAAaATATTA AACTTAAAGC TATTAATTTT GGAAAGTTTC TTTTCATATA AAAACTCCTT 360
 T 361

15 (2) INFORMATION FOR SEQ ID NO: 854:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 605 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 854:

25 ACnTTTACT AGAnAAAGnT ATTAAAGTAT ATCTGCTTTA CACCATTAAA AAGCGGCAAA 60
 TGCTATAGAT AATCTAAAGC ATCTGCCGTA TGAGnATTTA TTTATTTTTG ATTGTCATAA 120
 TCGTGTGGTT TTAAATTAAT TGTTTCTAGC TTTACAAATT TTGTTTTGTG AATGATTTTA 180
 30 TGAATAAAGT AAATCAACGC TAGAATGATT AAAGGTAAAA AGTTTTTAAA AGCATTTAAC 240
 CATTGATCTT TTAAATATA TTCAACTGAG CCACCAAATA GCAAGAATAA TAGTGTAGTG 300
 ATGACAATGA TTGGTCTTAA TGGATAAAAA GGTGCTTTAT ATGGTAGGAC CTTATTAGGG 360
 35 TCTTGACCTT GTTTTTTAAT AGCTTGTCGC AATCGTATTT GTGACCAAAT GCTTGATCCC 420
 CAAACAACTA TAATCATTGA ACCAATAATT TCAAGTAAAT TAAAAACGGC ATTTGAATTA 480
 40 AAGTTTGCAT AAATAATAAC AATAACAACG ACTGCATAAG TAGTTAATAA TGCTCTTAAA 540
 GGTAACCTAG TTGTCTTGTT TAATTTACTT AAAAATTGGG GTGCTTTTTT GTCTGAACCT 600
 AAGGA 605

45 (2) INFORMATION FOR SEQ ID NO: 855:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 515 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

TTGAACAAGC ATTTTAAAT AGTCAACAAC ATCAmGCAAT TAGTGAAGAG GCACAACTTT 60
 TAAATAATCC AGATGAATTA ATGGCATTIT TAAAGAATAA GCGTAAAAAT ATTTTAGAAA 120
 5 ATGCATACAT TATGCCGCAT AATATGAGAG AAATGTTACG AAGTTATTTG GAAAGTATGT 180
 CTCAAGACTT TAATGTTGGC GGATTTTTTA ATAAAAAGAA GAAAAAACTA CAAATTCAAC 240
 AACAGCGATT ATTAACAGCG ACAGATGCGT TACAAGAACA TGTTAATCAA CAAATTCGTC 300
 10 AACCAATGCG AGAAGATATG TCATTTGTTA CGCGTTTTAT CAATAAAAAA GmAGCTTCAG 360
 aTAmAGTATT AAATCAGCAT TATGACGTTA AGCCAGAAAT GATTGAAGGT TTATATCAAC 420
 CACAAACATC AATCAGCAAT ACTTATGTAC TTACATTTTC AGACGAAGTG GTTAAAGCCA 480
 15 TTAAGAAATA TGTGAACAA CAATCAACAC CAATT 515

(2) INFORMATION FOR SEQ ID NO: 856:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1607 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 856:

30 GTCATTTTTA AAATTCGCA TTCGCGTTTA ATTTTTTCTC TTTTTCTTT TTCTTCTAGT 60
 GACATACTTT CTITAGGTGT TTCAACCAAT TCAGATGTAT CTACATCATC AATTTTAGTG 120
 ATTTTGTCTA CATCTTCTT TAAATCTTCT GGGACGTTCT CGAAACGCTT ATATTGCTCT 180
 35 TTAGAGATAC TAGCAGCTAT TTCATTAGCT CCTAAATTT CATCTATCAA GCCGAAAGAC 240
 AAGGCTTCTT CTGCAGTAAG CCAAGTTTCT GCATCTAACA TCTGTTTTAA GTGTTCTTGA 300
 TCTAAATCTT TTGCTTTATC TAAATAAGCT GAATTACTAA CAGCATCTGT TTTTCAAGT 360
 40 AAATCCGCTG TCTTCTTAA TTCTTCTGCA TTACCTACAG TCATAACCCA TGAATTATGA 420
 ATCATTAATA AACTATTTTT GTGCATAAAA ATAGTGTCAC CACTCATAGC GATAACACTA 480
 GCAATTGATG CCGCTAAGGC ATCGACATAG ATATTAATTT TTGCAGGATG CATTTTTAGC 540
 45 ATATTGTATA TTGCATGTCC TTCAAATACA CTGCCTCCAG ATGAATTTAT ATGAACATCT 600
 ATTTCACTGA TGTCTCCTAG TTCATCTAGT TTATTTTTGA AATCTGTAGC AGTTACATCA 660
 CTTTCAAACC ATTTATCACT TACAATATCA CCATAAATGa ATATTTCAAC TTTACTTTTT 720
 50 GATTTTCTTT TCAATTTGAA ATACTTAGCT TTCATTGACA TTTTATCAC CACCTTTCaA 780
 AGATTTTCTT AATTCAAGTG GCGTGTCAAT TGGGTATAAA TCACCGCTTA ATTAGCGGCT 840

55

AACCACTACG AACTGCTTTA AAGTACACTT CTGCTTGTGT TGCACTATCA GCCCTTAAAT 960
 AAGATTTAAC GTTAAATTTA AAATACCTAT TTTTCTCTCT GTCTGTTTAA GTAAGTAGTT 1020
 5 TCCGATTAAn TTCTTCTTCA TACTGTTTGA CGATTGGCAA TAAGGTATGC TGCAAGTAAA 1080
 ATCTGTTTAA CTCTTCATTT TTCGCGAAAT TTGTATTGA TCTTGCAATT AAGAATATTG 1140
 AGGGCAATTG AAAAACGTTA GCTACTCTTT CTCTTGTTAA ATTCTCGCTT GCCACTATAT 1200
 10 CTTCAGAGAC ATATTTTTTA GGTAAAGGTT CGATTTCAAC ACCAGGCTCT TGGAATAATA 1260
 TTCCACCGTT TTCTTCATAG TACTGTTTGA AATCTTCTAA CACTTGTTGC CTTTTTCTG 1320
 TACTTACATT GGAACCATAT TTAAGCATAA AAGAATCTGG TTTTGCATT TCTGTAAGAT 1380
 15 TAAAGTTTCT TACTGCATTG TCAAAATCAG TTGTATTCTT CAACACATCA ATCGGACTAA 1440
 TACCTTGAAC CATATTAGAT GCCACGATGT GTTTAAATG CAACATGTCC aTATTATGAA 1500
 20 CAATCAATTT ATTTCCAGTT GCAGCATGAA TGGAATAATA AAGTTCACcG TGATTGGTTT 1560
 TCAATTAACA TTTCAACAAC ATCTGGATTT AATAAGAAAA GCTTTGA 1607

(2) INFORMATION FOR SEQ ID NO: 857:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 346 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 857:

35 ATGCGCATT AACC GCATT AGAGATGCTG AAGATCAGTA TAAAGCTTTG TTAGAAATTA 60
 CAACATTACC AGAAGGTAGA ATTTATGTTG CTCGCCAAGA TCAACTCATT GTGGGTTATG 120
 TCACTTTCCA CTATCTGAT GAAATTGAGC GCTGGTCTAC AGGTAAGCTT CCATATTTAA 180
 40 TCGAATTGGG GGCAATTGAA GTCAGCATCA ATTTTAGGCA ATTACAACTT GCAGAAAAGC 240
 TGATACAACT TAGCCTTCTA CACCAGAATT CGAGAATTAT ATCGGnATAA CTACAGGATA 300
 TTACCGGCAT TGGGGGTTTA AAAAATTCCC AGTnAGAGGT ATTnCC 346

(2) INFORMATION FOR SEQ ID NO: 858:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1176 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

ATTACTATAA TGAATCTTGC GAGATATTTA GAAGTAATGA TACAAnATAA TGGACTTTAG 60
 AATTTmAATT TACAATAAAT AGTTCTATGA TTAAATGTCA GTTTTATGAC ATTTATTTAT 120
 5 TGAAAAATACG AACGAATGaG CGATaTGATA ATATAGATAA GAATGATTTT AATTTAGGAG 180
 GCCTTTATGG TGCATAATAA AAACAATACA ATTTTAAAAA TGATCAAAGG TGAAGAAACA 240
 TCACATACAC CTGTTTGGTT TATGCGACAA GCTGGCCGTT CGCAACCAGA ATATCGAAAA 300
 10 TTGAAAGAAA AATATTCaCT ATTcGATATT ACACATCAGC CGGAGTTGTG CGCTTATGTA 360
 ACACATTTAC CAGTTGATAA TTATCATACA GATGCAGCAA TTTTATACAA AGATATTATG 420
 15 ACACCATTAA AGCCAATTGG TGTcGATGTA GAAATTAAAT CGGGTATTGG TCCAGTGATT 480
 CATAATCCAA TCAAAACAAT TCAAGATGTT GAGAACTTT CTCAAATAGA CCCCgaACGA 540
 GATGTACCAT ATGTATTAGA TACAATTAAA CTTTAAACAG AAGAAAAGTT AAATGTGCCG 600
 20 CTAATAGGAT TTAaCTGGGGC ACCATTTACA TTAGCGTCAT ATATGATTGA aGGCGGACCA 660
 TCGAAAAATT ACAATTTTAC AAAAGCGATG ATGTATAGAG ATGAAGCAAC ATGGTTTGCT 720
 TTAATGAATC ATTTAGTTGA TGTATCTGTT AAATATGTAA CAGCTCAAGT CGAAGCAGGT 780
 25 GCCGAATTGA TTCAAATTTT CGATTCATGG GTAGGTGCAT TAAATGTcGA GGATTACAGA 840
 CGTTACATTA AACCACATAT GATTcGATTA ATCAGTGAGG TTAAAGAAAA ACATGATGTG 900
 CCGGTAATTT TATTCGGTGT AGGTGCCAGT CATTTAATCA ATGAATGGAA TGATTTACCG 960
 30 ATTGATGTAT TAGGCCTAGA TTGAGAACG TCTATTAATC AGGCTCAACA ATTAGGCgTT 1020
 ACTAAAACAT TACAAGGGAA TTTAGATCCA TCAATTTTAT TAGCACCATG GAATGTcATT 1080
 35 GAAGAGAGAT TGAAACCAAT ATTAGATcAA GGTATGGAGA ACGGTAAACA CATCTTTAAT 1140
 TTAGGACACG GTGTATTCCC AGAAGTGCAC CAGAGA 1176

(2) INFORMATION FOR SEQ ID NO: 859:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 621 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 859:

ATTTGGTGCA TTGGTAnAAT TGACCTGAAA GnAAAAAGTG GTTTAGTTGC ACATTGAGTG 60
 nAAGTGCgCA GATGAATTAT GTTGAAAACG TGAGGAAGAG CACCTTTGCT GTTGGTGATG 120
 AAGTAGACGT AAAAGTATTA TCTATTGCTG ATGATGGAAA AATTAGTCTT TcAATTAAGA 180

AAAAAGCCGA AGATTTTGAA AAGAAATTAA GCAATTTCTT AAAAGATAGT GAAGATAAAT 300
 TAACTTCAAT CAAACGTCAA ACAGAATCTA GACGCGGTGG CAAAGGTTCA AGACGTTAAT 360
 5 TAAAATAAAT AAAGACTGTT TCGATAAGGA ATATATTTAG AATGATGCGT ATCGAATAAT 420
 CGATTGCAGC GTTAGACAAT CTAAGACTGT TTCTTAAATA AGGAGCAGTC TCTTTTATTT 480
 GTAATGATAT AACTAAGACT TATACCATTT TTGAAAATTG TAAAAGTGAG GTGATGTTAT 540
 10 GCAGTTAAAT AGTAATGGTT GGCATGTTGA TGACCATATT GTTGTCGCTG TTTCTACAGG 600
 TATTGATAGT ATGTGTTTnn T 621

15 (2) INFORMATION FOR SEQ ID NO: 860:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 544 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 860:

25 TATATTAGGT ACTTGTGTTG GTATTGTGCC AGCTTTGCTG TCTACTATAA TTTCTAAAAG 60
 GTTTGAACAT ATAAAAGGGA AAGTGCTAGG TGTATTTAAT TTTGTGAGAT ATATTGGAAT 120
 30 GACTGTCGGT GCATTATTAA TTGGTATCAT TTCTCAGCCG TTGGTAGCCT TTTACTTCAC 180
 AACTATAACT ATCATGTAA TAGTAATATT TCTTTATATA AAGATAGTTG ACTTTCAGCT 240
 AAAGTATGCC AAATAAACTT AAAAAAGCAG TGAATCATTT CAATAGATGG AATGATTCAC 300
 35 TGCTTTTTGT TTATAGAAAT TTAAAATTCA TGCCTTTATG CTAACAACAA TATTATTTTA 360
 CTCTCTTACG TTTGCCAATG TAACCGTATA GCAAAGTAAA TGACTGCGAT AATGATGACA 420
 ACATACATat ACGTGAATAG GAAGGAAGAC CAGTCATCAA CATAACAAAG CTATCGCTCA 480
 40 AAGTGCGnCT GGTAAATTAA ACCGAATTCC AnGATTGTAG GAACCCnATT AATGGAAATA 540
 ACGG 544

45 (2) INFORMATION FOR SEQ ID NO: 861:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 651 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 861:

55

AATTTAAAAT ACATTTATGC TTTACCCATT CGAAAATACC TTAATCATT TCATTTATAG 120
 CATTGTTTGA TTGAAGGATA AAAAGTTGTT GTTTACAATA AAAATAATGA GTATCTGAAA 180
 5 TGAGGGATT CAnTATGACAC ATGTGGAAGT AGTAGCGACT ATCGCGCCAC AATTATCTAT 240
 CGAAGAAACT TTAATTCAnA AnATTaATCA TCGTATTGAT GCAATAGACG TATTAGAATT 300
 ACGAATTGAT CAAATTGAAA ATGTCACAGT TGaTCAAGTG GCAGAAmTGA TTACAAAGCT 360
 10 GaAGGTTATG CAAGATTCAT TCAAATTATT AGTTACGTAT CGTACAAAGT TACAAGGTGG 420
 CTATGGGCAA TTTACAAATG ACTCGTATCT TAATTTAATA TCAGACTTAG CAAATATCAA 480
 15 TGGCATAGAT ATGATTGATA TAGAATGGCA AGCAGATATT GACATTGAAA AACATCAACG 540
 AATCATTACA CATTTGCAAC AGTATAATAA AGAGGTGGTT ATATCACATC ATAATtCGA 600
 AAGTACGCCT CCATTAGATG AATTGCaATT TATATTTTTn nAAATGCAAA A 651

20 (2) INFORMATION FOR SEQ ID NO: 862:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 401 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 862:

TGTGAGTGGT AATGACACAC AAGCAGATTT AATAGGTTAT AGTTTTAAGT TTGACGGTGC 60
 GATTTCaAGG CAAGAGGCTT CTAAAGATGT ACATGCAGTT ATATTATCGA ATAAACACT 120
 35 ATATTTATTA GATGTATTAC AAAAATTACC GATAGATAAA ATGAATTCAT TGAATATCCA 180
 TCAAGAAATT ATTGATGAAA TGTCAGATAT CATTTTAATG TTATATCGTG AATATGCAGG 240
 TATGTTTTTT AAAAGTCAGA AACTAATCAA CCAATTAAAA AGATTGGAmC AATAACaTAA 300
 40 AATAATAAAa GGTATTCAAG TAGCCACATA GATGTGTTTA TTTGAATACC TTTTkgAATA 360
 GaaAAGaGAT ACTGGCaATT TTAATAACCA GtATCTCTTT T 401

45 (2) INFORMATION FOR SEQ ID NO: 863:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 308 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 863:

ACAAAGTAAT CAATATGTAA TATAAAATAC ACTGGTACTC AATATATAAT GATGATAAAA 120
 TTAATTTTAA TTAGATAGAG TTGCTTTGTG TTTTAAACGC nGATGCTACT ACTTATCTTA 180
 5 ACAGTTGATT AAGTAAATCA TTTAACAGCG AGATTATnCA ACCAGGGGGA TGACTIONATG 240
 AATTTATTCA GACACCAAAA TTTAGTATCA GAAATTTAAT GTCGGTATTT TTTCAGCTTT 300
 AATGCCAC 308

(2) INFORMATION FOR SEQ ID NO: 864:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1827 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 864:

TTCTACATCT TAAACGATGA CCTAACTATG ACGTTAATCT GGAAAGATGG GGAGTTGCTA 60
 GTATGATGTT GAAATTTAAA GCTTGGGATA AAGATAAAAA AGTTATGAGT ATTATTGACG 120
 25 AAATCGATTT TAATAGTGGG TACATTTTGA TTTCAACAGG TTATAAAAGT TTCAATGAAG 180
 TAAAACTATT ACAATACACA GGATTTAAAG ATGTGCACGG TGTGGAGATT TATGAAGGGG 240
 ATATTGTTCA AGATTGTTAT TCGAGAGAAG TAAGTTTTAT CGAGTTTAAA GAAGGAGCCT 300
 30 TTTATATAAC TTTTAGCAAT GTAAGTGAAT TACTAAGTGA AAATGACGAT ATTATTGAAA 360
 TTGTTGGAAA TATTTTTGAA AATGAGATGC TATTGGAGGT TATGAGATGA CGTTCACCTT 420
 35 ATCAGATGAA CAATATAAAA ATCTTTGTAC TAACTCTAAC AAGTTATTAG ATAAACTTCA 480
 CAAAGCATTAA AAGATCGTG AAGAGTACAA GAAGCAACGA GATGAGCTTA TTGGGGATAT 540
 AGCGAAGTTA CGAGATTGTA ACAAAGAACT GGAGAAGAAA GCAAGCGCAT GGGATAGGTA 600
 40 TTGCAAGAGC GTTGAAAAAG ATTTAATAAAA CGAATTCGGT AACGATGATG AAAGAGTTAA 660
 ATTCGGAATG GAATTAAACA ATAAAATTTT TATGGAGGAT GACACAAATG AATAATCGCG 720
 AAAAAATCGA ACAGTCCGTT ATTAGTGCTA GTGCGTATAA CGGTAATGAC ACAGAGGGGT 780
 45 TGCTAAAAGA GATTGAGGAC GTGTATAAGA AAGCGCAAGC GTTTGATGAA ATACTTGAGG 840
 GAATGACAAA TGCTATTCAA CATTGAGTTA AAGAAGGTAT TGAACCTGAT GAAGCAGTAG 900
 GGATTATGGC AGGTCAAGTT GTCTATAAAT ATGAGGAGGA ATAGGAAAAT GACTAACACA 960
 50 TTACAAGTAA AACTATTATC aAAAAATGCT AGAATGCCCG AACGAAATCA TAAGACGGAT 1020
 GCAGGTTATG ACATATTCTC AGCTGAAACT GTCGTACTCG AACCACAAGA AAAAGCAGTG 1080

5 AGTGGTGTAA GTAGTAAAC GTATTTAGTG ATTGAAACAG GCAAGATAGA CGCGGGATAT 1200
 CATGGCAATT TAGGGATTAA TATCAAGAAT GATGAAGAAC GTGATGGAAT ACCCTTTTTA 1260
 TATGATGATA TAGACGCTGA ATTAGAAGAT GGATTAATAA GCATTTTAGA TATAAAAGGT 1320
 AACTATGTAC AAGATGGAAG AGGCATAAGA AGAGTTTACC AAATCAACAA AGGCGATAAA 1380
 10 CTAGCTCAAT TGGTTATCGT GCCTATATGG ACACCGGAAC TAAAGCAAGT GGAGGAATTC 1440
 GAAAGTGTTT CAGAACGTGG AGCAAAAGGC TTCGGAAGTA GCGGAGTGTA AAGACATCTT 1500
 AGATCGAGTT AAGGAGGTTT TGGGGAAGTG ACGCAATACT TAGTCACAAC ATTCAAAGAT 1560
 15 TCAACAGGAC GACCACATGA ACATATTACT GTGGCTAGAG ATAATCAGAC GTTTACAGTT 1620
 ATTGAGGCAG AGAGTAAAGA AGAAGCTGAG CGCAAATACG AGGCACAAGT TAAGATAAGG 1680
 AGAGATGGAG ATGCCAAAGA AAACGGTAAC GATTGATGTA GATGAAAAC TATTAGTAGT 1740
 20 AGCTAGTAAT GAAATATCAG AACTATTATA TGAATATGAC AGTGAGTtAA TGTCAGCTGG 1800
 ATGGAAGATG GCGATAATAG GGATATC 1827

(2) INFORMATION FOR SEQ ID NO: 865:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1379 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 865:

35 TCAATTGCAT CATCATATGA AATTCTaGGG AATGGTGTCTG CAACTTTTTTC AAGTTTTGAT 60
 GTATCACGCT CTAAAATTTT CAACTCTAGT TTACAATTTT CTAAAACCTGA TTTTACAACA 120
 TGTGTTACAT ATTGTTCTTG AATTTCTAAA CTTTCAGCAT GATTGTGAA AGCCATTTCT 180
 40 CCTTCAATCA TCCAGAACTC GATCAAGTGT CTACGTGTTT yTGATTTTTTC AGCTCTGAAA 240
 GTTGGACCAA ATGAAAATAC TTTTCCGTGT GCCATTGCTG CAGCTTCTAA GTATAACTGA 300
 CCACTTTGAG ATAAAAACGC ATCTTGATCA AAGTATTTAG TATGGAATAA TTCACTTGTA 360
 45 CCTTCTGGCG CACTTGCTGT CAAAATTGGT GGATCAACCT TTGTAAATCC ATCTTTGTTG 420
 AAAAATTCAT ACGTTGCACG AATAACTTCA TTTCTAATTT TCATTACAGC ATGTTGTTTT 480
 50 TTAGAACGTA ACCATAAATG ACGGTGATCC ATTAAGAATT CTGTACCATG ATTTTTAGGT 540
 GTAATCGGAT AGTCATGCGC TTCTGAAATA ACTTCAATTG ATTTCACTTG CATTTCTGAT 600
 CCTAAGTCAG AACGATTATC TTCTGTAATT GTGCCTGTAA CGTATAGAGA TGATTCTTGA 660

TGCATAAAGC CTGTTCCATC ACGTAATTGT AAAAAGGCGA TTTTACCACT TGAACGTTTA 780
 TTTGTTAACC AAGCACCAAT TGTAACGTCT TGGTTTAAAT GATCTTTCGC TTGTTTAAATC 840
 5 GTTGTTCCTA TAACCATTCT CCTATTATTT TTTTCGKTAT ACAATACTCA TTCATTTTAA 900
 CAAAATCCGC TTTCAAGTTC TAGAACTAGA CTAAAAGATA ACGtGTAAAT GGTAATGATT 960
 TACGCACAGA TTTGAACATT AATTTATATT AAGACAATAC AATCATCAGT TTCAGACAAA 1020
 10 ATATATAATA AAGCCTCAAT ATTATTAGTA TGAAGGTCGT TTTTGTnCT AACATTCAAG 1080
 TTTAATTATA AAACCTCAAT TTTGACGACC AAACAAAAT TCTTGTACA GAACATAGCA 1140
 15 CATGTCATGT CACAAGAATT TTACTTTGTA TTTCACTTTT TCTTTTTTTG AATTTGTCTT 1200
 AATAATTTTC CAACTGTTG AATGTCGCCT TTTTCTGAC GATAATTTTC AAGTGTGTGT 1260
 TCAAAAAAGT TTTTATAATT ACTGTTTATG AGTCGATCAT CAAATGAAAC TATTATGCCG 1320
 20 CGATCATTTT CATTCTAAT TAATCTTCCA AGTCCTnGT CTAnAACGTG TAACTGCAT 1379

(2) INFORMATION FOR SEQ ID NO: 866:

(i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 450 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 866:

TTACAACGAA TAGTTCACCG TTCGAGTCTA ATCTATGTAA TGCTTGCTCG AATATACGAT 60
 35 GCACGGTTTC TTTCCCTGCT CTTATTGGTG GATTGGTTAA AATAAAATCA AAACTTTGT 120
 CTTCCACAGC AGACAAAGCA TCACTTTCCT TTACGATCAC ATTATCAATA CCATTTAATT 180
 TTTTGTTCCT TTCAACTAAG GCTAGCGCTC TGTGATTAAC ATCTAGCATT GTAATTGAAT 240
 40 GATGTGGTGA TACTTTAGCA ATCATCAAAC CAATTGGTCC GTAACCACAA CCAACATCGG 300
 CAATTCGCTT ACTTGGACCA GGTGGATGCG CTTTTAAAAA AGTTTGAACA AGAACATCTG 360
 aACCATAATC TACTTTAtCt TTCGAAAACA CTCctTATCA GTTATTAAAT CAATTTTATG 420
 45 AnGGTTnTAA TTGGATATTG AGTACGTTGT 450

(2) INFORMATION FOR SEQ ID NO: 867:

(i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 867:

5 AATGTTAACA ATATAGCACC AATTAACCCC GACATGATAA TAACGTGTAA TGTTTTATTT 60
 CCTATTAATT GTCTCGCAAT ATGAGGTGCA ATTAATCCTA AAAAGCTAAT ACCACCGACA 120
 ACTGAAATTG CGGATCCTGC TAATATTACT GCTAAAATTA ACAATAGCAT TTTAATAGTT 180
 10 TTAACTTTTA AACCGAGTGC GGTGCAACA GCATCACCTA GATTCAATAC ATCTAATTGA 240
 TAACTCCATA AAATGATGAT GGATCGTTAT TAAACCAGG GAATATAGAT AATATCCGCA 300
 TATCAGGCCA TATAGCTACC TGTCACCAAC AGGTTTnTGC TCCAnGGTTC TnTAATAAGA 360
 15 CGCACACGCG ACAATGCGCT TGCTACCATT AGGCAGCTTG 400

(2) INFORMATION FOR SEQ ID NO: 868:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1132 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 868:

AATCCnTnAA GTGCCATAAC ATCATCTCCT AACATCTTTA TTATACATCA ACATTTTATA 60
 30 AAAATAACnT CTTATGATAA AAATGAAAAT ACTAATTTTA ATGAGAGCGT CTTAAGCCGC 120
 AATTGATAAA ACATATGCTA CAATATTTTT AAACACTATA CAGGAGGTAC TCTATGACAA 180
 AATATACATT TAAACCTAAA GATTTCAAAG CGTTCAACGT AGAAGGCTTA GACGCACGAA 240
 35 TGGAAGCTTT AAACGAATAC ATACGACCAC AACTCCGTGA ATTAGGAGAA TATTTTAGTG 300
 ACTTCTTTAC AAGTCAAAC TGTGAAACAT TTTATCCTCA CGTAGCAAAG CATGCTAGAA 360
 GAAGTGTGAA TCCTCCTAAA GATACATGGG TTGCTTTTGC AACAAACAAA AGAGGCTATA 420
 40 AAATGTTACC TCATTTCCAA ATTGGTATGT TTGAAGATCA ACTGTTTGTT ATGTTTGGAA 480
 TCATGCATGA AGCAAAGAT AAAGCAACaC GTGCAAAAGT TTTTGAAAGA AAATTTAAAG 540
 CTATTCAACA ATTACCTGAT GATTATCGTG TTTGCTTAGA TCATATGAAA CCTGATAAAC 600
 CATTTATTAA AGATTTAACG GATGATGATT TAATAGAAGC GATACAAAGA GCCATCAATG 660
 TGAAAAAAGG TGAATTCTTT ATAGCGCGTG CAATCACACC ACAAGATAAA AGATTAAAAA 720
 50 GTGACAAAGC ATTTATTGCA TTTTGAAG AAACCTTCGA TCAGTTCTTA CCATTTTATT 780
 CTGCATAAAT AACTTTGTTT AAATAATAGA GCACTTAATC ACATCCaTGA TTTCTGCTcC 840
 TTtTTTCTTA ATATTAAATC GAACGTcCAA CATAaTAATT CATACTTTTA AAAAAATTAA 900

GGTGAGTCAA AATGAATAAT AAACGACATT CAACAAATGA ACAATTAAGT TTAGACGAAA 1020
 TAAACAATAC AATTAAATTC GATCATCGCA GTTCAAATAA ACAGAAATTT TTATCATTTT 1080
 5 TTGGACCTGG GTTATTAGTC GCTGTTGGTT ACATGGATCC CGGAAACTGG AT 1132

(2) INFORMATION FOR SEQ ID NO: 869:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 418 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 869:

CTCAATGTAT TGA CTGTGAT TTGGAGTGTA TAATGCTACT GTTTTGTCTA TAGCCATCGC 60
 20 ATCAAATATT ATATTGAGT AATCAGTGAT AACGACATCA GACATTAGAA TTAAGTCTTG 120
 AGCACTAAGA TACTTTGGTG CAACAAGCGC CTCTTCTGGT AAGATCGCTT CATCAACACC 180
 CTGAACCACA ACATGATAGG CTTTGAATAA AGCATCTGAT AGTGGTAATT GTTGCGCACT 240
 25 CACTAATCCA ATAGGTGCGT ATAATAAAAC TGGTTTTTGA TCATTTATTT TAAACGATTT 300
 CTTATACTGT TGTGAGTAG TACTTTCGTT ACTCTTTTGT AGTAAATATT GGTGTTTAGG 360
 ATTACCATAA GCCAATACAT TGCTTGGTGA CTAGGAAAAG nCGTTTGATA AAAGGGCT 418

(2) INFORMATION FOR SEQ ID NO: 870:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 587 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 870:

GAGTnTGTTA CAGTAATTGA TTTTATTGGT AATTnTAAGA CAAATTATTT AATTCCGATT 60
 GCGCTTTCTG GGGATCAATC GCAAnATAAA GATAATTATA AGAAGTTTTT AACGAATAAC 120
 45 GATTGATTA ATGGAGTATC TACAATTAAT TTTGAAGAAG TTGCTAAAAA ACAGATTTAC 180
 AATTCATTGG ATGCAGTATC ATTAAATCAA AATAAATTAA TATTAAAAGC TTATGAAGAA 240
 GTTGAAAATA GATTGGGACA CATGCCGTTA CTAATGGATT TCATACAACA ACATTCTATA 300
 50 GATCCAAGCG TTATATTTTC TAAATTTAGT AATTATTACG AGTTCCTAGT GAGATATAAA 360
 AAAATAGATA CATTATTGAC GGAAAATGAA TCTAAAAATC TGGTTTTCTT TTCAAGGCAA 420

TTAACATATG ATGAATTAAA AAATAAAATG TTGAACGAAG TTAAGGATAT AACAGAAGAT 540
 GATATAGATA CTTCATTAAG AATTTTAGAT TTTTCATTTT ACAACGC 587

(2) INFORMATION FOR SEQ ID NO: 871:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 971 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 871:

CTACTAAACC AGGTAATCGG TAGAATCCAA GCAGGAATAA ATAAATTAAT GCAACACCAA 60
 TAAACGATGC AAACACAGTn TTATCTAATG CATCTTGACC AAATTGGGCA CCTACTGAGT 120
 TTGAATAAAT TTCTTTCAAG TCAACTGGTA AAGAACCTGC ATTTAACAAT KCGGCGATTT 180
 GTTTTGCTTT TTTAAGCCT TCTTGCTCTT TAAATCCACC CGAGATTTCT ACGCTATCAG 240
 AATTGATTGG TTGATCAACA CTTGCTGCAG AAATAAATTT AGGGTTTTTC TTTGTGCTT 300
 CTTTTTTATA GCTGTCACCT TTTTGAAT CTAACCAAAC AACCATGACA TTATCACGTT 360
 TCTTAGAGAT TTCTCCGTT ACTTTTTTAA ATTGTTTTT GTCTTTTACT TTAAGAGTAA 420
 CTGTAGGcTG GTTTGTTCCT tGtTTAAAT CTTGtTTGGC AGATCCCTGT TtAATATCAG 480
 AACCCTTAA TTTtACTTTA TCTTCTGCAT CGCGAATTGT TAAATTAGCT TGAGAAGATA 540
 AAATTTTACG TGCTTCATTC TGGTCTGTTA CACCAGCAAG TTGTAATCTA ATTCTATTAG 600
 GTTCTTCAAC TTGAATTTTA GGTTCGAAA CACCTAAAAC GTTAACACGA TTTTCTAATG 660
 TTTGCGCTGT TGATTGTAAG GCTTTTTTAT CTATTTTGTC GCCTTTATTT AAAGGATCGA 720
 CTTGATAAAG CACCTCAAAT CCACCTTGCA AATCAAGTCC TAAATTGmCA TTCTTTATAA 780
 CACTTTTATA AGTTGcAGCC ATTCCGGCAA ACAACAATAC GACTAAAAGC AAGAACGCAA 840
 TTATTCTACT ACTTTTCTTC ACATGAACAC CTCATTATTT ACGTATGTAT TTAGAATACT 900
 TGAATACTAT TTTATAACGC AAGTGAAATC TTTCTTACAA AATTTATTAG CCTTATACAT 960
 ATTAACATAC T 971

(2) INFORMATION FOR SEQ ID NO: 872:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 636 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 872:

5 CTGGTTAACA ACATCTGGTT CGGCATCTCC TTAGCTTCCA CTAAATCTTG GATGTCTTGG 60
 nnATCTTGCG CAAGTnGTGC TTGGCTTGT TCAATTtCTy CTTTAGTCAT CGCATTGTTA 120
 ATGCCGTTAT GACCTTGTTG AAGTATTTGA TTAATACGAT CTTTAAGTGC TTGTTTTTCC 180
 10 TTATCTGTTA GATTTGGATT TCGATCAATT TCGTCAATTA ATGCTTGaAC TTGcTTATCA 240
 ACGTCTTTAT TGGCATCAAT TTTTGCTTTT GGTATTTTCAT TGGCATGCAC TTGTTCAATC 300
 GCGTGGTTGC CTGCTGTTTG AACTTGAGAT ACAGCCTGAT TACTTGTTGC TTTATTAATG 360
 15 TTGTTGATGA TGCTGTTTGC CAATTCTTCT GCTTTATTTT TCGCAATAAG CTTGTCTTGA 420
 TCCGTCGCAT TTGAAGCTTC GATTTCTTTT AGCTTATTAG CTAAAGCTTG ATTAATAGAT 480
 TGAATTGCCT TGTCTTTAGC ATCTTGTAGT CGTTGATCAC CATTAGATT ATGGATTGCA 540
 20 TCATTGACTg CTTGGATTGC GCCATTGATA TCATTACAT tTGTGTaTCA CTATTTAGCA 600
 ATGTATTTGC TAGACGTTg CATCATCGAA GTTTGT 636

(2) INFORMATION FOR SEQ ID NO: 873:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 659 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 873:

35 CTCTGAATGC ATTCATTAAA CGCTGGGTCT CTTTATATGA AACGTATTTG TCATTTTTAG 60
 AACTCAATCC GTAAAAATTG TCAACTTTCT TTTTAATATT ATCGTAATCA ATGGTTACAT 120
 TACTTAAATC AATATCTAAA TCTATATTTT CTGCATCTTC TTTAAAGCCC GCTATACTGA 180
 40 AAAAGCCTTC AATCGGCTGA TCAATCATTT CAATATATTT TAAAGCTGTG ATTGAACCTA 240
 AACCATGTGT TACAAAATAT GTATCCTTTT TGCGTACATT AATTTGTTTC GTCATAGCTT 300
 CAATCCACTG ATCCACTGTC TTCGCTtCAG GGGATTCAAA ATTAAATAAT GTTACGTCAT 360
 45 ATCCTTCTAA AGTTAAGTTA TGCTCCAACC ACTGATACCA ATGATTTCTA CTATTTCCAT 420
 GCATAGAATG TACAATaTA CATCTGTCAT CTCATTCTCT CTTTTCAACT TACTACTTCT 480
 TTTCTATTTT TAAAAAATG ACTGATTACC TATAATTGTA AAATAAAAAC ACCTTAATTA 540
 50 GAAATGTTAT ATCGCAAAGT GACATTTCTA ATTAAAGTGT ATTGTCATCA TTTCAATATC 600
 ATTCAAAAAC AGCTAAACCT TTGTCTCTGC TTCAATTTCA CAAAAATAAT TCCCGCTGA 659

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 874:

AAAATAAAAT CCATTTTAGT ATTCGTAGAA AATGAGAGCC CCCTTTACCA CAATATGTAA 60
 ATGTATATTG TGAAAAGGGG GCTTTTATTT ATGACATAGT TGTTCAACTC ATACATTGAG 120
 TTGATAATGC AGTGCGTTAT GCAGCTATCC ACTTCATATG TATGGCGCAT ACTTGATATA 180
 TGAATATCAT CAAGATTTTA AAAGTATCGC TAATTTTAAA GTCTTAAAC GCAGATTAAT 240
 GGTGATGATG TGTTAATGCG TTTTGGTTTT TAGAGCCACG TAATTCCTCA GGACATATAG 300
 GAGGAACCAT TCCAC 315

(2) INFORMATION FOR SEQ ID NO: 875:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 482 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 875:

TACACGACCT TGTAATCAG CTGCTTTATA AGCTAGGGCG ATGTTATCAA AGTTGGATGT 60
 ACTTTGAGCT TGGCATTTCG TTAATACTTT TAACGAAGGA TTTAATATAT GTCTGATACG 120
 TATATATTGA TTATATTCAA TTCTTaATTt GGATAAGATT GCTGATAATT TGAAGCAATC 180
 GGAGCTAAAT GGATCGAAAA ATGAATAAAT TTCGATTTTA CTTACAGGTG ATAGaTTGAT 240
 ATCTTCACGA CTCTTATTTT CCATTaWTCG TAATTCTCCA GCCATGTTTA TTCACCTACA 300
 ATTAATTTTA GGAATTCACC ATATGaTTAG CAGTTAATCT TaAGCGCTCA AATAAATAAT 360
 CTCCAACACC TTGTGGAAAC caGcGCaTT AATTGCTGTC TGCATATTTT CTAGCCATGC 420
 ATCTCTTTCA AATTCAGTGA TTGTAAATC TATATGTCTT TTTCTTAGCA TAGGATGTCC 480
 GT 482

(2) INFORMATION FOR SEQ ID NO: 876:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 670 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 876:

5 ATGCAATTGA ATTCTTAAAA GCAGAAGGCT TTGATGATAT TGAATGGGGC GAAGATTTTG 60
 GTGCGCCACA TGAAACAGCC ATTGCTAATC ATTATGATTT ACCGGTGTTT ATTACTAATT 120
 10 ATCCAACTAA AATTAAGCCT TTCTATATGC AACCAAATCC TGAAAATGAA GAACTGTAT 180
 TATGTGCAGA CTTAATTGCA CCTGAAGGAT ACGGTGAAAT TATTGGTGGA TCTGAACGTG 240
 TGGATGACTT AGAATTGTTA GAACAACGCG TTAAAGAACA TGGATTAGAC GAAGAAGCAT 300
 15 ATAGTTACTA CTTAGACTTA CGTCGTTATG GTAGTGTGCC AACTGTGGA TTTGGTTTAG 360
 GTTTAGAGCG TACAGTAGCA TGGATTTCTG GTGTTGAACA CGTTCGTGAA ACAGCGCCAT 420
 TCCCAAGATT ATTAAACCGT TTATATCCAT AAGTTTTAGG GTCGTTACTT AAAATATATA 480
 20 TTATGATAAT AAAAAGAATC GTCCAGCCTA TATTAATAGC TGGGCGATTT ATTCTTGGA 540
 AATGAAGATT GATAACGGCT TATATTAATC ATGAACAAAC TTTTGTTTAG TAACTATATT 600
 TGGAAAATAA ATAGTTCATC TGATTATCCA TAAAGTTCAT TGTTAAAATC TAGTGTTAAA 660
 25 AAATACTGTT 670

(2) INFORMATION FOR SEQ ID NO: 877:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1258 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 877:

35 TGCAAAGTAA ACTGGATGGC TnTCTTGCCG CCAAGGATCT GATGGCGCAG GGGATCAAGA 60
 40 TCTGATCAAG AGACAGATCC TAACCTCTCA TTTCCGATAA GTTTAGTATG TATTATAGGT 120
 GATGATTnTC TTGTTTTCAA CCTTATACAC AATATTTTTG AnATTTATAT TATATTCATA 180
 TAAGGAAGGC GATTAATTAC GATTCATCAT TATATCTCTA CGTGTTAATT TATTAACCL 240
 45 TACTATATTT ATCaACaCAA CAGTTTTTCC TTTCATTTTC AAACATAACA GAAAAAGCCT 300
 GAGACATACT TAGTCCCAGA CTGTTCTATA ATTAATTAAA GATATAACAA GGTTAATCAT 360
 50 TCAAAGCTAA CATCACATTT ATTCACTGTA ATTTTTCATA AAGAATAATA ATGACTGTAA 420
 TTCTATACCT AAGTCAATTT GATGTACTTG CACATCTGAA GCGGTATTAA TTCTACCAGG 480
 AGTGAAGTTT AAAATACCTT TCACACCAGC TTGGACGAGT TCATCTGCAA CTTTCTGTGC 540

5 TAATTCATCG TTATCTTTAA CAATAACGTT CCCTATTTTC TGGCCAATAA CATCTTCTTT 660
 TACGTCAAAC GCTTCTGTAA TCGTCATATC GTCATGTATT GAAAAGTTAT ATGTGAGCAA 720
 10 AGCTTTCCCT AGGTTCCCAA CTCCGACAAT TCGGATTTTG ATCATGTCAC TCTCGCTTAG 780
 TTCAGATTTA AAGAAATCCA ATAACTATC TATATTATAT CCGTACCCTT TTTACCTAA 840
 TTCGCCAAAA TATGAAAAGT CACGACGAAT TGTtGCCGAG TCAATTTGTa ACGCATCGCT 900
 15 AATCGCTTTT GAATTTACAC GATCTATACC TTTAGATTTT AATGAACTGA CAAATCTATA 960
 ATATAACGGC AAACGTTTTA AAGTTGCTCG AGGAATTTTA ACTTGGTCAC TCATTGCTA 1020
 TTTCTCCTT CGTGTtTGAA TGAATTATAT CTATGTATTC AAACGAGTTA GAGCGTAATT 1080
 GTTAGATTAT TAAAAACAA ATAATTGaTT ACAATGAATA CCTAAATATT ATACATTAWT 1140
 GAATTTTAAA AATAAATATC TTAGTTTCAA GACGGCGTTT TATACTGTAA AATGGTTATT 1200
 20 AATGTAATAA TTACGCTTGT TTCATAAGTT AATAAAATT GAGAAGACGA AGGTGAAG 1258

(2) INFORMATION FOR SEQ ID NO: 878:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 546 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 878:

AAACAATCCC CnCCAAAGTA ATTCCACCAC TAATTGTAAA TAACTTTGGA AAGGATGTTG 60
 35 TTTGTnATTC CCATGGACAC ATGCCATCTn ATTTAAGTTT TATACTCTAT TTTTCTGATG 120
 TCTTGATTAA TTGCAAGCGG TCTTGCGAGC CTTGTTCAAG CTCATATTCT TTATCTACTT 180
 GCTTACCGTT TTTCTTATT AATCTTACGG TTGCGGTAAT CTTATCACCA TGTTTTTCAG 240
 40 CATCTATAAT TTGTGGTGAA CTAATCATCA TTAACTTCC TTTAGATACA CGCTTCTTTA 300
 CATCATCATA AAAAGATGAT CCTTTTTTTA TATATGATGA TACAAAATCA AAATCCGACT 360
 GATTAAACGC AGCATTATTC GCTAAAGAAT ATCCAGCAAA GAATTCTATC AATTTGTTCT 420
 45 TCAAGCTGTT TTCTTCTTTT TCTTTCTTTT CAACATAGTC TTCGATATCT TCACTGTCAA 480
 AATTCAAAGT TATCTCTGTA TTGTATTTTA AATCGCTAAG CTTTAATCGT CTTGCTCTGT 540
 50 GATGTA 546

(2) INFORMATION FOR SEQ ID NO: 879:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 870 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 879:

	TTTATGAATT TTTTATGTAA CATTAAATAT TCTTCTGGCA ATCTTCCTAC TATkTCTkTT	60
10	ATATATCCTT CTTCTCCTTC TTAAATACT AAACCTGTGG GAGTTAGTGT GTTATACCTA	120
	ATTTCTATCT CATTCAAATG ATCATTCAAA AGGTCAAAAA ATTCTCTATT AAAATAACCC	180
	CATTTCCCAT TTTGTAATAA ATATTTCTTT CCATCTTCCA TTTCAACTTC TGCATGTAGT	240
15	ATCGTTTTTA GTaGCATATT ATTTGAATGC CCCAAATTAT CkATCACKTc AATTCTGACA	300
	TCATTTATAG AATTAACGTC ATATTTTAAAT AAGTAATCTC CTATTTCACT GATGTAATCA	360
	ACTTCAGGAT CAGTCGCATC AAATGTCTCT AATGTATTAT TTTTAAAAGA ATTTATATAT	420
20	ATATTGACGT CGAGCATATC ATCTAACAAG AGTATCATAT TACTTAGTTC TAAAAATCTT	480
	GATATATCTA TAGATATATT TTCTGTAGTT GAAGAATTCG ATAGTTTTTT TAGCAATAAA	540
25	GTATCTAATA CTTCTATTTT ATTCAAGTCT TTAAAGTTA CTATTCTAGG AAATTCACTA	600
	ATTTTTTGAG GAAGATTAAT AATAGCGTTT ATTTCTTTGA TTATCACACT AATTTTATCT	660
	ATGaATTGCT GcTTTCTATT CGGtACACGC AATGaAATAC TTGTACCACA AGTCCATTGt	720
30	TTTTCCaAAA ATTTGAGGaT TCTGTGGaTG tCCTTGGaCT GGaTATATAA GATTCTGAAG	780
	GTCTAACGTA ATCTACACTA TTCCTTCTAT AATTAACAAT CTCTTTAAGC CTGTTTTGTT	840
	GAAAAAAATT AACATTTTTTA nTAACTAnGG	870

35 (2) INFORMATION FOR SEQ ID NO: 880:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 880:

45

	GGTTTACATC TTTATCAACT GTAATTTTCAT TGACATCTTC ATTCATATTT AAAACACCAT	60
	TAAATGTCCC TGAATATTCA CTTGTTTGTT TAACTGTGGC AGTGACTTGT CGTTACCATA	120
50	TGTCATCATA TATTGTGCAA ATGTTAAAGT CCCCATGAG TTGACCGACA AAGTTGAATT	180
	TATCGAATTG TATTCAGATT GTAAACTTAG CAGACATTTT TAACCACGAG CATTCTTATC	240
	CAAATAGnCC TTGTCTGTTA TTTTCAAGTC AATTTTCACA ATAGGATCAC TGCATCTTTT	300

55

ACCCTCTTTT TCTGCTnCTC ACATAACTTT CAGATGGACT

400

(2) INFORMATION FOR SEQ ID NO: 881:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 333 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 881:

AAGATATTGC	GAAAAGAAGT	GACAGTAACA	GAGGAATTC	AAAATTTATT	CAACCATATA	60
TTGAAAATTG	ATGAAATGGT	GCATGCAAAT	GAAGCGCGTA	TTGCGTATGA	AGCAGACATG	120
CGAAAACCTT	TATATAGTAA	ACGCATTTAT	CAAAATTTAA	CATTAGACTC	TATTGTTTTT	180
AGAAATACAT	TGAGATATAC	AGCGATTATG	ATGATAGCGG	ATATTTATTG	CGTTAATGTT	240
TGATTTTGAA	AAAGCATCCT	GGATCCCGTT	ATCTGCACAT	CCATATTCC	TAGGGCCATC	300
ANCCTATACC	ATGCCATCCG	AGAGAGGTAT	GGn			333

(2) INFORMATION FOR SEQ ID NO: 882:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 617 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 882:

ATTGAAAGAA	AAACTGACGT	ACTTATCCGA	TGATAAAATG	AAAGAAGTAG	ATAATGCACT	60
AATGATTAGT	TTAGGGCTGA	ATGCAGTAGC	TCACCAGAAA	AATTAGGCGT	CTATTATATG	120
TATTTTTCAG	AGATAAATAA	AATATTGATA	TAAAAGACAA	TAACTTTATA	ATAATTATAA	180
CTATTTCTAA	ATTCTGTACG	AAGAATTTTC	TTATAAACAA	AGATTTTAGC	AAATACCAGT	240
TATGATATTC	ATATTTTTTA	TTATAAAAGG	ATGTCCTAAG	TTTTTTAGGC	TTTAGGTATT	300
CCATCCTAAA	GTTTTTTTAG	CTTAAAAGTA	TCATCTACAG	CAAAATTGCA	AACGACAAAA	360
TTGATAAGTG	CAATTAAATA	AATGTTAGTA	AGTGAATCAT	AATTATCCTT	GCTTAAGCAT	420
TTGCTTTGTA	AGGGAAGTGA	GGAGGCAACT	AATCGTGGAA	GAATTAAAGC	AACATTATAA	480
GGGTTTAATT	GATGAAAGTT	TAACGTGCCA	AGATAAAGTA	GAATTGATAA	AAAAGTGTGA	540
GAAATACACT	GACGAAGTGA	TTCGTAAGGA	CGTCTGCCT	GAAGACATTG	TCGATATTCA	600

(2) INFORMATION FOR SEQ ID NO: 883:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 583 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 883:

GTTGTATTAA AATTTTTTCAA TTAAACAACG AAACAATACG AAGAAAAACA TATTACTGAG 60
 CCACTGGAAT TGACGAGCTT ATTAGGTANT ATTTCTCGTC TAGACAATGG TCATTTTGCA 120
 CATTTCATG CTACTTTCGG TACTCAATCA TATGAACTT TTAGCGGACA TCTCGCTAAA 180
 GCAATTGTTT CTGCAACTGC TGAAATCATT CTTACTGTAA CTGATTTGGA CATTCAACGT 240
 TCATTCAAAG ATGCAGTCGG TTAAATTTA CTTGATCCTC AATAGCACTT TACTTTTTAG 300
 TTGAATCAAA TTTTATTACA AGTTATCATA AGTGCTTAAT CATAAATGCA TTCTCGTTTA 360
 TTAATATACT TCAACGAGTT AAATTAAATT TTAAATTAGG AATTGAATGC GTTTATAACT 420
 ATCACAATTG aATGaATAAC TTTTTCATTA GAAAGTAGTG TkTATTATTG aTGCAATCTC 480
 GaTACATCGC TAGAATTTAC TTTtaATTTT ATTTATCGTC TCCCTTnTTG AGACTAAAAT 540
 TTTTCAATTT ATGACTTTAA ATTTGTTTTT AGCATATATA CCT 583

(2) INFORMATION FOR SEQ ID NO: 884:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 524 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 884:

GCAGAGGGCT TACGATAAAA AGAAAAATCT CGCTCGCCCG CTTCACTCAA ACTCACAAAA 60
 GcTAACGCAt GwyTGCTyCA TTCGTGCGAT ACACTTTTGA TACATCTACA CCAATACTTG 120
 ATATCGTTTC AATAATGCTA TCTCCAAATG CATCATTACC TAATTGTGTA ATCATAtACG 180
 CTTGTTGTCC tAACTTTTGA ACTGTACAAG CTACATTGCA TGGCGCGCCA CCAATTTGCT 240
 TCGTAAATGT TTGAACATCT TTAAATTCTG CATTGTGAAC ATTTGGAATA AAATCAATTA 300
 ACGCTTCTCC TATTGAAAAT AGACGTCTCA TTTATGTTCA TCCTTTAAAT CATATTTAGT 360
 AAATTGTAAA TATACTTGCC CTGATTCTGT TGATGTTTTA ATACCTAGTG CGTCTTCTGT 420

TGTATCAACA AAAATTTGTA ATTGCTTTAA TGGCGTGTCT AATA

524

(2) INFORMATION FOR SEQ ID NO: 885:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 885:

TTAATGATTT CTTCGTATGA TTCACCAACA ATTTGCGCTA AAATATCATG CGCAAGTACT 60
 TCACCTTCAA CCATAAATTC TATTTGCTCT ACTAAAAATT TTACGTCATT CATGCTATTC 120
 ATTTTGGAA TTAACAAAGC ATATAAGTTT GTAAATTCTT TTAAATACGC AGCATCAGCT 180
 TTTAAGTAAT GGCCTAATGC GTCACCTCTA TATCTCCGGA TAACATCTTC TGAATAAAGT 240
 CATCCTCATA AATATCATAA TGATGGCTTG CAGCTGGGTA CAATTTTGTG AAATTCCATG 300
 TAAAAATCCT CCCTAAATAA AAAACTACTT CCAACATGAA GAGTTGAGGC ATGTGCAAAC 360
 TAGCCCATCA CTCATAACAC TTCCACGTGG GACTACCAAC 400

(2) INFORMATION FOR SEQ ID NO: 886:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 857 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 886:

AAGAATTTAA ACAAATGAT ATTTGGAAAC ATTTTAAAGC TGTGAAAAAT AATCATGTTT 60
 ATGACTTAGA GGAAGTGCCA TTCGGTATTA CAGCAAATGT TGATGCTGAT AAGGCAATGA 120
 CTCAATTATA TGATTTATTT TATAAGGATA AAAAATAGTG AGTTGATATG ATGATAAAAA 180
 ATAAAAAGAA ACTACTATTT TTATGTTTGT TAGTCATTTT AATCGCAACT GCTTATATTT 240
 CGTTTGTAAC CGGTACAATT AAATTGTCAT TTAATGACCT ATTTACAAAA TTTACAAC TG 300
 GTAGCAATGA AGCAGTGGAT TCAATCATTG ATTTGCGATT GCCACGTATA TTAATTGCAT 360
 TGATGGTTGG CGCAATGTTA GCAGTTTCTG GAGCATTATT ACAAGCAGCA CTACAAAATC 420
 CTTTGGCAGG GCGAATATCA TTGGCGTTTC CTCAGGTGCA CTTATAATGA GAGCCTTTGT 480
 ATGTTGTTTtA TTCCACAATT GTACTTTTAC TTACCATTAT TAAGTTTAT TGGAGGTTTA 540

ATATTAGTAG GTGTTGCGTT ATTCGTATTA TTAAATGGTG TTTTAGAAAT TTAACTCAA 660
 AACCCTTTAA TGAAAATTCC TCAAGGCTTA ACAATGAAAA TATGGAGCGA CGTATACATA 720
 5 TTAGCAGTAT CAGCATTATT GGGATTAATA TTAACATTAC TATTGTCCCC TAAATTGAAT 780
 TTAATAAATT TAGACGACAT ACAAGCGCGA ALATCGGTTT TAATATTGAT CGTTACAGAT 840
 GGTTAACAGG GTTATTA 857

10 (2) INFORMATION FOR SEQ ID NO: 887:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 675 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 887:

GATTTATTTT AATATTATTG TTAGAAGGAA TTTTACAAA TTCAGCGAGT GCAATCGAAT 60
 ATTCAGACTT ACATCATAAA AGTAAGTTTG ATTCAAAGCG TCTAAGTAAT GCTAAGATGT 120
 25 CATTTCATCAA TCCAACCTCAG CTTGAAAATA AAAACACAAA CGATAGACTG TTGAAGCATG 180
 ATTTGTTTATT TCATGACATG TTCGTAAATG ATGATTGGAA AAAGGATTTT AAAGTTGAAT 240
 TTGAAAATGA GGCACTTTCA AAGAAATTTA TAAATAAGGA TATCGATATA TTTGCTGGAA 300
 30 ATTATGGATA CGGATGTCAT GGGGGAGCAA CCAATAAAAC GCAATGTAGT TATGGTGGTG 360
 TTACTTTAAG TGACAATAAT AAATACGATG ATTATAAGAA TATACCTTGT AATTTATGGA 420
 35 TTGACGGACA TCAAACAGAA ATAGAACTAA CTGCAGTAAA AACGAAAAAG AAAATTGTTA 480
 CTATTCAAGA ATTAGAGGTT CAATTAAGAA ATTATTTGAA TGAGAAGTAT AAGTTGTACG 540
 AACAAAGGTGG CGACATTGTT AAAGGGTATG TTAAATATTA TAATGATGAT GAACAAAATG 600
 40 TAGAATATGA TTTTATAAAT TTAAATGGTG AGTATGGTCG TGAGGTATTA AAAATGTATG 660
 CnGATAATna AACna 675

45 (2) INFORMATION FOR SEQ ID NO: 888:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 851 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 888:

AATTCCTAAT TAACAATGCC TTCATTAAAA AACCTTCTTC CATCAATTG ATGTTTTTCAT 120
 GATATTTAGG TTTTCTCTTT TCAAGATAAA CCATTAAAGT CTCAATAGAA GGTGTAGTGA 180
 5 TATTCACCTT AKGACCTAAA ATTTCTTTAA ATACTTGCGC ATGTTCTGTT AATCCTATAA 240
 ATCCCGTACG CTTATGTTCT TCATAAATAA GTAGAGATGC TAATTCATCA TCAAGGTCCT 300
 CAATAGCACC ACTCACAAC TGAATCTTT GTAACAGTTG TTCTTTTGAA CCATTTAAGA 360
 10 TAACTTCACC GTCACCTAAG TGAATGATAT AATCAGCTAT TTTTCTAAA TCTGAGATAA 420
 TGTGTGTCGA CATGAATATT GTTTTATTTT CATCAATTAA TTCTTGCTGA ATTATCTCCA 480
 ATAGCTCATT TCTAGCTAAG GGATCGAGAC CTGAAGTAGG CTCATCAAAA ATATACAATT 540
 CAGCATGATG TGAAAATGCA ACTGCTAGTG ATAATTTTCAT TTTCATCCCA GTCGAAAATG 600
 TTTTATTGA TTTATTGATT GGTAATCGA ACTTTTCAAG ATAAAATTCA AATACTTGAT 660
 20 GGTCCCATTT ACGATAAAAA GGGGCAATCA TTTTTCAGG TTGCTTCGTA GTCCATCTCT 720
 CATTAAATA ATTTTCAGAG TAAACnAACC CAATCTATT CTTTAACTCT ATCGGATTGA 780
 GAGCCATATC CTCTTCTAAT ACTCTAATAA CACCAGTTTG TGGGTGATAC AAATCCATAA 840
 25 TTAACCTAAT n 851

(2) INFORMATION FOR SEQ ID NO: 889:

(i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 362 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 889:

AATCACCTTC ACGCCAATAT TGATTTTCAT GCGTAAATAC TTGTGCCGTT TCATGATACT 60
 40 TTGTCAATCG TGCGTGTTGC TGGGGGGAAT ATTTTTCAGT AGCCCAATTG GCTGCATGAC 120
 CTTCAATGGC TAGTTCAATT GCAGGATTAA TTAAATCTTC CAATGACAAT TTAGCATAAC 180
 GCTTGTGAAT ATAATCAAAC AGCTTTGGAA TTGCTGGCAC ACGACAGTTT TACCATGTGT 240
 45 AGTCATATCA AAAAATGATT TATATTCGCC TGAATCACT AGATAAAATT GGTGGGCTAC 300
 ATGGTCAGGG TGCTGGCTCA CGGTGCATCA AACGcGnTA ATACTGCCCA GTACCTTTGC 360
 50 TC 362

(2) INFORMATION FOR SEQ ID NO: 890:

(i) SEQUENCE CHARACTERISTICS:
 55 (A) LENGTH: 747 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 890:

	TTGGTTTATT CCCAAAAAGT ACGCGAAAAT TAGTAGAGAA GAATTTGAAA GTTTAAATAT	60
10	TGTCAAACCC GCTAAAAATA ATACTTTCTG GCCTGTTGCA GGATTTGCAG TGTTATTAAC	120
	AACCTTAACA AGAAAATATA TCTATTTGCT TAACATCCAT TTAGAAAAAG AAATAGTTAT	180
	ATTAACATGC TGTATGATAC TTCTAGGTGT TTTCGCATTG TTTATATATA TAAATACAAA	240
15	ATTGAAGTTA CATATTTTGT ATAAAAATAA AAGTAATAAC GAAAAGATCA TATTAATACC	300
	TACATTTAAA AATATTTGTT TATCCTTATT TGCTTATATA TTATTTGGTG GATTGTCAAC	360
	AATGGCTCTG AGTATGTTAG TAACTTCATC CCCTCAAAAT ATAATAGAAT TTCTTGCTTT	420
20	AATTGGCATG ACTGCATGCT TCTTTCTACT GAATATGTCA TCGGTTCTAG ATAAAAAAAT	480
	TCATGTTATT TTAAAAACAA ATAAGTAGTA AAATTGATTA ACTTAGGTAG TATCGGATAC	540
	TTAAATGTTG GTTCATAAAA AGCAATGATT TTAAATCGAG GAGCTATCTT AGAACAGGGA	600
25	AATAAACAG CCAAAGTTAT AAAAAGTGAA TTAATACTA ATTATATTAT GTTAGCCACG	660
	CTTCAAATAA AaATAATTA GAATAAGGTG GGATTGATAA TCAAnGCTAn GCGAATCTAA	720
30	AATCATCAAT AAAAACCCCA AATATAG	747

(2) INFORMATION FOR SEQ ID NO: 891:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 526 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 891:

	ATTTTGAAAT TTTAGCAGAT ACTGATAATG GTCTCGATGC AATGAACTT ATTGAAGAAT	60
45	ATAATCCTAA CGTTGTTATT TTAGATATAG AAATGCCAGG CATGACTGGA CTTGAAGTTT	120
	TAGCGGAAAT TAGAAAAAAG CATTTGAATA TTAAAGTGAT TATTGTAACA ACTTTTAAAA	180
	GACCGGGATA CTTTGAAAAA GCAGTTGTGA ATGAtGTGGA TGCATATGTT TTAAAGAAC	240
50	GTTCTATAGA AGAATTGGTG GAAACCATTA ATAAAGTAAA TAACGGAGAG AAAGAATATA	300
	GCGCCACATT GATGACTTCA TTTTTTGTAG ATAAAAACCC ATTAACGCCC AAAGAACAAA	360
55	TTGTATTAAAG GGAAATTGGC AaTGGTTTAA GTAGTAAAGA AATAAGTGAA AAATTATTTT	420

ATCGTTTTGA TGCTTGGGAA AAGGCAAATG AAAAAGGCTG GACCTA

526

(2) INFORMATION FOR SEQ ID NO: 892:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 751 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 892:

15 TATTTGATGT AAAAGCGAa GAACCATACA ATGTAACAAT TACTAGTGAT AAATACATCC 60
 CTAATACTGA TTTGAAACGT GGGCAAGCTG ATTTATTTGT AGCGGAAGGT TCTATCAAAG 120
 ATTTAGTGAA ACATAAGAAG CATGGTAAGG CAATTATAGG AACGAAAAA CATCATGTTA 180
 20 ATATTAAGTT ACGTAAAGAT ATTAATAAAA TCTATTTTAT GACAGATGTT GATTTAGGTG 240
 GACCAACGTT TGTCTTAAAT GACAAAGACT ATCAAGAAAT AAGAAAGTAT ACAAAGGCAA 300
 AGCATATCGT CTCTCAATTT GGATTGATT TGAAACATAA AAAAGATGCT TTAGCATTAG 360
 25 AAAAAGCGAA AAATAAAGTT GATAAATCTA TTGAAACAAG AAGTGAAGCG ATAAGCTCAA 420
 TATCAAGTTT AACCGGAATA TTATTATTTG TAACATCATT TTTAGGTATT ACATTCTTGA 480
 TTGCTGTATG TTGCATTATA TACATAAAGC AAATAGATGA AACCGAAGAT GAGTTAGAGA 540
 30 ATTATAGTAT TTTGAGAAAG CTTGGATTTA CACAAAAAGA TATGGCAAGG GACTAAAGT 600
 TTAAATTAT GTTTAATTTT GGGTTACCTT TAGTTATTGC ACTATCACAT GCATATTTTA 660
 35 CATCATTAGC ATATATGAAA TTAATGGGTA CAACGAATCA AATACCGGTT TTCATAGTAA 720
 TGGGATTATA CATTTGTATG TATGCTGTTT T 751

(2) INFORMATION FOR SEQ ID NO: 893:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 357 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 893:

50 GGATGTTGTA CATTATTCAG ATTACTTTGA AGGGGCACAA AAATATTTGA GCTATTTAAA 60
 ATCAACAGTA GATGTTAACT TTGAAGGTTT GAAAATTGCT TTAGATGGTG CAAATGGTTC 120
 AACATCATCA CTAGCGCCAT TCTTATTTGG TGACTTAGAA GCAGATACTG AAACAATTGG 180

TAGCTGAAAA AGTAGTTGAA ACTGAAATGA TTTTGGGGTT AGCATTGGA CGGCGATGGG 300

AGACAGAnCC ATAGCAGTAA GATGGAGAnG GnCCAATCCG TTGACGGTGG ACCCAAT 357

(2) INFORMATION FOR SEQ ID NO: 894:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 464 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 894:

GCTAACCTGC AAAATTCGAT AAGCGATTCA ATaCTGAcGC CTGcATTTCG CAAATTTTTTC 60

ACTAAATAAA TCCAATTTAA ATCACTATCA TTATATATTC TATATCCATT TTCATCTCGA 120

TTAACTGGTG GAATCACACC AACCTTTTCA TAATATCTTA AAGTGTCTTG AGATATATTC 180

ATGAGCGCTA CGACTTCTTT AGTTTTCAAT GCGACTATCC TTTCAGTTAT GTTTGGTCGT 240

CTAAAGTAAT GTTGCTTTAT ATATTGTCAT CTTGTTTTGA ATACTTCTTA TTTTATTACT 300

CAAATTTAAA TTTGTCTCTT TTTTAACATT TTACATTTCA TCGTTTTTAA TTACTTTAAA 360

AATTGTATAA CTkAAATATT TAAAATGATA TAAWCACTAA GATTGAtAAT AtTTAATTnT 420

TTGGAAAATT ATnTTAAGTT GCCAATTTTG GGGGGATGCG GAAA 464

(2) INFORMATION FOR SEQ ID NO: 895:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 599 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 895:

AnAATGAAAG AGTTACAGGA CTTTATTGCT CGTTTCTCAG TAACGCTTCT AAATCTAAAC 60

AAGCAACAAG TCGTAAAAAA CAACTCGAGA AAATTGAATT AGATGATATT CAACCATCAT 120

CAAGAAGATA TCCTTTTCGTT AAATTCACAC CTGAGCGCGA AATCGGTAAT GACTTACTAA 180

TCGTTCAaaA TCTATCTAAA aCGATTGACG GTGAAAAAGT ATTAGATAAT ATTTCAATTCA 240

CAATGAATCC AAATGATAAA GCAATTTTAA TTGGGGATAG TGAAATTGCG AAAACCACAT 300

TGCTTAAAAT ATTAGCCGGT GAAATGGAAC CAGACGAAGG TTCATATAAA TGGGGTGTA 360

CAACGTCATT AAGTTACTTC CCTAAAGATA ACTCAGAATT CTTTGAGGGC GTTAATATGA 420

CGGGCTTCTT AGGCCGTATG CTATTTAGTG GAGAAGAAGT TAAGAAAAAA GCTAGTGTAC 540
 TTTCAGGTGG AGAAAAAGTA CGTTGTATGC TAAGTAAAAT GATGTTATCA AGTGCAAAC 599

(2) INFORMATION FOR SEQ ID NO: 896:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1057 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 896:

TCATTAGTAG AAATCGTTAA AGATGCATCT TGCAATTTAA AATCCGAATC TTGTATTGT 60
 TTGTTACACAC TATCTACGTT TAACAATGTT GTCATATCCA TGCTCCTCTT TGTTTAATTT 120
 TAATAAAAAC GTCTTCTCTT CAATAAATAA CTAAAGGCGT TATATATGAA AATAGCAATG 180
 ATAAGTAGCA AACTAAAGCT TTGATTAAAT CCAGTAGTAA TTGAATAAAG CCCAATTACT 240
 ACAATGGCTA GTATCAATAT TGATAAGTTG ATCGCATATG TTTTAAATAA TGCAATTAAT 300
 TCTATATGGC GTTCACCTTC ATCCAATATT TCCAATCGCT TTTCAGTGTA ATTCTTATCT 360
 GCAATTTTTG GTATTCTGTC ATCAAACCTT CTATTAAACA GTGTAAATTG TGTATTGAAA 420
 ATAGCACTAG CAAAAAATGG TATTAATAAG AATAGTATTG CATTTGCTGC AGCGTGCCCT 480
 ACCACAAAAA TAAGCAATAC TAAAAAAGCA ATAAGTGTCT GCAGAATACT TAATATACTT 540
 CCATTTAAAA CATAACGATT CGCAAGCAAC TCATATTGAT CTGTTGCATC GATATCTACC 600
 TCTTTATCTA CAAGTTGCTT ATACTTCAAT GCACGTCTTT GATTCATCAA AACAATGGCT 660
 TCTACCAATA TGATAATAAT CGTCGCAACA ATCGATATTA CAACGACATT GTTATATGTC 720
 GCAAACGTCA AGTTCTCAAT ACCAAATCCA TCAAAATTAC CTAAAATTCC ACCTATGATA 780
 CCCCCTACAA GTCCACCCAG TAAAAGATAT CCGATATATC TTAGTATTTT CATGCCTCAT 840
 CCTCCTCAAT AATAAAAACA GTTCCACCG TTTCATTGAA AATGCGAGCA ATTTTATTG 900
 CCGTTAATAC TGATGGCATA AAATTGTTTC GCTCAATTAG CGATATGGTT TGTCTTGAAA 960
 CGCCCGCTTG TTTAGCAAGT TCGGTTTGGT TTAAGCCATC TCGTGCTCGT AATTCTTTCA 1020
 ATCGATTACG CACATCGCAT CAACTCCnTA ATTACAC 1057

(2) INFORMATION FOR SEQ ID NO: 897:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 336 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 897:

5 AAATTATTGA GTTAACAAAT CATTACGGAG CACATAATTA TTTACCATTG nCAATTGTCA 60
 TTTCAGAAGC CGAnGGGGTA TGGGTAAAG ATCCTGAAG CAATAAATAT ATGGATATGT 120
 GAnCTGCATA TTCCGCTGTT AACCAAGGTC ATAGACATCC GAAAGGTATT CAAGCATTAA 180
 10 AAGATCAAGC TGATAAAGTG ACTTTACTCT CACGTGCTTT TCATAGTGAT AACTTAGGTG 240
 AATGGTACGA AAAAATTTGT AAAGTGGCAG GTAAAGATAA AGCTTTACCA ATGTAATACA 300
 GGTGCTGTAA GCAGTAGAAA CAGCTTTGAA AGCAGC 336

(2) INFORMATION FOR SEQ ID NO: 898:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 302 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 898:

25 TGnTGnTTTC TTATTGTTGA CCAATTATTT TGCATACCAA ATCCAAATCA TGTATTTATC 60
 AATCACATCC TAATGCTAAA TCCAATGTAT TACATTGTTT AATGGTATAG CACAATCTAT 120
 30 CATATTTGGT ATATCAAGTA TGGAAACAT TCCATATCAT TTTTACTTTA TTTTATTCTT 180
 ATGTTTAATA GCTGCAGTAA ATTTCTGATT AGCAGGTATA CGACACACGG CAATTTATAA 240
 TAAAACATCT AAAGTGACAC AAAGTATAA TCAACAGGGA GTTTCAAAGA TAGTCAGACG 300
 35 GG 302

(2) INFORMATION FOR SEQ ID NO: 899:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 469 base pairs
 (B) TYPE: nucleic acid
 40 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 899:

CCACTAGCTG ATACAGTAGG TATACCAAGA TTCGTCATCG TTACGACATA TCAATTTGGT 60
 50 CAATATGCGA TGTTATTCTT AGCGCCGACT GGACTTGTTA TGGCCACACT ACAAATGTTA 120
 AACATGCGAT ATTCACATTG GTTCCGATTT GTATGGCCGG TAGTTGCTTT TGTATTGATT 180

55

TATAAAAAAT ACTAATGGGG TTTTATGCAT CTCGTAGGTT TGTAGAAATA CTAACTAAG 300
 CGAGGTGCAT TATTATTTTT GATTAAGAAA ATAATGACGG TAATGATAAC ACTAGTAAGT 360
 5 AATTGATACA ATGCTCTATT TAATAATGAT ATTTTAAAAA TTTGTTTTTA ATGTAATGTT 420
 AGATCTATGG TATATTATAT TTAACGTGGT AAATATGTAT TTGCTGTAA 469

(2) INFORMATION FOR SEQ ID NO: 900:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 821 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 900:

CTTTATTGGT AACTCTCTAT aATATtTtAa CATTTTACG CTATCGTGCA AACGCAaCcA 60
 CTTTGAATTy TCTGATATTT TTAGCATATA ATTTACACCC TGCACGATGA ATTGTTAATC 120
 CAGTTGTATA AATATCGTCA ACGAGTAATA TTTCTTACC ATTTAAATCT AATTCCGTAT 180
 25 CAATAATAAA TGGATTTTCA TCTGCCAGAC GCTCTTTCTT AGTTAAATGA GACTGTTTTG 240
 GTCTATTTGA CATCTTTAAA ATCTTATCaA AGCGAATCCC TTTAGCTTTT AGTACTGCTT 300
 CTACCGGGTT AAATGTTCTA GATAAATCAT GTGCCGGCGA AGAAGGAATT GGCACAATAT 360
 30 AGTCATAAGA TGTTTGTGGT ATTTCAATCA AATGTGCCAA TAATTCACAT AAATAATAGT 420
 CTTTCAAAAA TTTATACTGA TGTATCATCT CTTTCATTAA ACCGTCAATAT TGAAATTGAC 480
 AATATAATTG TTCCATTAAA TTAAAGTGTG CCGATAGAAA CTTGCAGTCT AAACAATACG 540
 35 CTTCACTTGT ATTTAAGTGT TTTAAGCACC TTGAACATCG CCTTGCTTTA ATATCAAGTT 600
 TAATATTGTC CCAATTCTCT TTGCATCTGT CACATAATCT ATTAGGTTTC TTGAACAAAT 660
 TATAAATGGT TATATTTTCA TATAACTTAG CACCACAAC TAAACAATTA TTCATCAATC 720
 40 CAACCTCTTT TTAATGCTAA TTTGTTTCATC CTTTGAATCT CTTTTTTAGC TTGAATCATA 780
 TTCATACTTA CCCCTTnCAT GAAAAACAA TACTTTTTCC A 821

(2) INFORMATION FOR SEQ ID NO: 901:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 349 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

TGATGCTATT AAAAAGTTAA TTGTGCGGTC TGTTTTGATA ATTTTAATAA TAACTTCAGG 60
 TAACTAAAAA TCCTAATATT GAAAAACAA AGCCATTAA AACATAACCT AGTATATTCC 120
 5 ATGTATGATT GTAACCTATT TGCCAGTtnt GTACTGACTT GCATAATTCT GTCACGTnCG 180
 AACCATGTAC AAGCCTGCAA CTACTGCTGA ATGATTCTCTG ATGCGTGAAC AATTGAGCAA 240
 TTAAATACGT AACAAATGGT GTTACAATTG AATAATTGAA ACATATTAAAT GTTTCATATC 300
 10 CTCGACGnCA TCAATGTTAA TCCGGACCTT ACTAATGCAT ACCTATAAG 349

(2) INFORMATION FOR SEQ ID NO: 902:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 390 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 902:

CAAGTGGTAT AAATTTACCT GGTGAAATGT AGGTCGTGGT GGCGATGATA CATTATTGCG 60
 25 TAAAATCGAC GCGTTGTTA AATTCGAACG TAAAGGTCGC GACAAAAAAC AAGTTTCTGT 120
 ATATGCAGTT AGCTGAATAA TTTTGTCTAG TTAACACCAG AAAGTGAATC TTCTGGTGTT 180
 TTTTACTTTT TATAAAATGT TTTTCATTAT TATTTTCATT ATGnTATTTA AAAATGGGGA 240
 30 TTTTAGACGT TATACTAAAT GTGCACTGTA TAGGGGCCCT AATCACTAAC TATAGGGGGA 300
 CAAGGATACA GTGCAGCGTT AAGGATAACT GnCCACATTG GTCTGGGAAT ATAGGATTTA 360
 AGCAAGGTTA TAAAGTACTC nTAGGCCCTA 390

35

(2) INFORMATION FOR SEQ ID NO: 903:

(i) SEQUENCE CHARACTERISTICS:
 40 (A) LENGTH: 334 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 903:

GGAATAGACG TATATGCGCG TTGAATTTAA TCCTAATAAG CTTTCGCATG ATGAAGTGCT 60
 TTGGTTAAAA CAAAATATCA TCAGTTATTT GGACGATGTT AGTTTTACGA GATTAGATTT 120
 50 GGCTTTTGAT TTTGAATTTG ATTTAAATG ACTATTATGC ATTGTCAGAT AAGnCGGTAA 180
 AAGAACTAT ATTTTATGGC CGAATGTAAA ACCAGAACAA AATATTTTGG TGnCGAATA 240

55

GATTCACATT CTATGCGGTG GAATTGATTA AACG

334

(2) INFORMATION FOR SEQ ID NO: 904:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 334 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 904:

CAAGTTGGGG AACTGTTTC AAATGAAACA ACAGTGTGTA TTTAGAGGC AATGAACTA 60
 TTTAATGAAA TTCAAGCAGA AATTCAGGT GAAATTGTTG AAATCTTAGT AGAAGACGGA 120
 CAAATGGTAG AGTATGGCCA ACCGTTATTT AAGGTGAAAT AATGAAAAAG GTTTAATTGC 180
 AAACCGCGGT GANTCGCACT TAGGATTATT CGCGCTTGTC GTGATTTAGG CATCCAACT 240
 GTTGCATCTT TCTGAGGGGA TAAAGATCGC TACATCTCAA ATGCTGATGA GGCATATnCG 300
 TGGGCCnCTT GTCTAAGGTC ATATTTAATA TCCG 334

(2) INFORMATION FOR SEQ ID NO: 905:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 905:

TTCACCCTGT AATTCTTaAC CGTCAATTGA CCTTTATGCA GATTTAATAT TCCTAAAACG 60
 TATAACTCTT CTAATGATAA TTGACACATA TTTAATAAC ACTTTAGATT CCGCAATAAT 120
 TCTTGTGACG ACAATAAATA CGATACATAG TCACAATAA AATCAAATGT ATATTGACCA 180
 AAATAACGTT CAATATAGTC ACTAAGCTCA TTATTCATAT ATTCTATTTn ATCTATATGC 240
 ATTTTATTAA TCGAAATAGT CAATTTACGT TGATCCCTTT GATCTCTTCC TTTCAATAGC 300
 CATTGATGAT TGnTAAATA ACTTAACATT TGAAGAAGGG CCATCTTAGA TTGGATTCT 360
 TTCATnTTA ATAACTATC TATTGAAGCG CTATTGGAA 399

(2) INFORMATION FOR SEQ ID NO: 906:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1478 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 906:

5 TACATTTTTC AATTTATGGC AAAATnCACa ATTTTCATCG AAATATTTTG TAnAnAATCA 60
 TTGCTAGGTG ATGAGTTATC TCGGTTAATT ACAAAACGAT ATCTAAGTTT AAATATTAGA 120
 TCATAAAAAA ACTATCTACT GAATATCCCT CCATAGCATC ATTTTCTAA GCAAAAGTTA 180
 10 AACTTAGTTA GATAAAATGC AAATGAGATT ATTGTAGATA GTCTCTTTT AATGTTTAAA 240
 AATGATTAAT GCCaTAAACA TAAATAATTC CaAAGAATGC CGCGCCACTA AGCACTGTTA 300
 ATATAAGACT TAGTAACAAT GTGCGTTTAT AATGTTTaAC AAATGCTaTG AACAAAATGA 360
 15 CTACATTATA CGCAAAGAAA AGTCCGAATA ATGTCATTGT TATTTTCAAA TCAGAATAGA 420
 AAATATTTTAA TAGTAATACA ATGACTGCAT AAATTAAATA CGGAATAACT ATAAATTTGC 480
 GTTCAAATCT TAATTGTTTCG AAaCGtTTAT CTTCGTTTGT CATGATATGC TCCTTTATGA 540
 20 TTTAAAAGTA AATAAGTTTA CGGAAtaTCT TGTCCGaTAA TAGCGGTGTA AAtGTCAAAC 600
 CACGACTGAT CATCtAAATT AAGTkGTAGC CctTCGATTG CTTGATCAAT ACGCTTTAAC 660
 TGACTTGTTT CAAGTATCCG GCATGATACG ATGCGGTATT TTTACTAACC ACGCTATCAT 720
 25 CACAGCTGTG TCACTyACAC mATATTTGTG AGCTATTGAT TGAACAACtT TCATAATACG 780
 TTGCGCTTTA ATATCTTCCT TGTCGAAAAT TTTACGCGCT GCAAAAGGAC TCCAAGCCAT 840
 GATTtGAACA TGGTTTTGAT ACATTGAATC CATTGTTCCA TCTTGTAAC TATCAACGTG 900
 30 ATATGGCGAT AATTCTAATT GATTGATGCT AATATGTAGT CTTTCTTTCA TAATATATTG 960
 ATTTAACAAT TGGTATTGTG AATGATTAAA ATTCGACACC CCGAATGACT TCAACTTACC 1020
 35 TTGTTTAACA AGTTTAGTTA ATGCATCAGC AACTTGTTCT GGaTCCATCA ATGGTGaAgG 1080
 ACGATGAATG AGTAGACTAT CTAAATAATC TACATTCAAA TTGATTaATG ACTGTTCAAC 1140
 AGATTTcACG ATGTGCTTAC TACTCAAATC ATAACGATGT CCATTTGTAA AATCAAATTG 1200
 40 CTTAGAAGGC AAAATGATAC CACATTTcGT AACAAATTGA ATTTTATTTC TTAATTcGGG 1260
 TGATAAATCC AAAGCATTAC CAAACAGTGA TTCACATTGA TAATCTCCAT AAATATCAGC 1320
 ATGATCCATC GTTGTAATTc CACGTTCAAC TAATTCATTt ATAAATAAT TTAACTCTTT 1380
 45 CGCAGTCATC TTCCATTcAT TTGCACGCCA AAAACCTTGT ACAAGCCTAG AAAAATGAAC 1440
 ATAGTGATTA ATCATTATTT GTTCCATATn TCATCCAC 1478

(2) INFORMATION FOR SEQ ID NO: 907:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 907:

	TTCTATGAGA ATAGATATTG TTAAATTAAG AAAGTAGTCA ATTTTATTAT GACAAAAGAA	60
10	AGAAGAACAT TTAGTTCAGA GTTTAAGTTA CAAATGGTTA GATTATATAA AAATGGTAAG	120
	CTTAAGAATG AaATTATACG AAAGTATGAT TTAAAACCCCT CAATTATCTC AAATTCGATA	180
	AAACAACACC AAAATACTGG ACCCTTCAAT CATCAAGATA ATTTAAAAAG TGATGAAAAA	240
15	GAGTTAATAA AATTACGCAA AGAAGTTCAA CATTTAAAAA TGGAACATGA TGTTTTAAAG	300
	CAAATTTTAA AGTTGATTTA GAACTGAAC TTTTAAAGA TACATTACTT GAAGATTCTA	360
	ATTTATCTGA TTATATAAAG AACCAACTTC CAACGATATG GAGAGACTAA GAAATAGTTT	420
20	AAATTTATTT AATTTCAATA AAAACTCATA ATAATATTTA AGTAAGTTAG ATTGCTTTTA	480
	GCATTAGGAA TTCGCTTATA ATTAGATGTT CAATATAGAC TTTTACAC ATACATGAAC	540
	TAT	543

25

(2) INFORMATION FOR SEQ ID NO: 908:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 751 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 908:

	TTTnCGGTnG AACATGATtn CGATATTACT TCTAAAACGA TTCAACCAAA ACCAGGAGAA	60
	AGTGTGGAC GTTTAGTTAA TGTCGATTGT AAAGATACAG TACTGGAGTA TTCAGAATTA	120
40	GATCCAGAAG TTGCAAATCA ATTTAACAAT GCTAATATCG GTATTCATGC ATTTAAACTT	180
	GGTTTTATTT TAAACGCTGT AAATCGTGAG TTACCTTATC ATTTAGCTAT TAAGAATTTG	240
	AAGCAATTAG ATGAGAACTT TGGTGTATT GAACAACCAA CTTTAAAATT CGAATTGTTT	300
45	TATTTTGACA TATTTACTTA TGGAAGTAGT TTTGTCACGT TACAAGTTCC TAGAGAAGAA	360
	GAATTTTCAC CTCTTAAAAA TAAAGAAGGT AAAGATAGTG TTGCAACTGC AACAGAAGAT	420
50	TTACGTAGAA TGGGTTTAAT TTAATGATAG GCGGTGACTA TGATGTCCAA ATCAACTAAA	480
	AAAAGTAATG ACACAATCGT AGAAACATTT AAGGATATCA TCCCTTTATC ATTTGGAGAA	540
	GAAATTGGTA ATGCGGCATC TCATGGTGTT GCAGCTTTGT TAACATTATT AGTGTTACCC	600

55

TATGTTATTT CTATCTTTAT GATGTTTATT TCATCTACTA TTTATCATTG TATGCAAAAT 720
 GAAACACCTC ATAAATATAT TTTAAGGATT A 751

(2) INFORMATION FOR SEQ ID NO: 909:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 909:

AAATTGTCAC AATTCAGTCT AATTTAAAAG GAAGTTAATT AAAAAATTAA AAAATCATGC 60
 ATCGTGTGTCAT TTAGATATTG AAAAAGATGA GTCAAATTTG TTTTCATATAT TTTCCAATAA 120
 CAAATTATTA TATCTAAATA TACAATCAGA AATATATGAA TAGAAATATC AATAAGAAAA 180
 ATAATATGAT TAAAAATGAT GAATGGCATA CTTATAAAGT GTCTAAATAT TGGCGGTCAA 240
 TATTACTTAC AAACACGAAT GTTAAGTAAT GTAGACAATG cTGaAAAAAC AATaAGTAGa 300
 AACTCTATAT AGAATGTTTCG TTTTATTTTT TaATAATTGa ATAATCaTTA TTGGGTGGAT 360
 TTTAATATGG nATTCCGGTC GGCACACGGA TTAC 394

(2) INFORMATION FOR SEQ ID NO: 910:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1022 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 910:

TAAAGCTAAG AAAAAAGAAA AACGTAAAGC TTATAACCAA CGAATGAAAG AACGTAGGAA 60
 AAATCAACCT AGCGCAGTTA GTCAACGTCG AATGAATTTT GAAGAGCGAC GTCAAATTTA 120
 CAACAATGAT ATTTCTGAAG AACGCAATTC AAGTGAAGTT AAGGACAAAA AAGAGCAAGA 180
 ATAAATATTG ATGATACCCA GATTAGTAAG CAGAGGTCTT TTGCTTATTA ATCTGGGTTT 240
 TTATATGAGG TTAATTATCG ATAACGTTTA ATTAAAGTGT TTAGGTGTCA TAATTTTAAA 300
 TGACGATTTC CCCATTACKA TACACCTAAA TtATCATCAA TCTGaATTCA GATGTTTATk 360
 ATAAAAATTA GATGAAAAAT ATGTTAATAT ACAAGkaATT TAATGTGcGk ATATCTATAG 420
 GCTCGTAGTA TACTTATTAA AATATTAAAC AAAGAAAGGA TTTTAAGATG AATAGAAAAC 480

GGCTATCCTA TTTTCTTTA AAAAGTGGTA ATGCATCACA ACGTGAAGAA TTAGCGAAGC 600
 AATTATCTCA GAACGGTGGC AAGGTTTCTT TAGATATGCT TCAGACAACA ATGGGTGCAT 660
 5 TAGCAATTAT TTTATTAATT TCAACACTTT ATGGTATATT TGCACAAATT TGTATTAAAG 720
 GACGTAGAAA ATTATCGATT ATACTTTTTG TTATCGCGAT AATTGTAAGT TTGATGGCTC 780
 10 TTAATTTAAT TGCAATTGTC TTATGGGTTA TCGTGATGAT TATGTTGATT TCTAAAAAG 840
 AATCAAAAGA AACAACACAT AAGGACGATG AGTATATTTA TCATTAATAT GTTCATAGCA 900
 AAAGAAAAAC CATTAAAATG TTAAGTTGTT AATTATTAGA TACAACCAAC ATTTTAATGG 960
 15 TTTTATTTTT AACTTTGTAG TTCTTGAAT GTTGAACGA TTAAATAGAC ATTTAAAATA 1020
 CT 1022

(2) INFORMATION FOR SEQ ID NO: 911:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 274 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 911:

30 GGTACTTTAT TTTTATTTTT GGTCTTAGTG ATTTTACAT TATTACATA TAAAGCGCCT 60
 AATGGTATGC GTGCCATGGG TAGCATTAGC TAATGCAGCA ATCGCAACAT TTTTAGTGGA 120
 AGCATTTAAT AAATATGTnG GTGGCGAGTA TTCGGTATTA AATTTTGA AGAGCTAGGA 180
 35 GACGCTGCGG AGGTCTAGGT GGTGTCGCTG CCGCTGGATT AACAGCATT GCTATCGGTG 240
 TGTCAACAGT ATATGCATTA GTTATAGCAG CCGC 274

(2) INFORMATION FOR SEQ ID NO: 912:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 679 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 912:

50 CTTTAAATTGT TCACTTTATC ATTATTATCA ATCAACTCTA TGATATGAAA TTGTGTTAGA 60
 GATAAATCTA GTTCCTCATT TTCCTTTTTA TACTGTCTTC TTTATCGGC ATTTTCTCTT 120
 TCTATAATAA ATTGTTGTAA TTTTAAAATA ACTTCTCTCT TCATATCCAC TATTATTCCT 180

TTGTTGTAAT TAACATATAT TTAAGTAGTT ATGCACCATT TATTCGCGAT TTCACAAATC 300
 CACCTATAAT ACAAACTATC ATTTTAAATA AGCAGAATTT ACCCTGTCTT CCACACAAAT 360
 5 TGTTCGAC ACTTTAATGT GCCTACTATC CAAAAAGTT ACTCATAACA AATTGCGCAT 420
 TTTAATAAAT GTAGCAATGC CATTGTCATG CTGCAAAAAT ACTCGTATTC ATATTTATGT 480
 10 TTTATATTAT TAAAGTTTTT TAAATCCATA TTTTATAAC ACTTGCTATG TGATAAAATT 540
 AATTTTATAT ATAAATCTT AAGATTCAGA TTATTTAATA GCAAAGGAGA TAGTGATATG 600
 GATGTTTTAA CAATAGAACA TTTAACAAAG AAGATAGGCA ACAAACGAT TCTCGAAGAT 660
 15 GTATCATTTA AGCTGAAAC 679

(2) INFORMATION FOR SEQ ID NO: 913:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 564 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 913:

GCCATCCTTC AGTGCTAATT TTTTCAAATT CTAATGCAGA TGCACCACTA ACCATACCAT 60
 30 ATATAAAGAA AAAGACCACG AAAGnAATAA GTAGTAGTAT AGATGCACAT ATCATAATGA 120
 CATTGGAGTT TTCAGCGCCA AATTTAAATG TCAATGsGAA AATGATAGAG TATGCCCTTA 180
 TGATACCAA ACTCATTAAC CCTGACATAA TACCAATCAT CACACTTTGG GTCACAATCG 240
 35 TAGTCACAAC TAATCCAATC ATTAAACTTG CACCGAATAA GATTAAATAA AAGGCAAAGT 300
 ATGACTTAAT ATAATCACTA CGTTTAACCG GTAAAGTAGA TACATAATAC ATCCATCTTG 360
 AGTCTTTTTT aTGTTTAATA TTATCAGTAA TAGGTGTGAT TAACATAACC CCAGCCATkG 420
 40 CCGAACTCAT CAACGGATTA AATACTGCAA AGTATCCTGC AGCTATAATA GCGACTATAA 480
 AATAAATATA TGTTTGcNTT CTCGTTGCAT AAAAATACT TAGGAACATA CCTTTTATTA 540
 45 TACTTCACCT CGCATTATGA TTTT 564

(2) INFORMATION FOR SEQ ID NO: 914:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 643 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5
 10
 15
 20

AAAGAAGAAG CAAGTGCAAA TAATTTAAGT GATATATCAC AAGAGGCACA AGAGGTTCAA	60
GAAGCTAAAA AAGAAGCACA AGCAGAGAAA GACAGTGACA CATTAACTAA AGATGCAAGT	120
GCAGCAAAGG TAGAAGTATC AAAACCAGAG TCACAAGCTG AAAGATTAGC AAACGCTGCA	180
AAACAGAAGC AAGCTAAATT AACACCAGGT TCAAAAGAGA GTCAATTAAC TGAAGCGTTA	240
TTTGCAGAAA AACCAGTTGC TAAAAATGAC TTGAAAGAAA TTCCTCAATT AGTTACTAAA	300
AAGAATGATG TATCAGAGAC AGAGACGGTT AATATAGATA ATAAAGACAC TGTTAAACAA	360
AAAGAAGCTA AATTTGAAAA TGGTGTTATT ACACGTAAAG CTGATGAAAA AACAACTAAT	420
AATACAGCTG TTGACAAGAA ATCAGGTAAA CAATCTAAAA AAACAACACC TTCAAATAAA	480
CGAAATGCAT CAAAAGCATC TACAAATAAA ACTTCAGGTC AGAAAAAGCA ACATAATAAG	540
AgATCATCAC mAGGTGcAAA GaAACAAAGT AGTTcmAGta AGTCaACTCA AAAGAATAAT	600
CAAACTAGTA ATTAAGAnTT CAAAAACAAC AAATGCTAAn TCC	643

(2) INFORMATION FOR SEQ ID NO: 915:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 915:

35
 40
 45

AGGTTTAAAG GAAAAATCCT ACCTGGAGAG ATTCAAAAAC TAAATGGACT ATTTCCAAAA	60
ATAAAGGGCA ATCACATGGT GGTTCCTTATT GGAAGTTAAT AAATAACAAA GGAAAAAGAA	120
TAGCTTCTTT AACTAAAGAA GGAAAAATCT TAAGGGAATA AGGTGTTGTA AATTATGTGC	180
TTTGATATAA ATAATTTGGA TATAAAAAAG TTGAATTTTA GAAAGGTAAA AAACGCAATT	240
CATTTAAGTT TAAGTGGCGA TAAATACCAG TTTTGGAGG ACGGTAAAAC TATCAATAAT	300
ACTTATTTTT TAGCTGTATA TGATAATGCA ATAAACATTT TTGAAGATTT GTTTcCGCCT	360
tCTGATTCTA TAAATTTGGn TCATGTAGTA TATGTTTATA	400

(2) INFORMATION FOR SEQ ID NO: 916:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 120
 180
 240
 300
 360
 400

CCAAAGTTC GATAGTGTAT TAGTATTATC TTAATAAAAT GTTAGGTACA ATAAAGATGA
 TTATATATCG GAGGTTAGTA TAAAAATGTA TGTAGATCGA AAACCATCAC TATATTTAGA
 GGATTTGCCA CATGATTTTA AAAATAGTTT AAGTAAATTT GAAAATGGTG ATGAAGCATT
 TGATACGTTA TTAGGTTTCG TAGAGTTAGA TCATATTTAT TCGTCAGCAC TAAAGGAAAT
 AAGCACTAAA CTGAGTATTT TAGATGACAA TTTCAATCAC ATTTATAWAC ACAATCCTAT
 ACATCATATG GAGCGACGTG TGAAAGAAAT GCSTAGTTTA ATAGAAAAGC TTAATCGTAA
 AGGATTACAG ATTAGCGCAG AACTGCCAA AGAACATATA

(2) INFORMATION FOR SEQ ID NO: 917:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 917:

25
 30
 35
 40
 45
 50
 55
 60
 120
 180
 240
 300
 360
 400

GAAAATTACT ATGAAGATTG CATTAGGATG CGACCATATT GTTACAGATA CAAAAATGCG
 TGTATCTGAA TTTTAAAAT CAAAAGGACA TGATGTCATT GACGTAGGAA CATACGATTT
 CACAAGAACA CATTATCCAA TTTTGGTAA AAAAGTTGGC GAACAAGTTG TTAGCGGTAA
 TGCAGACTTA GGTGTTTGTA TTTGTGGAAC AGGTGTTGGT ATTAACAATG CTGTAAATAA
 AGTACCGGCG TTCGTTTCAGC ACTAGTACGT GATATGACAT CAGCGTTATA CGTAAAAGAG
 GAATTAAATG CGAACGTTAT TGGCTTCGGT GGACGTATTA TAGGTGAGTT ATTAATGTGC
 GATATTATCG ATGCCATTAT TAATGCTGGA TTATAAACCC

(2) INFORMATION FOR SEQ ID NO: 918:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1220 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 918:

50
 55
 60
 120
 180

GTCTCCAGCC ATTWCACCAC ACATACCTGT CCATTACCT TCITTATGTG ACGCTTCAAT
 AACTTGTTTA ACTAAACGTA AGATTGAAGG GTTATATGGT TGGTATAGAT ATGATACAG
 CTCTGACATA CGGTCAGCAG CTAATGTGTA TTGAATTAAA TCATTTGTAC CGATACTGAA

5 GATTCCTAAT TCTATATCAT CCGAAATGTC ATGACCTTCA TTTTAAAGGT TTTCTTTTTC 300
 TTCTAATAAT ATAGCTTTAG CTTCTCTAAA TTCGTTAATT GTTGCAACCA TTGGGAACAT 360
 10 GATATTTAAC TTACCATAAA CTGATGCACG TAATAATGCA CGTAGCTGTG GTCTGAAAAT 420
 ATCTTGTTGC GCAAGGCATA AACGAATCGC ACGGTAACCT AAGAATGGAT TCATTTCTTC 480
 AGGCAAGTTT AAGTATGATA ATTCTTTATC TCCACCTATA TCTAAAGTAC GTACAACAAC 540
 ACGTTTACCG CCCATTGCTT CTAATACTTC TTTATAAGCT TCAAATTGTT CTTCTTCTGT 600
 AGGCATTGGG TCACGACCCA TATATAAAAA CTCAGTTCTA TATAAGCCGA TACCTTGTGC 660
 15 ACCATTTTCa ATAACACCTG GCAAATCATT AGGTGTACCA ATATTTGCAG CAAGCTCTGC 720
 GTGAACACCA TCAACTGTAA CAGTATCAGC ATCACGTAGW TTTTGTAATT CTTTCTTGTC 780
 AGCAAAATAA CGCTCACGTT TATCTTGATA AGCGATTAAC TCATCTTCAG TTGGATTAAC 840
 20 GATTACATCA CCATTTAATC CATCTACGAT AATCATGTG CTTGTTTAA CTTCTTGAGT 900
 AATTGATTTT GTACCAACAA TTGCTGGaAT TTCTAAAGAA CGACTCATAA TTGCAGAGTG 960
 aCTTGTTCTT CCGCCAATGT TTGTAGCAA ACCTTGACG AATTCTTTAT TTAATTGAGC 1020
 25 AGTATCAGAT GCGGTTAAGT CATTCCCTAC AATAACAACG CTTTCATCAA TCATACTCGG 1080
 ATTCGGTAAT TCTACACCTA AAATATGTGA TAACACACGT TTAGAAACGT CGCGAATATC 1140
 30 AGCCGCACGT TCTTTCATGT ATTCGTTATC CATAGATTCA AAAATTGTAA CAAATTGTGT 1200
 TGThACATCC GCTTAATGTG 1220

(2) INFORMATION FOR SEQ ID NO: 919:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 315 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 919:

45 TTAGATTGAC TGCAAATGCC GACTCAAAAG AACAAAGCACA ATCATTGATT CAACCTGTTA 60
 AACAAAGAAAT TCTTGATCGT ATTGGAGAAT ATTATTATGG TTCAGATGAC ACATTAATTG 120
 AGCAAGCTGT AATAAAGAAA ATTCATGAAC CTTTGTGAAT ATATGATGGT ATTACTAATG 180
 50 GTGCTTTATA TCATCGATTG AAAGAAGTGG ATTTAAACGA TGTTCTAAAG GGTATGATTA 240
 ATCACAATGA AAACTTTGTt GATATTAATa AACCTATTGA GCaGCAATTA AAAGATGCAG 300
 TGCAATTTGT TAATA 315

55

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 433 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 920:

ACTTTGTAAA	CCTACAAATG	TAATGAATAA	TCCTATACCT	GCTGAAACAG	CCATCTTCAT	60
TTGATAAGGA	ATTGCATTAA	TAATAACTTC	CCTAAACCCT	GTGACGGTTA	ATATCGCAAA	120
GAATATACCT	GAGAATAAAA	CGCCTGTTAA	ACCAACTTGC	CAAGGAATAC	CCATGGTTAA	180
CACAACAGTA	AATGCAAAGA	ATGCATTTAA	TCCCATACCT	GGTGcTAACG	CAATTGGATA	240
TTTAGCTATT	AGCCCCATGA	ATAGCGAGCC	TACAAATGCT	GCTAATGCAG	TCGCTACAAA	300
AATGGCACCT	TGGTCCATTT	TCATATCTTC	TGATACGCCT	TTAACACCTG	CTAAACTTAA	360
AACTTGCGGG	TTAACTGCTA	AAATATAGGC	CATAGATAAG	AAAGTTGTGA	TACCGCCTAA	420
GATTTCTCTT	TTA					433

(2) INFORMATION FOR SEQ ID NO: 921:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 439 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 921:

CTATTTTATT	TGTATAGCGC	TTGATGTAAA	GTTATTTTAG	CTGTCATACT	ATGTGCTAAA	60
CCAAATTCAG	TTGTGTCATA	TCTTCGCTTT	TTTGTCACTG	TTATAATAGG	TATAATCGGA	120
TATAATGAAA	AGGkGAGGGA	GGATTCAAAA	TGCGTAGGTT	ATTATATTCA	TTTCTTTTTT	180
ATATGATCAT	AGGTTTATTT	AGTGGCTTTT	TCTATAGAGA	GCTAACAAAA	GCATATGATT	240
TTACTGGTAC	AACACAATTA	TCACTTGTAC	ACACACATAC	ACTTATTTTA	GGTATGTTTA	300
TGkktTTAAT	ATTATTACCA	TTGGaGAmGT	TATTTAAATT	ATCAAGTTAC	TACTTATTTA	360
ACTGGtCTTC	TATGTATATA	ACATAGGTGT	TATCGTTACT	ATAGGTATGA	TGGTGACAAA	420
AGGATTCTTC	CAAGTCACT					439

(2) INFORMATION FOR SEQ ID NO: 922:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 474 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 922:

	GTAATATCTA TATACGTATA TTAAATGAT ATATCTGGGT ATTTTCTTTT TAATAGCGGC	60
10	TGTAGCCAGT CATAAATATC TtTCGaTGTT GCGCATTAA CGCAACTTGC ACATATAACA	120
	TCTGCCCCAT AAACCTACCAC ACTCACGTGC TCCATATATT TATCCCCCAT TGGTTTGATA	180
	GATTTTATT ACACATCTA TTATAATATA sTmATAAAGA TTATCAAATT CATCTCTCGA	240
15	AAGGAGACTT GCCTGATGCC TACTGAAGAT ACAACGATGT TTGATCAAGT AGCAGAAGTT	300
	ATTGAACGTC TTCGTCCATT TTTATTACGT GATGGTGGCG ACTGmTCATT GATTGACGTG	360
	GAAGACGGTA TTGTAAATTA CAATTACATG GTGCATGGTG GnACATGCCC CAAGTTCTAC	420
20	AATCACTCTT AAAAGCTGGG TATTGAGCCG TGCATTACAC GAAGAAGTGC CTGG	474

(2) INFORMATION FOR SEQ ID NO: 923:

25	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 554 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 923:

	TCAACTGTG AAGAAAATTT ATTAACCTTA GTAACCTTC CAACAAAACC TTGTGATGTC	60
35	ATCAGAGCCA TATTGAAGT TATACCTGAT TTAGATCCCT TATCAATTAC AATTGTATTG	120
	ATCCACTGAT CCGGATTCTT TGCCAAAACC GTAGTAGAAA TAGGATCAA TTTTGAAATA	180
	TCTTTTAAAT CAAGCTCTTT TTTTAATTTT TCATTTTCCG CTTCTAATTG TTGTTCTTA	240
40	GATTCTAACT GGCTAATCTT ATTTTTAGAT TCTTTAGAAT CTCCTTTTTT AAAAAAGTCC	300
	CCAATCGTAC CAGCAACAAA ATTAACCTGA TAACTCACA CTCGTTGTCC AAAAGACACA	360
45	GAATCACCTA TATATTGTTT AGGAGGTGAT TGAGATTGTG AACGTATGGA CAGCCCAATT	420
	AATGCAATAA AAACGATAAT TGCACATAAA ACAACAATTA ATTTGGTATT TTTAAAAAAC	480
	TTAAGCACCC AGAACACCTC TATTATGTCA nAATATTGTA TATCCTTTTC TAATTnATAT	540
50	TACTCCCAT ATGA	554

(2) INFORMATION FOR SEQ ID NO: 924:

55	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 575 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 924:

	ATCTTTAGGC ATTAAGCAAG TTTATGTAGA AGACTTTGAA CATAAATCCT TTAGCAAAGC	60
10	TAAAAAAGCC TTAGAAGAAA AAGGGTTTAA AGTTGAAAGT AAGGAAGAGT ATAGTGACnA	120
	TATTGATGAG GGTGATGTGA TTTCTCAATC TCCTAAAGGA AAATCAGTAG ATGAGGGGTC	180
	AACGATTTCA TTTGTTGTTT CTAAAGGTAA AAAAAGTGAC TCATCAGATG TCAAAACGAC	240
15	AACTGAATCG GTAGATGTAC CATACTACTGG TAAAAATGAT AAGTCACAAA AAGTTAAAGT	300
	TTATATTAAA GATAAAGATA ATGACGGTTC AACTGAAAAA GGTAAGTTTCG ATATTACTAG	360
	TGATCAACGT ATAGACATTC CTTTAAGAAT TGAAAAAGGA AAAACAGCAA GTTATATTGT	420
20	TAAAGTTGAC GGTAAACTG TAGCTGAAAA AGAAGTCAGC TATGATGATG TATAAATATA	480
	ATTGAAGTAA ATGTACCGAG GTTCTATTTT GGAAGTCTCG GTATTTTAT GTTGGAGATT	540
25	GCGGTAGTTT TAAATGCnT CnTGTCnTCA TATAC	575

(2) INFORMATION FOR SEQ ID NO: 925:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 307 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 925:

	AAAAACGCTG CTGAGTTTAA AAAGCAACAA TTAAATGAAC AAGGTATTTT CAAGAAACCA	60
	GTGATTACAC CTATTAAACC ATATAAAAAT TTCTATCCAG CTGAAGACTA CCATCAAGAT	120
40	TATnACAAAA AGAACCCCGG TACATTATTA CCAATATCAC CGTGGGTCAG GTAGAAAAnc	180
	GTTTATAGAT CnCATTGGGG GGATCCAAAG CTTAAAAAGA TAAAAGTGAC TACCAGATAT	240
45	AGATATATTG TTACACAAGA AACGGCACTG GACCACCATT TATGAATGAT ATTGGACCAT	300
	TTGCTAA	307

(2) INFORMATION FOR SEQ ID NO: 926:

50

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 926:

5 CTCTACGCAC TCATTGAGAT ACACACATTG CGCATTGCTC ATAGAATCAG GTGTACACAT 60
 TAAAGAAATA CAAGAACGAT TACGGCATAA AGATATCAAT ACCACTATGA ATATCTATGC 120
 TAAAATCACA AATTCATACA AAAAAGACGC CTCCCATAG TTTAGTCACC AAATGGAAGA 180
 10 CGTCTCGTAA TTAAATATA TTTGCGTGCA TTCTAATTTA TACTTAGAAT GAATCATACT 240
 CGTGCATAAT GTAATTTTCT AGTTAGTCAA AACTATAAAC AGTTTTACAT CATTCTGGC 300
 ATGCCACCCA TGTTAGGTTG GTCATTATTT TTTTCTGGAA TTGATGCTAC AACCGCTTCA 360
 15 GTCGTTAAGG ACATTGCTGC AACACTTGCA GCAGTTGnAA 400

(2) INFORMATION FOR SEQ ID NO: 927:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 809 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 927:

25 TTTAGGTGGT TTTAATnGTT ATGGTCATTC TGGACACCTA CTTATAATAA AATTTCAAAT 60
 30 CAAACTGAAC nTTTTGTACC GTTTGATTAA ACTCAAATTT ATTAGCTCCG TATTTAAATT 120
 TTGGTTGGGC TATATTCGTT TCGGTACTTA TTTCAACACC GTTTTTATAA ACTCGGAAGC 180
 TATCATAAAC AATTcTGTCT CCAGCTTTTA GTTTGATCCC TTCGATTTTC ATTATTTTCA 240
 35 CATGCGTTAA ATTCCATACA AACGATTCTG TATCTTCGCC TAAAATAATT GTTATCTTTT 300
 TATACATGTT GAATTGGTCG TTAGGAGCAC TACCATGATA GTAAACTGTA CCTTTGCTCA 360
 AATTTTCAAA TGTATACTTT CTTTGTCTC CGCCTGCATG CCAATCAATA TTAAATCAA 420
 40 ACGACCACAA TCCAACCTTT TTGTTTTCTT CTAACCTAG GCTTGTTCCTA ATACTTTCAC 480
 CGTATGGTAA TTCTGTAGTT TCGAATTTTA GTTCAAAAGA AACTTTATTA CCTTTTGT 540
 TAGGGTTTAT AACTCCGTTA AAAATAACTT TATACTGTTT ACCATTTACA TAAATTTGTT 600
 45 GATCGTGTCT TGAATATTCG TAATCCGGGA AGTTGTTTTT ATCTAATTC ACGTAATCAT 660
 CAGAAGTTGG TTGAGTAAAC CTGTAATTCA ACTCTTCTTT TCTTCTGATn TCTCGCAAAT 720
 50 ACATAGGTTT TATGTCTGTC GTTAACGAAT ACAACATATC TCGCATATAA GCAATGTCTG 780
 AACGATTTTT AACTTTACAA AAACAAGGA 809

(2) INFORMATION FOR SEQ ID NO: 928:

(A) LENGTH: 1016 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 928:

10	nAAAACTATT ATCGCATCTA CaGTaATCGC TGCAGGTCTT TTAActCAAA CTAATGATGC	60
	TAAAGCTTTC TTTAGTTATG AATGGAAAGG TTTAGAAATC GCAAAAAATT TAGCAGATCA	120
	AGCTAAAAAG GACGATGAAC GTATTGATAA GTTAATGAAA GAGTCTGATA AAAATCTAAC	180
15	TCCTTACAAA GCTGAACTG TTAATGATCT GTACCTTATT GTTAAAAAT TAAGcCAAGG	240
	TGATGTGAAG AAAGCAGTTG TCAGAATTAA AGATGGTGGT CCTAGAGATT ACTATACTTT	300
	TGACTTAACT CGTCCTTTAG AAGAGAACAG AAAAAATATT AAAGTTGTTA AAAACGGTGA	360
20	AATCGACTCG ATTACTTGGT ATTAAAAAAC ATACTGAATT AAATAGTTGT ACGCCAAACG	420
	TTAGAAAACA ATGCTAACGT ATTGGCGTGC TTTTTTATTT TAAGTAACTT CCAATTTATT	480
	TAGCATCTTT ACGACTGTTT AATAAAGCAC GTATGATTAA CACGGTTGCC ACTATATCCG	540
25	TTACAATTTT TATGATTGTT AATACATTCG ATTTCTTTT CACAATAGCC ACCTCGCTTG	600
	TTCAAAACAT AAACAACAT TGCATTCAT TTTAAGTAAT TATTTATATT ATTTATCCCA	660
	ATAAGCTCAC CATTCAAATA AACACAATAT TATAAAATAA TTATACCTTT GGATATAGCA	720
30	AAAAGCCACA CTCTATAGCA TGGCTTCTAT CAATTATTTT AAAGTATTAT ATTTTAAAC	780
	TAGATCGATT TGTCTTTGTA ATTTTTTCTT TTCATAACTG TGTGGAAT GAATTAAATT	840
35	AACAGCTCTT TGTGCTTTAC GGTGTGTTGC AACAGTTCTT GTACGTTGA AAAAGTTTAC	900
	AGCTTTTTGT GCATCCACAA CTTTTtATT TACTTGWTTT KTAAAGTTTG TTGATACTGA	960
	TCATATTTTT TAGCAGCTTC ACCGTTtTTA GTTGTGCATG AGATCACC GG CAACCG	1016

40

(2) INFORMATION FOR SEQ ID NO: 929:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 929:

50

	AAATGAAATC ATACCTAAAA GAGATATTAT TACAGAATCA ATGATTGTG ACTGTATTCA	60
	AAATGCAGGT ATTGATTAG AAGTATTAA AGACGACTTA CAAAAAGTA AACTAACCGA	120

55

CGTTTTCTTT AGTGAAGATG TTCATGAAGA AGGTTTAAAA GTCGAAGGAT TATACCCATA 240
 TCACATCTAT ACTTATATAA TTAATGAATT GATGGGTAAA CCTATCGAAA AGAATCTTCC 300
 5 TCCTAAATTA GAAACTTATA TACAGCAACA ACAACTTGTA ACGATGGAAG AATTACTTAC 360
 TATTTATGAA TGGCCAGAAA AACTTTTAAA CAAAGAGTTA 400

(2) INFORMATION FOR SEQ ID NO: 930:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 930:

20 TATTTCTGAC ATGTATCATG CTCCTTCGTG CTTTATTCTA ATGTAATATA CACAATTATA 60
 CATTGAAATT CAAATTTGTG AACACATTGT GAACTGACAT AAATTATACA CAATTTTATA 120
 GCGCATAATG TTTTAAAAAT ACACATTATA GCAATTATAA TTAAGTTCAT TGTGTTGTCA 180
 25 TCAAAAAAAG AAAAAGGTGA TGTTTTAAAT GGATACAGTT GAATCAGTCG GTTTTTGACA 240
 GCATGACTTA GCAGTTCATA TCATTTTGCA ACGATGGTGT GGTATGCTTA ATGTCGCAAT 300
 TCAGATTTTA GGATCGCAA ATGACTCATA TATAGCATGC CAAGTGGCAA GTATCATACG 360
 30 ACATGGnGTG TCAGGCATAn GGCTCATACT GTGCAATATG 400

(2) INFORMATION FOR SEQ ID NO: 931:

- (i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGTH: 543 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 931:

CCAATCAATG TTGCCAATGG ATCACCTAAT GACAACCAAG TAATTGCTAC CATAGTAACT 60
 45 GCTGTTTTAA TTTTACCTAA TTGACCAGCT GCACTTACGA ATCCTTGTTT AATTTGTAGT 120
 AAACGTAAAC CAGTTACGGC AAATTcTCTG GCAATaATAA TGATTGCTAC TACAGAATTT 180
 GTTAGTCCTA GTTGCAACAAG TACAATTAAA GCACTTGCAA CTAATAATTT ATCCGCTAAT 240
 50 GGATCCAAAA ATTTCCCAT ATTTGTA ACT AAATTCCATT TTCTAGCTAA ATAACCATCA 300
 ACAAATCGC TAAGGGAAGC CAATATAAAA ATAAACCAC TGATTAATAA CTCAATTCTT 360

ATAAAACTG GTATTAACAC TACTCTAAAA ACCGTAATCT GGTTCGGAAT ATTCATTATA 480
 CATCCTCATT TCTCACTAAT TTATTCTGT TAAAAATATT AAAACTAACC ATGATCCATA 540
 5 ACC 543

(2) INFORMATION FOR SEQ ID NO: 932:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 932:

GATTGCGAAT GAGATTAGTG GGATGATACC GGTAGAATGG GAGCAAGTAT TTACAATAGC 60
 20 TTATGTAACT GATCAAGCTG GAGAAGTCAT CTTTAATTAT ACTAAACCAG ATAGTGATGA 120
 ATTAAATTAT TATTCAGmCA TACCTAAAGA TTGCAATGTC TCAAAAGATA TTTTAAAGAA 180
 TTCATGGTTT AAAGTTTATC GAATGTTTGA TGAGTTAAGA GAACTTTTA AAGAAGAAGG 240
 25 GCTTGAACCA TGGACATCAT GCGAATTTGA CTTTACAAGA GATGGCAAAT TGAATGTATC 300
 TTTTGATTAT ATAGATTGGA TAAATACAGA GTTTGATCAA TTGGGCCGTC AAAATTATTA 360
 TATGTACAAA AAATTTGGGG TTATACCAGA AATGGAATAT 400

(2) INFORMATION FOR SEQ ID NO: 933:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 759 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 933:

ATTTTACAA AGAAATAGTA GTCTTATATA TCTTAACATT TAATAACTAA ATCAAACATA 60
 TTTTGAGCCT TTTTGTGAAA ATTAACATTT TAACCTTTTT GATTTTACAA CAAATTATAG 120
 45 CTACGTATTG AAAATTAAAG CmTTGGtTTA AGTGTGTGT TAAAgyTTT TATGTTTAGA 180
 TTGTTATATT ACTATGTATT CTTAAATTTG TTTAAAAAAT TAATGCTTTA AATTGATGTA 240
 TAATGGGAAT AAGAAATAAA TAAAAACGAC CCGCAGGATT AACGTACGGG TCCACTACTA 300
 50 AAGGGAGTCA AATTTTACC TcGTTTGTAT CATGCAGCGT TTTACACATA CTTTAAAGAG 360
 ATGTTTATTC GTTATCGAAG GTACACCTTT ATTATAACTT ATATCATTTT TATTAAAATA 420

ATAATGATTA TTAAATAGTA ACTAAATACA AAATTACATG GGGTGAATGA TAATGAAACA 540
 ATACTTAATT ACTGGTGGGA CTGGTATGGT TGGATCTCAA TTAGTTAATG AAATTAAAAA 600
 5 ATCAGATTCA CATATCACGA TATTAACGCG ACACGACCAA ATTTCAAATG ATAAGAAAAT 660
 TTCATATGTC AACTGGGCTA AATCTGGGTG GGAACACAAA GTTCCTCAAA ATATCGATGT 720
 GGTCATCAAC TTAGCAGGTG CTACATTGAA TAAACGATG 759

(2) INFORMATION FOR SEQ ID NO: 934:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 934:

GATCAAGCTG GGGAAGTCAT TTTTAATTAT ACTAAACCAG GTAGTGATGA ATTAAATTAT 60
 TATTCAGACA TACCTAAAGA TTGCAATGTC TCAAAAGATA TTTTAAAGAA TTCATGGTTT 120
 25 AAAGTTTATC GAATGTTTGA TGAGTTAAGA GAAACTTTTA AAGAAGAAGA GCTTGAACCA 180
 TGGACATCAT GCGAATTTGA CTTTACAAGA GATGGCAAAT TGAATGTATC TTTTGATTAT 240
 ATTGATTGGG TGAATTCAGA ATTTGGACCA ATGGGAAGAG ACCATTATTA TATGTATAAA 300
 30 AAATTTGGnA TTGGCCTGA AAAAGAATAT GCCATAAATT GGGTTGnAAA AATAAGnTT 360
 ATGTTAAGAG CAAGnTTGAG CTGAACTATA GGGGAGATAA 400

(2) INFORMATION FOR SEQ ID NO: 935:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 935:

CTGAATATAA TTTTnCAAC TACATCTCGT TTATTAGACA CCGTGCACTG ACTAAGAAAA 60
 TTTCTCTTGA CTCCATTAGT CCTGACGAAT ACTAACATTT AACTACCTTG CTCATCGATA 120
 GAAACAACTT GTAATGTTAA TTCCCTTAT TTTCTTAGTT TTAATCTATC AGCGATTAAAT 180
 50 TCGATTGCAT CTTTTTCATA AGCAATTGGA TAACTTGAC CGCGGTACAC CTAACGCTCG 240
 AAATATGATT TnTATCGTC ATAATCTAAA ATATTATnGG CAAAATCACA GCAGTTTTCA 300

(2) INFORMATION FOR SEQ ID NO: 936:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 619 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 936:

GTTATCATCA TCTTCTGAAG TATCATCTTG CCCATCGACT TGACTTGCAT CAGTTCTCTT 60
 TAAATCATCA AACTCTGGAC TAGCTAATGT ATTATATAGT GTTTTAGGTA AGTAATAATA 120
 TTCGTTTAGC ATATCTTCTT TTAAATATC ATCTACTTGA TACATAATTC TTTGAGCCAG 180
 ATACATATTA TCTTCAGAAT TTTGATTTTG GAAAAAATTA GGTAAATATA AGAACATATT 240
 TACTAAAATA TCATCTAAAT CTGAATGTAT AGATGGTATA TCAAAGAAAT CTTGGCTTAA 300
 AAATGCATGT TCTAAATACA AAAATAATAA ATCTGTATAT TGTGTTTTAG TACGATACAC 360
 TTTAATTGGA GATTCCGTAT ATGATATACG TGTATCTAAG CGAAGATCAA TTAATTTAGC 420
 AGTACTTGGG CGCTCAACTT TAATAGAATT TAATACGCGC ATATCTTCTA ATAATTTAAA 480
 AAGTTGTTGG ATAAAATTTA GGGtGTTTAA AAGTTTTATC TTGtACTACT TCATTTACAA 540
 TTTGtACATC CaTCATATGA TAACCGTAAG CAGCTAACAT AACATCTGTT TTTAAACCAG 600
 CCATTTTCGAT ATGGCTTGG 619

(2) INFORMATION FOR SEQ ID NO: 937:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 615 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 937:

TAATCCCTTT CATAATGGGC ATCAATATCA TATTAATCAA TCTAAAAAAC TTACAAATGC 60
 TGACGTTACT ATTGCAATAA TGAGTGGTAA CTTTGTCTAG CGTGGCGAAC CAGCAATCTA 120
 TAATAAGTTT ACTCGTGCAA AAATGGCATT ATCAACAGCT GATTTAGTTA TCGAACTACC 180
 AGCAACTGCC AGTTTATCAT CTGGcGATCA TTTTGCCGAA CTAGCAGTTA AAGTCGCnGA 240
 TTATATGAGT GTCGATACAA TTGCATTGGG TAGTGAAAAAT AATGATATCA AAACATTAAA 300
 rCmATTAgCA CACAGCATTA ATGaAATTGA mCAATCTGaA TCCTTTTCAC aAAAAGTrAA 360

AAGTCCTAAC AACATACTTG GTATTAGTTA CCTGAAAGCA ATTGCTAAAA ATGCTAAAAA 480
 CATCAATGCA ATTTCTATCA AACGAGAAAA TGCTCAACAT CATGATTCAT TAATTCAACA 540
 5 CCATCAGTTT GCAAGTGGTA CATCTATTAG AACATCAATC ATTAGTCAAG ATGATCATTG 600
 GCATCATGTG GTACC 615

(2) INFORMATION FOR SEQ ID NO: 938:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 938:

20 AGAATTGAAG ATGGTAAAGT TTCAAAATAT CATTCCGGTTA TCATAAAAGA CGCACAAGCA 60
 ACTTCACCAT ATTCAATTTT TATCAGAGGT GCTATTTATC GCTTTGAACC ATTAGTATAA 120
 ATATACGTAA GTGCTATGAG CGAGAATGCC CATATGAATA ATGACAAGCA CAATGGAAAG 180
 25 AATCGTAATA TATTATTTAA TCGTGATGCT TAATTAAAT GAAAAAGATT GATAATATAA 240
 ATGTGAAAAA GTAAGTATAC CCGTAAACTA AAGTATTCAC GGTGAGAGGT GCTCATGTCA 300
 TAATGATGCA CGTGTCAATA TTATAATAGC TTAATATGT CCATACAACA CCATGTAGAT 360
 30 ATGCATATAC AGGGTATGAT AGACATAGnG TCTGnGAAAT 400

(2) INFORMATION FOR SEQ ID NO: 939:

- (i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGTH: 466 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 939:

GTACAGGTAT GATTGGTGAT CCCATCAGGT AAATCAGAAG AACGTGTGCT ACAAACAGAA 60
 45 GAACAAGTAG ATAAAAATAT CGAAGGTATT AGTAAGCAAA TGCACAATAT TTTTGAATTT 120
 GGAACAGACC ATGGTGCACT GCTTGTTAAT AATAGAGACT GGTTAGGACA AATCTCATT 180
 ATTAGTTTTT TACGTGACTA TGGTAAACAC GTCGGCGTTA ATTACATGTT AGGTAAAGAT 240
 50 TCAATCCAAA GTCGTTTAGA ACATGGTATT TCATATACAG AATTCACATA CACGATTTTA 300
 CAAGCTATTG ATCTCGGTCA TTTGAATAGA GAATTGAATT GTAAGATTCA AGTAGGTGGA 360

GTAGATAAGC AAATGATTAA TAGTTGAGGG GTATGTCGAT GAAGCG

466

(2) INFORMATION FOR SEQ ID NO: 940:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 329 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 940:

15 ATGAATCACC CATTGATACC AATCATTTAC ATCACTGTCA TACGAACATA TATTTAAATA 60
 GAAAAAATT ATTTTAAAGA TTATAACTAC TCTTAATCAT TTTAGTGAAT TAAAAAAGT 120
 AGTGCAAAAA GCAAAATATA CTTTATACAC TACAAATCAT TTATTTATAA TAAAGTTTCA 180
 20 CCAAAAAATG TTCCAATAA TGAAACCGCT TGTTCAGCAG TATGATTATn ACTGTCAATC 240
 AATGGATTTA CTTCAACTAA ATCCATTGAG GAAATTAAAT GTGATTGATG CAGTAATTCC 300
 AATGChAAAT GGCTTTCTCT ATnACTAAG 329

25 (2) INFORMATION FOR SEQ ID NO: 941:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 315 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 941:

35 CGATATAGTA CCAATTTATT TGAAGAAAAA TCCGAACTTC AAATTTCCGA TGAAGCAAGA 60
 TATACCGGTT ATTATGATTG GACCAGGTAC TGGAATTGCT CCTTTTAGAG CATATTTACA 120
 40 AGAACGTGAA GAACTTGGTA TGA CTGGAAA AACATGGTTG TTCTTTGGTG ATCAACACCG 180
 TAGTTCTGAC TTTTATATG AAGAAGAAAT AGAAGAAATG CnTGAAAATG GAAACTAACA 240
 CGCGTAGATT AGCATTTC CA AGAGnCCAGA CACAAGATAT GACAGCCnCG ATATGGAGAA 300
 45 GTAACGTTCC ATGAT 315

(2) INFORMATION FOR SEQ ID NO: 942:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 699 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 942:

5 TACCAATAAT CAACTTGTTT TACATCATCG TGTGTTTCAA TCACTTTCCT TTTTATACCT 60
 TTCTTACCTT TAAAGAATGC ATCAACATCT TTTATTGAAG ATCCTTTTGA ATCTTTAGTA 120
 AATCTATAAT CTGCAAGTTC CACTTTACCC ATTAATTCTT TGTAATCTTT CCATCCTTGA 180
 ATCCCTTTAC TATTCGTAAA CGACTTTTCT GTCATTAATG GATCCAACCT ATCAAGAGAT 240
 10 GGCGCATCTT TAATTCCCTT CATTTTCTTC TCTATTTCTT CATATTTCTC TAAATCTGGA 300
 TCTTTCTCTT CTTCGGATTC TTTTTTTACA GTTCTTGTT TTGATTCTTC TGATTTCTTT 360
 GATGATAATT TATGATTCCC ACCAAAAGAA CAACCTGCTA CGACTAAAAG CATAATTAAA 420
 15 AGTAAACATC CCAGTGTCTT TTCATAAAC TTCTCTCTCT TAAGTTATTT TGTTTGAATC 480
 AAATAAACAT ACATAAATTT AAAAACTTA TTTATATTGa TATTTTACAT TAACCATCAA 540
 20 TATTTTAA TACTTTAAWT GaTAATTTAA GAAATTGTT TTGTCTAATA ATTTAAGTAG 600
 TTA AACATA GATAGATATA AAATAGTTTA TAGCCATTAA TTTATAAGTT TAAAATTCGC 660
 ACTATTTCAA TTGCCATTT AAAAAAGTT TGTTTAATC 699

25 (2) INFORMATION FOR SEQ ID NO: 943:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 30 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 943:

35 CGCTCCATAA ATCAACAATG TTCGAGTCTG TAATCTAGAT GATAAAACGC CAAGAATGAT 60
 CATAAATGGC ACTTCTAAAc CAGCACATAA ACTAGCTAAA TAACCGACAT GTTGTTCATT 120
 40 TTCTTTTAAA TAATCAGTAA CAAATAAAGG CATATTCATC GTATACATCC ATTGTCCAAT 180
 GTGTAATAAA ATAAATGCAA TAAATGGTAA TAAAAGCGTT TTGTCTTTAA ACATATTAGG 240
 AGCAATTTTT TCAACATGTT GTTGCCTACT AATAGGGTGT TTAATGTTTA AATCCTTATA 300
 45 GAnAAACACT TGAAGTACTA AAGTAAATAA AATGATACTT ATTGGGCCAC CAAACAATCC 360
 AGCATAGCCT TTAAATCCGA TTAATTGGGC ACCAATAAAT 400

(2) INFORMATION FOR SEQ ID NO: 944:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 570 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 944:

5 AAGATGACAA TACTTTGTTT CAAGAATTGA AAAAGAATT AGAACAATGG GATTTTAATG 60
 TTGCTGGTAT TGAAGATTC GGCAAAGTAA TGGATACATT TGAAAGTTTT AATCCTGAAA 120
 TTGTTATATT GGATGTTCAA TTACCTAmAT ATGATGGGTT TTATTGGTGC AGAAAAATGA 180
 10 GAGAAGTTTC CAACGTACCA ATATTATTTT TATCATCTCG TGATAATCCA ATGGATCAAG 240
 TGATGAGTAT GGAACCTGGC GCAGATGATT ATATGCAAAA ACCGTTCTAT ACCAATGTAT 300
 TAATTGCTAA ATTACAAGCG ATTTATCGTC GTGTCTATGA GTTTACAGCT GAAGAAAAAC 360
 15 GTACATTGAC TTGGCAAGAT GCTGTCGTTG ATCTATCAAA AGATAGTATA CAAAAAGGTG 420
 ATCAGACGAT TTTCTGTCC AAAACAGAAA TGATTATATT AGAAATTCTT ATTACCAAAA 480
 20 AAAATCAAAT CGTTTCGAGA GATACAATTA TCACTGCATT ATGGGATGAT GAAGCATTTG 540
 TTAGTGATAA TACGTTAACA GTAAATGTGA 570

(2) INFORMATION FOR SEQ ID NO: 945:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1241 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 945:

35 CCGAGnCCAC CGTTCCAAAG TCCATTTTTA TCCCCATCCC TCTCTGAATT GAACTATAAT 60
 TTTTGAATTT ATTAATAATG CTATTTTtnt TATTTTATCA AAAACTGATT ACAAATACAC 120
 ATAGAAAATG AAATATATTT TTCTTTGACT ATACATAACT GTTATTTCTT TGGATATTTA 180
 40 TATTAAATGC ATTGGTAGTA GTTGTAGCGC ATTAATGTTT GACGGTGTAT AGTATTAATT 240
 ATATTGAAAG TTAGTTGGAA GTTGATAGTA GGAGTGGGAG CATTGAAGCG ATTTGTGGCG 300
 ACGGTATTAT TATTACTAGT CTTTATATCA GGatGtGGTA ATGrTaAATA TGTGAAAGAA 360
 45 ATaGATGAAG CAGTTAAaAt TCAAAATCaA AAACAAGAAC ACTTGCCCAA AAAAGGCAAC 420
 GGTGATCGTG TTGATCATT TGAACGCAAA GATGCTAATA TTTATGTCTA TGATAAGGAT 480
 AAAATTATCA TTTTAGCTTA TAAACCTTTG AGTAATGATG ATGAAGTGCA TTATTATGCA 540
 50 TATGATTTTA GTGATAAACG TGTATCATAT AAGCAAGATT TTGATTGAG ACGATATTAT 600
 CAACAACATG ATGCGGATTA TCATGAAGAA AATATGACGA ACTAGATATG AATAGGAGTT 660

GGTAAAAAAG AATCAGCAAC GACATCTTCG AAAAACGGCA AACCATTAGT TGTCGTATAT 780
 GGCGACTATA AATGTCCTTA TTGTAAAGAA TTAGATGAAA AAGTCATGCC AAAGTTGCGT 840
 5 AAAAATTATA TAGATAATCA CAAAGTGGAA TACCAATTG TCAATTTAGC TTTCTTAGGT 900
 AAAGACTCAA TTGTTGGTTC GCGTGCGAGT CATGCAGTAT TGATGTATGC ACCTAAATCA 960
 TTTTtagatt TTCAAAAGCA ATTATTTGCT GCCCAGCAAG ATGAAAATAA AGAATGGTTA 1020
 10 ACAAAGAAGAC TATTAGATAA ACATATTAAA CAACTGCATT TAGATAAAGA GACGGAAAAT 1080
 AAAATTATAA AAGATTACAA GACAAAAGAT AGCAAGTCTT GGAAAGCTGC AGAGAAAGAT 1140
 15 AAAAAAATAG CGAAAGATAA TCATATAAaA aCGACACCAA CTGCATTTAT TAATGGCGAG 1200
 AAAGTTGAGA TCCATATGAT TATGAAAGTT ATGAGAAGTT A 1241

(2) INFORMATION FOR SEQ ID NO: 946:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1057 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 946:

TAATACAATG ACACCATTTA GCATGACCGT TATCCCTGTA ATTCAGCTGA TATTATCTGT 60
 30 TGCAATTTTA TGTGACGAAC TGTGCACTT AATTTGATAA ATCAACAAAT ACAAAAAATC 120
 TAAGTTGAAC AATTATGATA CAACCGTGCA AACGATATGT AGTATAACTT GTCAACTTAG 180
 ACTTATTGAT AAATATATTA ATAGTGGTTT ACCATAGCAG GAGATTTTAC aTCAAAATTT 240
 35 TGAAGTAGCG TATCaATCTT TGaATCaTCa ATATATACct TATGTAAATT TTTCaTATAC 300
 ATCGAATGAG AAAGTGCTTC aTAATTTAAT GAAAAAGATA TATGATCTCC AACTTGATAG 360
 40 TGTCTTGAC CATTTAAATC AAGCATTAAA TGATCACTCG AAGCGCCTAA AATATTGATA 420
 TGCTGATCCA TAGGTGAAAT ATTATCGACT TTTGTAYCTA AATAACCAAT ATCTACAATA 480
 GCTTGTAAGA ATGATTCATG CGTTTGTGTA TTAACCGAG GTTTAATTTT TAAAATCTCA 540
 45 GCCTCCAATG TAATCGCATC TTGATATAAC ATAGCGATCG CTTGATTTGT CGTTGTATCA 600
 ACACCTCTAA ACAACGTTTC ACCTATTCGC AATTCATTTA TTTTACCCAA ATCATTATAT 660
 AAAAGTTGTG GTAACATGCT CGAATTACCA CCTGAAATAA TTTTCAATCG ATATCCTATT 720
 50 TCTCTTTCAA CAGCTGAGAC GAATCGATTA ATCATAAAGA TATCATCATC ACTTGGCGCA 780
 TCAGATTTAA AACACATAAA ATTGAATGCT AAACCTACAA AATGGATATT TTTCAAGTGA 840

55

CAATCTACCA TTAATAAAAT CTTATGTTTT TTTCTTAAAA CTTCTGgCTA CTTCAATTTAT 960
 TTGATGTATG GGTAGATAAT TCCGGTTTGG GATACCTCCA TATCnAACCT TTTTTCCTAA 1020
 5 TCAATATCCT GGAAAAhCCC TTTTTTTGGG CAGGGCG 1057

(2) INFORMATION FOR SEQ ID NO: 947:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 649 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 947:

GATATCCGCT CGATGAATAT ACCTTTAAAG GTGTTGGTGC AAATTCTGAA CAATTAAGTG 60
 20 CTATGAATCA TGATTCTTTA AAAAAGGAGT ACATTTCAAA TGA CTGTAA ATATAATCAA 120
 AATGGCGAAT TAACAATGGA TGGTATTAGT TTA AAAACGA TTGCACAAAG CTTTGGTACA 180
 CCTACCATTG TTTATGATGA ACTACAAATT aGAGAACAGA TCGCGCGTTA CCATCGCGCA 240
 25 TTTAAAGATA GTGGATTAAA ATACAATATT TCATACGCCT CAAAGGCATT TACTTGCATT 300
 CAAATGGTCA AACTTGTAGC TGAGGAAGAT TTACAGTTAG ATGTTGTTTC TGAAGGTGAA 360
 TTATATACAG CTTTAGAAGC AGGTTTTGAA CCGAGTCGCA TCCATTTCCA TGGTAACAAT 420
 30 AAAACGAAAC ATGAAATTAG GTATGCTTTA GAaATAATA TCGGTTATTT TGTATAGAT 480
 TCATTAGAAG AAATTGAATT rATAGACCGC TATGCTrATG ATACGGTTCA AGTTGTATtA 540
 35 CcGAGTTAAT CCcAGGTGgT TGrAAGCcAC mTrCaCaCgA aTTTATTCaA ACyGGGcaaG 600
 AAgGATAGGT TAAAGTTTTG GGATTAAhCC nhTTTCCCAT TAATGGGCC 649

(2) INFORMATION FOR SEQ ID NO: 948:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 948:

TACAAAGCTT AAAAACAGAG CTTAGTyTCG TTGAACCATT AATTAGCCGC TTAGAATTAG 60
 50 AAGAAGCTAA TGATAAACTA GCTAATATCA ATGATAAGTT AGATGACATG TATGATTTAA 120
 TTGAACATGA AGTTAAAGCT AAAAATGATG TCGAAGAAAC AAAAGATATC aTTACGGATA 180

AAACTACTA TATAAATGAA TCTGATGCTC AGAGTGTTCG TCAATTTGaa AATGrAATTC 300
 aAAGTTTAAT TTCTGTATAT GATGATATTT TAAAAGAAAT GTCTAAATCT GCTGTGCGAT 360
 5 ATAGCGAGGT TCAGGATAAT TTACAATATT TAGAAGATCA TGTACAGTT ATTAATGACC 420
 AACAAGAAAA GCTACAAAAT CATCTGATTC AATTGCGTGA AGATGnGCAG AA 472

(2) INFORMATION FOR SEQ ID NO: 949:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 949:

AGAGGATCCC AGGCGTTTGT TGATCCAGCA GCCGATnCAG AATACATCTT CTATGCATAT 60
 AAACGTTTCA ATTTGGAATT GAAAGATCCA AACTTCACTA GTGAAGAAAA TATGTTTAAA 120
 GATGTATCAG ACAAACCATT AATACCTGCT CGTAAAGCTC AAATTACAAA TGCGAACTAT 180
 25 AAACGACATG GTATGAAGTT GATGATTCTT GGGAAATGACG AACCATATAA CAATGAGTTC 240
 AAGAAAGATG ATCGAAAATG CGAAGAATGA GAAATTTCAA AGGTGCGCAn GTCAAGCCAT 300
 GTGTCAGGGG GGCATATGAT TTCnGGTCAT CAGTAAGGAC TGTGTCATGC CATGGCCATC 360
 30 TGTGTACnG ATCTCTAATA GAATCTCAGA TGCCCCGGCAA 400

(2) INFORMATION FOR SEQ ID NO: 950:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 556 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 950:

AAAATCTTCG ATGACAAAGA AATCTAATGA TGTTTCTTTA AAATGACCAA ACTGCAACGG 60
 45 AATGATATAA CCACACTTGA GATATGGCGC TTCTTTTTTC AACTTAGTCA TCACATCATA 120
 ATCCAAAGAC ATCACACGAT ATTGATGTTT AACACCATGC TTTTCAAAA TATCAATAAC 180
 ACSTTGTTGTA TAATCTGCTG GTTCTTTACC ATGTGGCTTT AACTCTACTA GTAGCTTCAC 240
 50 ATTTGATTGT TTAGCCGTTT CAATAAATTC GTCTAAGGAT ACAAATTTTG CTTCAATGCC 300
 ATTTTGACGC ATTTTCAAAC CGACGATATC TTTGAAATTA GATTCAGAAA TATTTTTATT 360

CATAATTGTA TCTAACTCAA CGTATTTCGAC ATTGCGCTTTT GCAGCAGCTT TCAATGACGG 480
 AATAGAATTT TCAACACCTT TATCTTCGAA ACCACGTGGC CAATAATGGA GATATTGTAA 540
 5 TTGATAGTAT TATTGG 556

(2) INFORMATION FOR SEQ ID NO: 951:

(i) SEQUENCE CHARACTERISTICS:
 10 (A) LENGTH: 469 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 951:

TTGTGGTAAT GATGTTAATC AGTATATTTT TACTTTTGAC AATGACATCT AAAGGATTAA 60
 20 GCAATCTTAG AGTAATAGAT GATGAGGCAA ATATCATTTT TTTTATTACT GAATTGAATT 120
 ATATTAAGTC GCAAGCTATA GCAAATCAAG GATATATCAA TGTTAGATTT TATGAAAACA 180
 GTGACACTAT TAAAGTAATA GAGAATAATA AAATACGATT TCTAAAATTA AAAGTAGGCA 240
 25 AAATAATTAA TGTTGCAAAA GTTGATATTA TTGCCTTTGA TAAAAAAGGG AATATCAATA 300
 AATTTGGTAG CATAACAATT TACAATAACA ATTCAATTTA TAGgaATAAT ATTCCATATT 360
 GaAAAAGGca AGAATTCGTT ATGAAAAGCT ATAAGTGTA AGGTTTCATTC TTAATAGATA 420
 30 GTATGGCTGG nTTTTTGCCA ATTGGGATTG nATTACATTA CTATTGATT 469

(2) INFORMATION FOR SEQ ID NO: 952:

(i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGTH: 730 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 952:

ACATnTATAA TnCTTATTTG CTGCTCCTT TCACTTTTTTA AGAGTAAAn ATAACCTGAA 60
 45 ACATAACTTT AACATCCAAG TTTCAGGTGA GTGTAATGTA AATGATTTAC TTTGTTnTAT 120
 TATGATTAAA CATCTAATTC TCTGCTTGTA AGTGCTTTAA TCATCTCTGG ATCAGCATGT 180
 GAGAAAAATT GGCTATTACT TTCTTCTAAT GTCGCAATTT GTTGTTTATC TTCTTCTGTA 240
 50 AGTTCAAAAT CAAATATATC TAAGTTTTGA GCCATACGTT CTGGATTTAC TGATTTTGCG 300
 AGTACAACAA TGTCACGTTT TACTAACCAG CGTAAATAA CTTGTGCAAT TGATTTGTTG 360

55

AATGGTGCCC AAGCTTCTAC CACAACATTT TCTTGTTGAA GTGCTGCAAC TTGTTCTTCT 480
 TGTGATGGA ATGGATTAAT TTCTATTTGG TTCACTTGTG GTTGAATTG ATTGTGAATA 540
 CCTAAATCTA CAATTCGATC AACGCCGAAG TTAGATACAC CGATTGCTTT AAaTTTTACC 600
 ATTTTCtTTT AACTCCTCCA ATGCACGCCC ATGGAGCCAT ATACnGCCAT TGGTAAAGGT 660
 TGGGTGGAAT nAAGCACCTA GGATCCAGGA TAGGTCCCAA ATTTTAATCC GGTGGGnAAT 720
 GGATCGGTTC 730

(2) INFORMATION FOR SEQ ID NO: 953:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 953:

ACAAATGACA ACGTACAATG ACGTCTTCAC AAACATGCAT TTTAGGAGAT TGATTATTAC 60
 AAATATCCAG TTTTTTGTTA CATCTATTG CAAATCTGCA CTTATTAACG TGAAATGCTT 120
 GGATTGATGG CACAATTCCT TCTATCGTTT CAAGTCGTTT CTGCTCTTCT TTAAGTTTTG 180
 GAATTGTTGA TAATAATTTT TGGGTATAAG GATGTTTCGG ATTGTGTAGT ATTTCTTTTA 240
 TGCTCCAAA TTCTACAATT TGACCTGCAT ACATACTAA GACTTTATCG CAAAACCTCAG 300
 CAACTACACT CAAATCATGT GTAATCATCA TAATTGCCAT TTGCGTTTCC TTTTGCAATT 360
 CTTTTAATAA GTCTAATATT TGCGCTTGAA TCGTGACATC 400

(2) INFORMATION FOR SEQ ID NO: 954:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 954:

TAAATCTTGA GGACGTTTGG TCTGAAAGCG ATGCTCAACT GTTGATAAAT CCAATTCGAT 60
 AACATCTGTA TAATTAGGAT CTTCTTTCTC AACATCAAAa ACATATGGTT TTGTTTCAAA 120
 TATTCTTTTA CTAGCGCGAT ATGTTGCTCT GATCTACCAG TTAAC TTCAT ATATTTAAGA 180
 ATTCTCATCA ACTGGGGAAG AATCCGCAAG TTGCTCCATA,CTCTGGTGCC ATGTTTnCAA 240

(2) INFORMATION FOR SEQ ID NO: 955:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 514 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 955:

```

AAACATTTAC ACAAGTAGTA GAAGCAATTT ATGATGAAGA AGGTAACAGC TTAGamgCGG      60
CGCGCCATCC attACAAATC GTTCAAATTA AAGTGGATCG CCCGATATAT CCAAACAACA      120
TGATGAGAAA GGAAATTGGC TAATGAAAGC TACTACAATC ATTGGCATAG CTGGTGGATC      180
TGGCTCAGGA AAAACAACCTG TAACTAACGA AATTATGAAA AACTTAGAAG GTCATAGTGT      240
CGCTTTACTT GctCAAGATT ACTATTATAA AGATCAAAAG CACTTGACTT TCGACGAGCG      300
CCTAGAAACC AATTATGACC ATCCATTGTC ATTCGATAAT GATTTATTAA TTGAAAATCT      360
TAAAGACTTG AAAAATGGTA AAGCAGTAGA AGTACCGACA TATGATTATG CtAGTCATAC      420
AAGAAGTGgA CATTACCATT GGATTTTtAA ACCTAAAGAT GtTTATTtAT CGTAGAAGGG      480
CCTTTTTTCGC CTTTAGGAAA ATAAGGGTAT TTAC                                514

```

(2) INFORMATION FOR SEQ ID NO: 956:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 388 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 956:

```

ATTTTAAACGA ATATCTTGCC GTACGATGCC CAAGCTGCAT CTGAAAAGGA TACTGAAATT      60
ACAAAAGAGA TATTATCTAA GCAAGATTTA TTAGACAAAG TTGACAAGGC AATTCGTCAA      120
ATTGAGCAAT TAAAACAGTT ATCGGCTTCA TCTAAAGAAC ATTATAAAGC ACAACTAAAT      180
GAAGCGAAAA CAGCATCGCA AATAGATGAA ATCATAAAAC GAGCTAATGA GTTGGATAGC      240
AAGCCATAAA GTCTCACCTG AATGACGGTC AAGTGTATAG CATAATAGTC ATGCTAAGTT      300
AATGAGTCTC AATGTGTAGG GCACAAGTGC GGGCGTCTAT GCATGAATGT TGCCACGCTC      360
ACAACTGGGA AGTGTAATG TGA CTGTA                                388

```

(2) INFORMATION FOR SEQ ID NO: 957:

- (A) LENGTH: 394 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 957:

10 AACAAAGTAC ATCGCAGTTA CGAGAAGAAA TAAATGCATT GTATAATCAA GTTATAGAGG 60
 TTGGACAAGC GTTATCAAAG CAGCGTCGTA TAGTCGCTAG GGAGTTAAGw GACCACATTG 120
 TATCGGAAAT TCAAAACTTA CAAATGAAAG ACGCAAATCT TGAGATTTCA TTTAAAAAAT 180
 15 TAGAAGAACC GAATATTGAT GGAATCGAAT TTGTAGAATT TTTAATCAGT CCAAATAAAG 240
 GGGAACCATT AAAAAGTTTA AATAAAATTG CGTCAGGTGG AGAACTTTCT AGAATTATGT 300
 TAGCGTTAAA AAGTATTTTT GTTAAATCGa GAGGTCAAAC TGCAATTtTr TTcGATGAGG 360
 20 TTGaCTCAGG TGTATCtGGr caAGCTGcAC CAnA 394

(2) INFORMATION FOR SEQ ID NO: 958:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 503 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 958:

CCGCCATTTT TTTCGATTTT TTTTCAAGT TTTTCACGGT CTTTGAAGA TAGACTATTG 60
 35 AAATCTTTTG CACCACTATC ATCAAATCA ATATCTGCTT TCAATTTTGT GCTAGATTTT 120
 AAAATGAAAT TAATTGGTTC TTCAGcATAt TTGATGCCGA TATyTAACGT AGATTTCTGA 180
 GTGTTATTyT TTACGTCAGA ATCTATATTA TTTTCAAAG TGAATTCATT TTCGTGCTA 240
 40 TATTTATCTA ATGCGACAGT GATTTTACCT TTATCTTGAC GTTTGTGCC ATCTACTTTT 300
 TCTTGGTTAT CTAATTTGAT TTTTGATTCA TCATATTCTG TCTTTTACC AAATTCGTAT 360
 TTATCATTAT ATTTATtATC TTTWCTTTA GAAGATACGC CTTTAATTGT ATATTTGCT 420
 45 TCAGCATACG TGTATTTATC TTGATCGAAA TCAAGTGCCT AATCTAGTTT TAACTTATCG 480
 TCTTCTAAAG TATTAGTACC TTT 503

(2) INFORMATION FOR SEQ ID NO: 959:

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 307 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 959:

TCAAACGTAC AGAGCTTGTT AAGTGTITAG ATTATTTACG AGAGAGCGAC ACATTAGTTG 60
 5 TCTATCAACT TGATCGGTTA GGTAGAACGA CAAACTATT AATTGAATTA TCACAATGAT 120
 TCGATGATAA CGGAATTGAC TTACCAATT AGTAACCTGA ACATTTCAC GAAAGACCCC 180
 10 ATGGGGCCAA AATGTTTTTA CCGATGATGA GTGCCTTTTT CCGGGTTAGA AGTTAATTTA 240
 CTATGTGAGT GTATAAATA GACTnAGCAG CCACAAGnCG AGAGGCCGAA AAAAGCGGGC 300
 GCCCTCT 307

(2) INFORMATION FOR SEQ ID NO: 960:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 960:

25 ATAAGGACAT TTATAGTCGC CATATACGAC AACTAATGGT TTGCCGTTTT TCGAAGATGT 60
 CGTTGCTGAT TCTTTTTTAC CGCAAGCTGT TAAATTAAC ATGCTCACTA TAAATAATGT 120
 30 TAGTAATTTT TAGTCCATAA TTACTCCTAA TCCTAACCA n GTCnGCnAAA TTTCCTCCAT 180
 GATAATCCGC ATCATGTTGT TGATAATATC GTCTCGAATC AAAATCTTGC TTATATGATA 240
 CACGTTTATC ACTAAATCA TATGCATAAT AATGCACTTC ATCATCATT CTCAAAGGTT 300
 35 TATAAGCTAA AATGATAATT TTATCCTTAT CATAGACATA AATATTAGCA TCTTTGCGTT 360
 CAAAATGATC AACACGATCA CCGTGCCTTT TTTGGCAAGT 400

(2) INFORMATION FOR SEQ ID NO: 961:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 45 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 961:

50 ACAATTAGTT ATACCCAAAA GAGTTATGTC AGTGAAGTAG ACAAGCAAAA CTCAAATCT 60
 GTTAAATGGG GTGTAAAGC AAACGAATTT GTTACGCCTG ATGGAAAAA ATCTGCGCAT 120

55

GCTCCTGATA ATCAATTGCC ACCTTTAGTT CAAAGTGGCT TTAATCCATC GTTTATCACT 240
 AACTATCAC ATGAAAAAGG TyCagTGATA CGAGTGaATT TGaAATTTCa TATGGTAAGA 300
 5 ACTTAGATAT TACCTATGCG ACTTTATTCC CTAGAACTGG TATTTACGCh GGAAGGAAGC 360
 ATAATGCATT TGGAAATAGG AACCTTGTAh TTAnGTATGG 400

(2) INFORMATION FOR SEQ ID NO: 962:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 358 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 962:

CTTTTACAAA GACAAATTAG CGGTAATCGG TAAnGATCGT TGTTGATCCA GTTGCACTAT 60
 TGAAAGAAIT AGACGGATTA AATGAACGTG GCATTCCTAC AAGTAATTTA CGTATATCTA 120
 ATCGTGCGCA AGTGATTTTA CCATATCACT TAGCACAAGA TGAATATGAA GAACGTTTAC 180
 25 TGGTGACAAT AAGATTGGTA CACTAAAAAA GGTATCGGTC CAGCATATGT AGCCAAGTCA 240
 CGTATCGGTA TCGTATGCAG ATTACTTGAA AAGAACATCG AAGATATTAA ATCAACATGA 300
 TATAACCAGC TATTCAAGGA TGTTACGAAC ATGhCACATT GATGAAnCTT GAGATATA 358

(2) INFORMATION FOR SEQ ID NO: 963:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 963:

AGTGAGCAGG TTGAAGTAGT AGGGCCGATT CAAATGGAAA GAGATACACA TAAAGATGGT 60
 AAGGTAAAGT GGCAATTGCT TTATATAATG AATCAGGATG ATGATGAAAT TAAGCCACCA 120
 45 TTTTTTATTC AATGGGAGGA AAGTGATTCC ATTGCGTACT AAAAAATTGC AAAAATATTT 180
 TCAAAACCAT TTTCAATTGA AACTGTATTG TGAAAGTAAA ACCGTCACAA CAGTATCGAT 240
 50 TGGTTGAAnT GGTTGhAT 258

(2) INFORMATION FOR SEQ ID NO: 964:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 388 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 964:

	TTATCAAACC GCCATAGCCA CATCTCGATT CAAGGATAAA TTTGTTTTCT CATATCTTAT	60
10	CTCTCATGAC AGTGCACGAT ATATATCTTT TCCAAATATT CATCACTTTT TTCCTGTAA	120
	AATCATTTCT ATTATTAATG CTTTAATGAC AACTTATTTA ATTTATTGCA TTTAATGCCC	180
	TGCTTCATTT TCAAATATTA ACTTTAATCG ACTAGCTCAT CACACTATTT ACCCTTTCCA	240
15	CCTATCACTA CTGAATTCAA AACGATTTGT TTACAAATCC CCTTATGCCA ACGGTCGAAT	300
	TTCATCATGT TTACATTTAA TTGCAAATGC CCTCTTGAAA TATTTTAAAA TTAATATTAT	360
	TCTTAGCATA CTAGTCGGAA TTATATTC	388

20

(2) INFORMATION FOR SEQ ID NO: 965:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 602 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 965:

	AAAGTGAGGT ATAGTCOGAC CATACTCTAA AAACGTAGCG AGATAAATAT ATTTCAATCC	60
	TAACTTTTAT GTTTTGAGGC ACTTGCCATT TAGGATATTG TCGTTCGTAA TACGACACTT	120
35	GTTGTATAAA TACACCTAGT CCAAATGGCA GCATCATGAG TAAGATACTT CTTAAATAAC	180
	TTAAACCAAT ATCATGCCAT ATGTGTCCAA TAATCAATTG AAAGACAATG ATAGATACTA	240
	TTAAAACGAT TATATTTATT GTCACCTGTT CAAACGCACT CCTTTTCCAA ATAATAGAAT	300
40	TGCTGCTTGC ATGACAACCA TAAACATAC AAACATAGCA GTTTTAAGCG TTAGACTTTC	360
	TAGAATGTGa TTTAGAACAT GTAAGGGCTC ATTAAAGAAA TAAACGGaAT GTAAGCGTAA	420
45	GAAACGACCA ATATAAATTC CGAATCCATT TAAAAACATT AGcACGaCAA CrATTAATCT	480
	ATTaAGCCAA CGGTGaGaAG TCAATGtTAG TATTTCAAAA TAGATTnAAA TCATCACATA	540
	AACCGCTAAG AAGACACCAA GCAGTAAATA GGTAAAGTAT TTCCACTCAC TTAAATTTAG	600
50	TC	602

(2) INFORMATION FOR SEQ ID NO: 966:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 744 base pairs

55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 966:

	GTCCTAAGGT AGATTTATCT ACAACTGTTG TTATTTTAnC AATTATTGGA GCGATTTTAT	60
10	TAATGTTTGC TAGCATATAT TTTATHAAAA TTGTGACATC ATACTTTACT AAGAACTTT	120
	TAGAAATTAA ATTTAATTCG AAATAGTATT TTCAATGAAG TGCTTCAAAG CTAAACTTTC	180
	ATGACATTTT TTTGAAACTA GAGAAAATGG CTCATCAACT TCTAGCTGAA TGCCTTCTTG	240
15	AAAATGTATT AATGTTGAAT TCTCTATGGC AGTCATACCA ACAATATAAT GTGCATTGAT	300
	ATAAATTTGT ATCGGTGCAC GTTTATGTTT TAACGGAAAA AGTATAGTTT TTGGGTTGAT	360
	GTAAATAGGA ACCAGTTTAT TAATATTTAG TATGTTTTTT GCAGTCTCAA TTAGCACACT	420
20	TTGTGATGAT AAATGCAATT TCGCATAATA TGAAACGAGT TTGTTAATTG GAACTTTAAG	480
	AATGTAATCG CAATGTAAAA AGACACAATT TGTTTCGAGA TGGTTTGAAG TAGCAGTTTG	540
	GATATAAAGT AGATATTTAG TAGAATTGTC TTGCAAATAT AGTACTCCTC GTTTAATAAG	600
25	TTGTTAATTT AATGGTTAAA ATAAATTTAT ATGAAGATGT TAGGGTGTTT CAACATCATT	660
	AACACATCGA CCGAAATTTT ACAATAAACT AATTAAAATT ATGTTGGCAT TTACGCTTAA	720
30	CGGACATCTT AATTAATTTA ATCG	744

(2) INFORMATION FOR SEQ ID NO: 967:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 503 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 967:

	TACTTTGCAG GAGCATTAGG TATTCTAGCA TTTATAATGT CAATTGTATT GATTCACGAT	60
45	CCGAAAAAGT CTACGACAAG TGGTTTCCAA AAGTTAGAGC CACAATTGCT AACGAAAATT	120
	AACTGGAAAG TGTTTATTAC ACCAGTTATT TTAACACTTG TATTATCGTT TGGTTTATCT	180
	GCATTTGAAA CATTGTATTC ACTATACACA GCTGACAAGG TAAATTATTC ACCTAAAGAT	240
50	ATTTTCGATTG CTATTACGGG TGGCGGTATA TTTGGGGCAC TTTTCCmAAAT CTATTTCTTC	300
	GATAAATTTA TGAAGTATTT CTCAGAGTTA ACaTTTATAG CyTGGkCaTT AWTATATTCA	360
	GTTGTTGTcy TAAAwTAWT AgTTTTtGCT AATGGACTAT TGGGTCAATA ATGGTTAATC	420

55

TTTTCCnAAn ATTGCnGGGG AAG

503

(2) INFORMATION FOR SEQ ID NO: 968:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 465 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 968:

TACAGGTAAA AAAGGTGGTC CGTTTATCGA TTCATTAAAC CATCGTAGTT ACACGAAAGC 60
 ATATGTTTCA TCAGAAACCG ATATTCCATT TTATGCTGAA GCATTGGAA TTAGAGAAGA 120
 AAATGTTGTA CCAACAGGTG TACCACGTAC TGaTGACTA TTTGATGAAG CTTATGCAAC 180
 ACAAATTAAA CAAGAGATGG AAGATGAATT GCCAATTATA AAAGGTAAGA AAGTTATTCT 240
 ATTCGCACCG ACATTTAGAG GTAATGGTCA CGGTACGGCA CATTATCCAT TTTTAAAAAT 300
 TGATTTTGAA CGTTTAGCAA GATACTGCGA GAAGCATAAT GCAGTTGTGT TATTmAAAT 360
 GCATCCGTTC GTAAAAAATa GmCTTAATAT TcaCgTGaAC CATAGACCAT ACCTTAATCC 420
 GATGTGGTCC AGATCCATCC GTGGAAAGTT AAACGGATAT TCCnC 465

(2) INFORMATION FOR SEQ ID NO: 969:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 969:

AATTCGGAAC AACGAGCTGG CAACAACATA AGATGACAGA GGCGAATGGT CAATCAATTC 60
 AAGTTAACTT AAATAATGCG CAAnGCGATT TGGGATATTT AACTGCTGGT AATTACTATG 120
 CAACAAGAGT GCCGGATTTA CCAGTAGCGT TGAAAGTTAT GAGGGTATTT ATCGGTATTC 180
 GTAAAGATGA TACAACACh ATTTACTCAC ACCTATACTC TAAAAGATTh CACACGTCAT 240
 CACAACGGCG ACTTGAGCAA CATGGACAGT TCCTAATGAA CATAATCAAC GGTATTGTTC 300

(2) INFORMATION FOR SEQ ID NO: 970:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 970:

5 ATACGCGAGC GAACGGTGAA GCGACTTGCT GCTGCAAAAC GTCTGCGACC TTGAGCAACA 60
 ACATGAATGG TCTTCGGTTT CCGTGTTCG TAAAGTCTGG AAACGCGGAA GTCAGCGCCC 120
 TGCACCATTA TGTTCCGGAT CTGCATCGCA GGnATGCTGC TGGATTATCA TCCGATTTTT 180
 10 AAAACGTTAT TAGGCGTTGT CGTTGCAATT CAATTTACAG TACCTGCTTT GATTGGTGCA 240
 TTGGTTGCCA TGAAATTTGA CCTATCGCCA TTAGCAATTG CTGTTGTGGC AAGTGCAGCA 300
 TATGTAGGAA GCGGTGCAGC GCAGTTTAAA AATGGTGCTT GGATGATTAC GGGTATTGGT 360
 15 GATTTAATna ATACAATGAT TACAGCAGCG ATTGCAGTTG 400

(2) INFORMATION FOR SEQ ID NO: 971:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 251 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 971:

25 AAGTGTAAGA TTAGAATAAA AAATTTATAA ACGAATTGAG AAAGACCAAA GCTAGCAATT 60
 30 ATTCAGAAAT TAAATCGAAG AGTCCTACAG TTAATGACTT ACCTGAACGA GATTTATCna 120
 ACAAATCTCA GTATGACCAA TATGGCAAAT GCCGTGTCCA TGTTAACTGA TGAAAAGTAC 180
 TTTGGTGGTA GTTTTGAAAG ATTACAAGCA TTGACGACAA AAACAACATT ACCCGTATTA 240
 35 TGCAAAGnCT T 251

(2) INFORMATION FOR SEQ ID NO: 972:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 414 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 972:

45 TGAAGCTCTT aATAAATCTT TTCTTGAACA CACACCAACA AAATCGTTAT CTTCATTAAAT 60
 50 AATAAATAAT GTACTTACAT CTTCTAAAAA AATTGTACAA ATAGCGTCAT AAAGTGTGTG 120
 ATTCTCTCTT AGCACAACAG GTTGAGACAT ATAGTCCTTA ACCTCAAATT GTCGAAGTTT 180

TTCTAAAAAA CCTGACATTG TTAATATCGC TAAATCTGGT CTAAGCGTTG CTCTTGTTAA 300
 ATTCAACTTA TCTGCTATTT GTTCACCAGw AATGGGTCCT TTAGTTTTAA CAATTCGAT 360
 5 GATTCGTTCT TGTCTTTGAC TGAGTTCTrw AGGkCTTCAC CCCTTTtTAT GCnC 414

(2) INFORMATION FOR SEQ ID NO: 973:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 973:

GAAATCGATA CTAATATCTT CTTTATTTAC ACCTGCTAGT TCAGCTTCTA AGTAATATAA 60
 20 TTCGTCTGTt CATAtATATt CGTtACAAaT GAAGGTGTtG AAAaGTAaTT AAaTaCTTGT 120
 TTaCCTAAAT CTTTAAaTGT ATCACTTGGA TTACCGTTAA AAAAGTTTTG ATTCTCGAAT 180
 TGATTGAAGT TCAtATCATC ACTTCCATTA TAGAAATTTG TATTACAAAG CCATTATATG 240
 25 AGTAGTATTA AAAAAGTTAA AGTAAAAGGG GAGATAAAAA AGTAATAAAT GGGTAGATTT. 300
 GTGCCTGATA AGGAATCAGG GCAAAAGAAC TGAGCATCGA GATTTATACA AGCAATAAAG 360
 ATAATGAATA AACATATCGA TGCATACTTC TGCCCATCAC CTTATGATAG GGAGTTATTT 420
 30 TTTGTCTGGAA TAGTCAAAAT TTTTAGAAAC AAAATTCCGC AAATTATTAA TTATATCTGA 480
 TGCTTCTCGT TCTGAAATGT TGAATTCATT AAAGACcTTA ACAGAAATTT CAGCAAGAGG 540
 GCTTTTTTATT GCTTTACCTT GTTCTGTTAG AGAAATTTGT AGGTTTCTTT CATCTTTCTC 600
 35 TTCACGTGTT CGAACAACGT AATCTTTCTT TTCTAATTTT TTTAGTAATG GkGTCAGTGT 660
 TCCAGAATCT AAGAACACAC GTTCACCTAA TTTTTTGATG TTAAGTTTTT CATCATTTTC 720
 40 AATCGCCAAT TAAACCAngG TanCnGG 747

(2) INFORMATION FOR SEQ ID NO: 974:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 257 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 974:

TGCTTCCAAG TAATGTGGAC ATTTAATACG AGCTGGGATT GAGATACACC TTGCAATGCT 60

TCTTGGnCAT TTGCATTGCT TCTCCCAAGT GAATGCCTTT AACAGCTTGT GTCATCATCG 180
 ATGACTTGCC ATTGAAATCG ACAACCTTCA CCTTCAAAC TAGCATCTTT TATAATGCCG 240
 5 nCTTCTATAT CAAATGT 257

(2) INFORMATION FOR SEQ ID NO: 975:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 975:

CAATGGAGCA TATTTTACGT TCGGCAATGG ACATTGATCA nACAACAAA CATTATATAA 60
 20 mTAATAAAGT AGAAGTTATC AAAGAAATTA TTAaaaaAGT AAGAGGGGGA AATCACGTGA 120
 CATCACAAAA TCCAGAAGTT AATTACGAAG CATTAGCTAA ATATGGCCGC GACTTAGTAG 180
 AAGAAGTTAG ACAAGGTAAA ATGGATCCTG TTATAGGAAG AGATGAAGAA ATTCGAAATA 240
 25 CGATTTCGTAT TTTAAGTCGT AAAACTAAAA ACAACCCTGT GTCATTGGT GAACCAGGTG 300
 TTGGTAAAAC TGCAATTGTT GAAGGATTAG CGCAACGTAT AGTTAAGAAA GATGTGCCAG 360
 AATCATTATT AGATAAAACT GTTTTGTAGT TAGATTTAAG CGCATTAGTA GCGGGCGCTA 420
 30 AATATCGTGG TGAATTTGAA GAGAGATTAA A 451

(2) INFORMATION FOR SEQ ID NO: 976:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 571 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 976:

AACTATTAAT AATATTAAGT TCACTACAGA TGTTGCTAAT GGACCATAAG TTTTAAAGAC 60
 45 ATCTTCACCTT TTATAACCAA CAATCGCATC TAAAAATTGA ACTAAGATCA TTGCAATGGA 120
 TATAGTTATC AAAAATATAG CACTATGAAT GACTAAAGAA AnaATAGCTA ATAAaAATAA 180
 AGGTAAGCTT CGACTAAGTG cATAATATGC ATTTATATTA TGGCTAGATG CACATGCTTG 240
 50 AATTGAATAA CCTAAACTTA CACTGGCACT GATTATTGTA AATATTGCTA AAACAAAATA 300
 CATGTTAATC CTTCTTTCTA TATTTGGATA TAAACAAGTA CTTGTCTAAA GTTATTTAAA 360

TATAACTATG TCATAAACTG AATTTGTTGA AATTTTTCAT TATGCAAATT TATTAATAAC 480
 AAACAGCTCG AACTATAGCA TCATTTTACT AATGAATGCA TTAAAGTAAC TATGACTAAA 540
 5 AATGCATATT AATTATCATT ATTAAGACTA T 571

(2) INFORMATION FOR SEQ ID NO: 977:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 274 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 977:

CCCGAAAAAA TTTAGACATT TTCAAACAGT AACATTTGAA GATTTGTCTA AGCTGGAAAA 60
 20 GAGTAGTATG CCATCAGGGG TCGGTATATT AGGCTATGCT GTTGACAAAG GTGTTGCTTT 120
 AAACAAGGGG CGCATTGGTG CAAAAGAAGG ACCAGATGCG ATTAAACAAG CATTTCAGG 180
 TTTGCCGGAT TTGAATCCAT GTGnACTTTA GTCnGATTAC CGGAATGTTT ATCCTGGTCA 240
 25 TGGGGGAATT AATTGGATAC CCCAAAAAGG AATT 274

(2) INFORMATION FOR SEQ ID NO: 978:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 877 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 978:

nTnTTACTTC TTCTTTTAAT ATTTCTCGTA TAGCTTCAAA TTTTTCCTG TCAAACAGTG 60
 40 TTCTTAATAT CCCTTGCTTT TCACGACTGT TTGATATTAA AAAGCGTTTG AATTCACCTT 120
 GAGGCAAAAT AAACAATTGT CGGAATTGAT CTGCATTTAC TCCTAATAAT TCAATAATGA 180
 ATTGGGTACC TGAAATTACT TTACTTTCTC TAATTTCATA CTTGCCATCC ACCATCTCAA 240
 45 ATaCATCAAA TTTAGCGTTC GTTTTGTGTT TATTACCTTC TTTGATATAA GGGCCTTGTC 300
 TATGCACTTT ATAAATTCTGA TGATTTAATT GAAATTCAAA TGTTACTGAC ATCGGCTGTT 360
 TACCATCAGC GAAATGACTT CTCAAATCAT TTTCTTCTCT TTGTTTCAGTT GATGCTTTAC 420
 50 CAAACAAGGC ATAAGTCATT GCATCAAAAA TCATTGTTTT ACCCGATCCA GTCTTACCAC 480
 TTATTAAAAA CAATTCATTA TTATCAATTT TAGAAAAATC AATTTCTTCT TTTAAAAAGG 540

CGTTTTCTAA AATATTTTTT ATCTTTTATG ATTGGATATC CgATAATTCT TTATCAGTTA 660
TATGTTTATA AAACATTTTCg ATAATTGACA TGTCaTCTTT TTCAC TTATT TCyATAGCAT 720
5 TATTTTCTTC ATnGTAATTA AAAGTTTCAT TCGTCAGCGC TAAAGTATTA GGATAAATTh 780
GTTTTAAACT CATCATTGGA TCAGTAATAT GAGACATATT TTTAAGTTTA AAATGTAAAA 840
TAATTATCGT TATnTTTCAC ATGAAC TTnT TCATTAA 877

10

(2) INFORMATION FOR SEQ ID NO: 979:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 624 base pairs
(B) TYPE: nucleic acid
15 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 979:

ATTGGTACTA TACCATTTGT CCATTTTAGC GAAGACGTTA AACTATCGAT TAGCTAGTGA 60
AGAAATTAAT CTCACTATAC CTTCTATACA TAAGCAAATT AAAAATTTAG AACAGCATCT 120
25 CAATGTGAAA CTATTTGAAA CTTATAAAAA TCAAATTATC TTGACTGAAG ATGGTCATAC 180
ATTTCTTCCA ATTGCGCAAA GTTTTATTGA ACAATATGAA AGTGGCATCA AACATATCCA 240
ACTTAAAAAA ACGATGTTTC AATCGAAATT AAATGTTGTC GTGTCTTCAT ATATCGCGAC 300
30 ATTCATTATG CCCAAATTTT TAAATCTTT CTTTAACGAG CATCCTTTTA TCGATGTTTC 360
TCTTCATGTC AAAAATGAAA ATATTGAAAA AGATATTAAT AATCATACGT ATGACATCGG 420
GATTAGCAGA AATCAACCTA AACTAAGAGA AGTTCATTCT GAAAAAGTAT GCGAAGGTAA 480
35 AATTGTATTG ATTGCTCCCA ATAAAGAGAA TAATCATCTT CTAAGTGGG CATCTTTATT 540
TGAAAAATAC AAAATAATAA GTGATAACCA TCCAGAATAT TGGGTCATCA CTTAAAAATA 600
40 ATATTTTAAA TATTTATGAA AAAG 624

(2) INFORMATION FOR SEQ ID NO: 980:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 457 base pairs
45 (B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 980:

TTTGTCTGT TGCATAAACT GCAGCTAGCA TAATGTTAAT TGTTCACCT ACGCTAACCA 60

55

TKGATTCATC AATTTTCAGCA CCTAAAGCTT TAAATCCTTT AATATGTTGA TCAATTGGAC 180
 GTGGCCCAAG TGGACATCCT CCCGGTAAAC CAATCACACA TTTTTTAAAT CTACCTAACA 240
 5 TGGCACCCAT CATATAGTAT GAAGCACGTA ATGATTCAAC TTTATTATTT GGtAATGCAG 300
 CATTtTGtAT TTCAGTTGTA TCAACTTCTA AwTCCGTACC ATTTaGTGaw GcCTTAAtAt 360
 TTAAAtCtKc tAAtAAACTC aCTAAAGTTT TAACATCAGA GATTtNgCGG TAACCCTTCT 420
 10 AATTTACATAT GTCCTTGAGC TAATAATGTT GCAGGAA 457

(2) INFORMATION FOR SEQ ID NO: 981:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 997 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 981:

ATTGCAGCAA CGCCCCCTGT AAAACTACGT ACTCTTACAT TTATGAAACC TGCTTCTTCA 60
 25 AACATGCGCT TCAACTCTTC TTTTCCAGGA AAATTAAACG TAGATTGCTG TAACCATTCA 120
 TATTCTTCTT TTGATTTTGC AAATAATTTT CCAAAAATAG GCATAACAAA TTTAAAGTAT 180
 AGTGATACA TTTGTTTAAA GACTGGCAA GTTGGTTGGC TCGTTTCAAG ACATACCACC 240
 30 ATACCACCTG GTTTAAGTAC TCTATTCATT TyTTTTAaCG CGACTAAATA GTCTGGCACA 300
 TTTCTTAATC CAAACCCAAT TGTTACATAA TCAAAAGAAT TGTCTTCAAA CGGCAATTCC 360
 ATTGCATCAC CATGAACAAG TTTAACATTT TCCATTGAAG CAGTTTTTTC TTTTCTACT 420
 35 TCTAACATAT TCTCACTAAA GTCAATACCA GTAACCTCAC CTGTTGGTCC TACAGCTTTG 480
 CTTAATGCGA TTGTCCAATC ACCAGTACCA CAACAAACAT CTAATGCTTT CGTCCCTTTT 540
 40 CTAACACCCA TGTCTTTCAT GACGCGTTTT CTCCATACTT TATGCTGCTC AAAACTAATA 600
 ATATTATTTA ATCTATCATA TTTTTTTGAA ATATTTTGAA AAACGCGATG TACTTGCTCT 660
 TTATTTGCTT TATTGTCAGC CATGCTTAAT TACCTCTACT TTTTAAATAA CTTTTTTGGA 720
 45 TATCGTGTA GTAAATGCTTT ACTTCATTTT GATTATATTT CTTGAAGTAT GATGGATAGT 780
 AATCAGACAT ATCTTCAAAT AAATAATTAT ATATTTCCGA CTCATCAATA TTGATACCGA 840
 AATGAGATAA CGTAAtATAT GGgAAAAGTG tTTCAATTTT TACTATTGCT TGAGAAATTT 900
 50 CATAATCATT TAAAGCTTGA TGATGTAATG AAGATTTCAA TTCATTAATT TCAACAATTG 960
 CTTTACTAAT TTCATTTTGA AATGATAAAT CATTGAT 997

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1118 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 982:

10 CATTCCCTTT ATCGGnACAA TTATTTTAAA TAATTGGTTA ATAAGTTTTT CTATTAATTG 60
 CTGTTTTACT TTCTAGAGCA ACAATAAATG TTTTCAGTTT AAAATAATTA AGATTGTCAT 120
 15 TTAATAATTT GCTTTTTCCA TTCTTTAACC ATAAATCAAC TTGTAATTGC CACTCAATAG 180
 GATGATTTTT AATATAAATT TGCTCTGGAA AAATATAACC ATAATTGTGC ACTACTTCTT 240
 GATCAGTTAA CTGTAATTGA TACATTGCAC TTAAAGTGGG TTCTAAAACA GAATTTTGCC 300
 20 AGCGACAATA TTTGATATAT TGATTAATTG CGAACTTTGA TAATTTATAA ACAGCATGAT 360
 AATCACAAGG CGCCTCATT AACAACCTCAC TAATTTTGGC ATTTCTTTTT TCAGCGACAT 420
 ATTTGTGCCC GCCTAAATAT TGTATTTGTT GAAACATCAT TATTTGTTTT TTCTCTAAAT 480
 25 TGAACGTATG AAGTTTTCGA GTGAATGGAT TAATAACAT ACTTTGAAA TGATTGAACT 540
 TCACTTTATT ATGACAATAA TCAACATCAT TTAmTAACCA ACTTACTTTA TATCCCATT 600
 TCTTKAAACC TTCCGTTCTG TGAAGAATCT GTTTATAAGG AATTGGCGAA TACTGTAGCT 660
 30 CAATAACATA TTTATTATTA ATCAAAATAT CCGGAATTG TTTTATTTCT TTTAAAAATG 720
 GCTCTATCTC TACTTTACAT CCTTGTTGCT GTATCATCTG TGCTAAAATC AATTTACTTT 780
 35 TATAATGTTT AATCGTTTCA TTATTAAAC ATTTTGTTC CGCTAAATGT TTATGTGCAA 840
 AATGACTAAT TACTTTGAGC CCACGCTTTA AAATAACTTG CTTGCCACAC ACCGGACAAA 900
 AATATTGTGT CTTTCTCAAT GCAGTAGTTG CTAATACGCG TTCCTTTTCT TCATTTAAAG 960
 40 CTACTAACAT ATCTTCACCT CAAAATATTA CACGTAAGTT TAACTACTAT TACTTTTTCA 1020
 GCTTTGGAAA TCCATTAAAA AAAAGGGACC CCCAAAGGAG GCCTCCTTTC ATAnGTGCGG 1080
 TGGCCTAATT AAATAATCCG TTAATTCCAG GTGGnChC 1118

(2) INFORMATION FOR SEQ ID NO: 983:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 269 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

CATCCCCTGT TAAAAATTGn TATTTGTGTA TTGATGCTTC TCAGnCATT TTACGAGTCC 60
 ATATTCATTG CTTCACTGAA ATATCAGTAG CAATAACATT TAAATCCGGC TTTTCACATT 120
 5 TCAAGGTAAT TGCAGTACAC CACTACCCGT TTCCGATATC TACGGATTGT TGCATCATCT 180
 TCTAACTGTT GGTAAGAAAT GCAACCATT CTGCTTCAGT TTCAGGTCTG GATCAAACA 240
 ATTTGAGTTT ACATCAAACG GTCTACCTT 269

(2) INFORMATION FOR SEQ ID NO: 984:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 298 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 984:

AACCCAAGTG GAAnnGGTTT TGGTGGAAAA ATTTTAGGTC CTTTATCTGT CGTGCGAATC 60
 ATTGTATTTT TTAAACTTTC CGCATAATCT TGATGGTCGA TTAACTTTTC GCGATAACAA 120
 25 TATTCAATCA CTTCAGAAAT CGCTTGCTCT GATATCTCTT CTTTTGTAA ATATTGTATA 180
 ACTTCTTTT CAGTTCTCTT TTTATATGAT AAATATTGGA TTGCTTTATT TAAACCTATG 240
 CGATAATGAT CATACTTTTG AATCTCTGCC ATGTCAGCAG CnTCAAGTTG TnGCCCTT 298

(2) INFORMATION FOR SEQ ID NO: 985:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 431 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 985:

CCATGCGATT CCACCTATTA ATAGTAAGAC AATAATAATA GGAATTATAA TTTTAACTT 60
 TTTAGACATT TTGCTTCCTC CCAATTATTT ATGAGTTAAT CATATCAGAA CAATACATTT 120
 45 TATTAAATGG TTTATTTTAA ATTTTGTTTA AATTAATAGA ATTTTAGTTA TAATAATGTT 180
 TAATAAGTTA TTGGrAATCy AATAAACTAC mAAAAATAGT TTGATTACaT AATGATTCTt 240
 GAAAAATGkT GGkTAACyTA ATAATATGCA TTTTTTkGGC GArGARgATT tATTtAACyT 300
 50 ATAAAAATAT kGAaGTAAGA TtGGGGAGAT tATGAATTtA TGGrATTgAA AGTCGATgAT 360
 TTTGTAAAGA ATATAAAAAG ACCATACTTG ACTGTATTGG GAGTATTTGT AGTTGCAGTT 420

(2) INFORMATION FOR SEQ ID NO: 986:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 986:

AAGGTTAGGn TGGAGTACCC AAAGTAGGTT TTAAAGGTTG GAAATACnAG GAGGATGAAT 60
 AACATGAATC AGTCAGTCAA ATTACTTAAA CATTTAACAG ATGTAAACGG CATTGCTGGT 120
 TATGAAATGC AAGTTAAAGA AGCAATGCGT AACTATATAG AGCCTGTCAG TGATCAAATT 180
 ATTGAAGATA ACTTGGTGGC ATTTTGGAA AGAAAAATGC TGAGAATGGT CAATACTCAT 240
 TATGATTTCT GGTCTATGGA TGAAGTGGTT ATGGTAACAA GATGTAAACT GTTATTCATT 300
 ACGCGTGGTG TGTGATCAGC nGCACTCAA 329

(2) INFORMATION FOR SEQ ID NO: 987:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 545 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 987:

CATATTATTT nTATATGAGT TTACCAGTGT CAAAAAACA ATTGCTTAAT GCCAACTATA 60
 TTACATGCAT TGTTTTAAACA TTAATTGGAA CGCTTGTTAT TAGTTTATAT GCTTATGAAG 120
 CAGATGTGAT TGAACCAAAT TCAATCTATT TTTCAACTGC ATATGCATTT GTCATATCTA 180
 ACTTCTTGTC TATACCAATT GCATTTAGTC AATTACAGA ATTGCGTAGA GTCAAAGTGC 240
 CATATGGTAT ATACGTGTTT ACTATTATCA TTTTAGTTCC ATTTTATTT TCAATTGCAA 300
 TAGTATTGGT GAATTATTTT GTTCTAAGTC AATCATCATT CCCAGATTTA TATTATATA 360
 TTTTAAATAT TGGTTTTCTA ATTATAAGCA TTGTTATACT TATTGTTAAT TATTTAAAC 420
 AGCTCAATAA AATAAATACT AGAAAAATTA AAGGAGGCAG TCGATGAAAT TAGAACATAT 480
 TACAAAAAAA TACGGCTCAA ATGTCGTTT AAATGATATT GATTTTGACT TTGGCGAnAG 540
 TAGAA 545

(2) INFORMATION FOR SEQ ID NO: 988:

- (A) LENGTH: 607 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 988:

ACCTAGCATG ACATAATCAA ATATATCAAA AACACCTTTC TGAAGTGAAT CAACCGTTCC 60
 AATAGTTAAT GGTCTTTCAG AATCAGATGT CTTTCCACCT TGAACAGTCA GATGACTAAT 120
 TAAAATATTG ACGGCATCTT CATCAATTTT CGGTGCTATT GTTTCAATAC AGCGCGTAAT 180
 TCCCTGTTGA TGTGTTTCAA TGGTGTCAAT TTCAAAGTAG TGTTCATTT CGCTCACAGT 240
 AGCATAAGGG AGTGATATAA AATTAACCCC ATTTATCTCT ATTGGTGAAT TAATCGATGT 300
 AAAATCTGTT CTTATAAATA ACTGATTATG TTCAAACCAA CTCGCCCCAT AGTTTAACCT 360
 CTCTTTACCA TCGTGATTTC CACTAATAAT AATTATTGGT ATACGCAGTT CTAAATTTAA 420
 CTTTCCAATC GCTTGTTCTA ATAACATGAT TGCATCTTTA CTTGGATATG TTGTGTCATA 480
 TAAAcTCCAG CTATCACAAT GTATCAGGT TCTTCTTCTT TCATTTTTTC TACGAACATA 540
 TCTAAAATAT ACGCTTGaTC TTCTAAAAGC TGTGTGCCAT CTAATATTTT CCCTAAGTGC 600
 CAGTCTG 607

(2) INFORMATION FOR SEQ ID NO: 989:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 722 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 989:

ATCGCGTTAA GCAACAATTA CAACAATCAT CATATAAGCT AACGCCACAA CGCGAACTAC 60
 TGTTAGAGTT CTAATTGAAA ATGAAAAAGA TCATCTAAGT GCTGAAGACG TATATCTGAA 120
 AGTAAAAGAT AAAGCGCCTG AAATTGGCTT GGCGACAGTA TACAGAACGT TAGAGTTGTT 180
 AGCTGAACTA AAAGTTGTCG ACAAATTAA CTTTGGTGAT GCGGTCGCTC GTTTTGATTT 240
 AAGAAAAGAA GGCGCAAAAC ATTTCCACCA TCATTTAGTA TGTATGGAAT GTGGTCGTGT 300
 AGATGAAATC GATGAAGATT TGTTACCAGA AGTTGAAAAT CGAGTTGAAA ATGAGTTCAA 360
 TTTTAAATT TTAGATcATC GTTTAACTTT CCATGGTGTG TGTGAAACGT GCCAAGCTAA 420
 AGGTAAAGGA TAGTAAATTG CGTAGGTAA ATTAACCTTC GCTTTTTTTA GAGGTGTGGT 480

TAACACAATT GGTGCTTATA GACGTGATTT GAAAAAGTAT CAAGATTATA TGA CTGAACA 600
 TCATATCTCG CATATTGATT TTATAGATCG ACAATTAATT CAAGAGTGTh TGGGGCATT 660
 5 AATAGACCAA GGGCAATCTG CTAAATCTAT TGC GCGATTT ATTTCAACAA TCCGTAGTTT 720
 TC 722

(2) INFORMATION FOR SEQ ID NO: 990:

10

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 366 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 990:

20 ATCCAAAAAA TACGAAAGCG CTTTCTATAT TGGTATGCAA GTATTTCAAA AAGAATAAAT 60
 TTAATTTTCC TACTTTTCTA AACATTTATC TTTATGTATA ATGTTTTC AA GTA ACTAAAT 120
 TATAAATTAA ATAAAGGGAG TGT TTTATCAT GCTTACAATG GGGACAGCAT TAAGTCAACA 180
 25 AGTAGATGCC AATTGGCAAA CTTATATTAT GATTGCCGTC TACTTCTTGG GTnATGACTC 240
 CAACTTATTG ATAGTGTTnT TATGTTTCAG ATAATGCCCCG ATGACTTTGT CATGCCAGCT 300
 CCACCGATTT TGAGnAACGA CAGCGACTTC CGTCCCCAGC CGTGCCAGGT GCTGCCTCCA 360
 30 GTATTC 366

(2) INFORMATION FOR SEQ ID NO: 991:

35

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 338 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 991:

45 TGACTACAGT AACATTACCT GTAAATAATG ATTGAATGGC TTCATTTTCA TCATGAATAC 60
 CTGTAATAAAA TTCTAATTCT TCATCTGAAA CTTTTACAAT ATGCGCCAAT GGCAAAAAGG 120
 TATGAATAGT TTGTCTTAAA TCTTCAGCGT TATCCCATAA TGGCAATCTT ACATTAGGAT 180
 CAAATACGAC AGTACCATTG GCATTTAATG TTTTGGTAAT CAATTGATAA TGTGCATCTC 240
 50 TCATCGGACT ATCGACCAAA TCAACAGAAC AAAAATGTAC TACATCATTT TCATTACAT 300
 CaATATCATT CACAAAACtk GGkTCaAaTA aCATATCn 338

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 371 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 992:

AGTTCATCGT ATTCTTAATA CCTATATGCT ATCGTTCACA AACTTTTGTT CTTTCTTCC 60
 AACATACCGC CTATAATGCT CAAAAAGAAA TCAGCAGTGG AACCACTACG TTCCCTACCG 120
 CTTTATCATT TTCAGGCTTA AATATTTTAA AAATTTTGTT AGGAATCCGC TTAAAAAAC 180
 TCATATCTTC AAAAAAATG ACAAATTTA AATTTTGA TAACATCATT TTTTAGTAAA 240
 TGTATAATT TCTTATAAT ACAAATATCG AAAGAAGGCG ACTGACTTTG AAAAAATTAT 300
 GTTCATTAAT TGTTAGGTAGC ATTAGTTTGT ATCCTTGCCA TTATCAGCTT GTGGnAAAGA 360
 ACnAACcnAA A 371

(2) INFORMATION FOR SEQ ID NO: 993:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 530 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 993:

ATAATGATTT ATCAATCGAT GTTAGATGCG GTTCAATTAT ATCCTAAGTC TATCGTGGTT 60
 GATTATACTG TGGATGGACA ATAcAAAAAT GATTGTCACT ATTTCGGTCA ATCATCGATC 120
 AACATTGCCG ATTGGGCTCA AAACAATAAT TATTATCCTA ACCTGATATA TCGGATTCAA 180
 CAGACTCTTG ATTTAATACA TTATTACTCT GTAGAAACGA TTTTCGATTT AGCTTTATTG 240
 ACACTTCTGA AAGGTGATTT ATCAATCGAT GGTCAATGTCG TTTTGGATTT TAAAGCCCCT 300
 ATAGCAACAA GTGCTTCAAT ATGGGAACT ATTAAACTA TTGAAGATTT TGATATGATG 360
 TCcAGTTTTt TTGAtAAAtG GCTTATAtG ATCATCATCC TATACCTTTT CGCAACCTTT 420
 TTATCGAGGA TTCAGAACAA CTTAATTCGC CAGATAATTG GTTGATTCTT ACCAAGTTTA 480
 TGCTACCTAA ATGGGnTATA TAAATAAGC AAAGCAACGT GCTGATAACA 530

(2) INFORMATION FOR SEQ ID NO: 994:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 994:

	TATTGGATTn TnCCACCAA AGGCAACGGT TCCTTTTCTA ATTCTAATAA TCAAATCCCC	60
10	ATTAATATCA ATCCCACTCA AATTCTTTCA CATCATGAG TATAAAATAA TATGGTGATA	120
	TACGTTGAAG TTGATTnTCT TATTTAAGTA TACTCCAACA GTATTTTATA AAATTATTTA	180
	GTCATCATCT TCAGTGTTTA ATTCAAAAAT ATAGAATCTA CTGTATTGCA TAATAAATTG	240
15	AAATGCCACA AATTGTGTAC AAGCAATAAT AATGATACTT ACCATTATGA CAAACTTGGA	300
	TTCTACGACA TTGGGCACCA ATATATATTC GATAATTCTT TCAATATATA TAAATGCGCT	360
	CAACGCGGTT AATAATAAAC CGAAATGTGT TTTaGTTTTA CCACCCCAAC GTTTTGtyAC	420
20	TTTAGGTAAT TTTAATAACG TGAACATTCC GCCAATTACT AATAACAAAT AACTAATTAT	480
	GATTGAAATC GTACCCATCA ACCAATAATT GAATAGGTCA TAATGATTTT GTACAGCAAC	540
25	AAAATAATAA A	551

25

(2) INFORMATION FOR SEQ ID NO: 995:

(i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 729 base pairs
30	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 995:

	ATTGCGCCACC ATTAGGTAGT GGAACACGAA CTTTCATCATC TAAACCTTTA CCAATTAACG	60
	CTTTAGCCAT TGGTGATTCA TTTGAAATCT TACCATTAAA TGCATCTGAT TCAGCTGAAC	120
40	CAACGATTTG ATAACCTTCC TCTTCATCAC CTGGTAATTC TACAAACGTT ACTGTTTTAC	180
	CAATTTTAAC AACGTTGTTA TCTCCAGTAT CTTCAATGAT TAATGCATTT CTTAACATAT	240
45	GCTCAATTCT TTGAATATCT TGTTGATGA ATCCTTGTTT ATCTTTTGCT GCATCATACT	300
	CAGAGTTCTC TGATAAGTCA CCAAATGAAC GTGCAACTTT AATTTTCTCT ACAACTTCAG	360
	GACGCTTAAC TGTTTTTAAT TCTTCAAGTT CACGCTCTAA TTTTCAAAA CCTTCTTGAG	420
50	TCATTGGATA TTGCTTTTGA TTTTCCATAT TGTCACTTC CTTTACTGAA TTATACTATT	480
	GCTTGCTAAC TAAAGACTGA ATTTTTGTTG TCATAATATC TATTGCAACT TTATTGCTCC	540
55	CACCTTCAGG AATAATTATA TCAGCATATT TCTTAGTCGG TTCAATAAAT TGGTCATGCA	600

TTAGTATCTC GTGTTGAAAC GGCCTAATAT nCTCAAGTCT GGCATCTGTA TCAACATATA 720
TnTTnAACA 729

(2) INFORMATION FOR SEQ ID NO: 996:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 325 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 996:

GACAAATATA CTGCTTCTAC TCAAATATTA GTGAACATGA AAAAGTCCTC AAGTGATTTA 60
GCTTTCCCAA AATGTTCAAA GTAGTTTGCA GTCTGTTAAT ACATATACGG AnATTATCAA 120
AAGTCCTAGA TTCTAGATAA AGTATCTAGA GAATTGATGG TCAGTATTCA ACAGCTGAGT 180
GAAATCATT TAAAAGTAAC CATCCAACCA ACTCTCAAAT TATTACTGTA CAGTTACACT 240
GGGATAATCT GATCTGACAA ATGTATAAAT ACTAAGTTTT GCCAGCAGCT AAATAGAGGT 300
GTATGTACGA CTTCCAnCTG TAGnG 325

(2) INFORMATION FOR SEQ ID NO: 997:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 997:

CTTCAGTTGT CGGCAAATCT TCTATATTGG AAATACCama TACATTTAAA AATAAATCAG 60
TAGTAATTAA CTGTTGGCTA CGCTGTTTCAAT TAACCACTTT AGCCTCAACT AGTCCTTTGG 120
CAATCAATGT CTTAACTGCA CCATCTGAAT TGaACTACG AATTAAATTCA ATATCACTTC 180
TTGATAATGG CTGGTTATAA GCAATAATTG gATAGTACTT CCATTGCTGC TTGTGATAAT 240
TTCATTTGTG aCTTTTGTTC AATTGAATTG T 271

(2) INFORMATION FOR SEQ ID NO: 998:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 498 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 998:

5 TTTATGTGCT AATTTGTAAT TGAGATTCAT AATGGTATTT AAATCATTAT GGTGCGGGAG 60
 TTATGGTATT GCCATAATAT GAGAGTGCAT CCACCTTCTAT AAGTAATGCA TATTGCGAGT 120
 GCAGGaATGA ATGAGTGCTT GATTAAAATC CTTATGGGTG GTTGACATAA TTAAAAGAAA 180
 CCACATTTAA AATTTCTTAA TCACAAGCGG TTAAGTAGGT ATAGTTTAGT TTTGAGTAAA 240
 10 TCTTTTAA GGTGTACTTG TGCATTTTAC TTAATTAAAG ACATAAGACA TTTAATCGGC 300
 CTAAAATAAA TAAAAAACTA CCTGTTTAGG TAGTTTTTTT AATGTAATAG ATTAAAACAC 360
 15 TAGTTCATTT CTTATTAAAG ATGGATAGTT ATTTTATAGA TAAATTTGTC CTTTAGTGTA 420
 GCGGTAATTT TTAGGACTTT TTGGTGGA AAATGTTCTT AATAAAAGTA ATAGTCCTAC 480
 TTTACCGCCA AAGCCATA 498

(2) INFORMATION FOR SEQ ID NO: 999:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 568 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 999:

30 TGAGCTGGTT GTTTATTATC ATCAGTTTTT TTCTGATTAT CATCTGATTT ATCTTGAGAT 60
 GCCTTGTCAT CCTTATCTTT CGACTCATCC TTTGATGCTT TATCATCATT CTCTTCTTTA 120
 GCAGGACGTT GCTTTGTAGG TGCTTGTTGC TGAGGTGGCA CATATTGATA TTGATTCTGC 180
 35 TGTGGCAATT GCTGATATTG GTTTGCTTGT GAATCTAGCT CATCTGCTC TTTTCTCTCT 240
 TGTGCTTTT TCTTCTCTC TTTTCTGCA TTTCTTTTT GACGCTTTT TTTTGTTC 300
 40 TGCTTTTGT CGTTCAACAT ACGTTCTTTA GCTTTATTCG AATGATCTAC ATATGCAAAA 360
 ATTGCAATA CTAATCCTCC GAGTAATAGC AATACGTCG CAATGCTAAC TACTTTCGCT 420
 GTTGTACTCA TTTTCTTCTT ACTTCCATCT TCATTTTAT TGTAGTCGTA ATTTTGATTC 480
 45 GACATACTCT CCCTCACTTC AAATAATTT AAATATAGAA ATTATCACGC AACCATTAAC 540
 ATTTTCACAA CTAATTGCG ACCAGTA 568

(2) INFORMATION FOR SEQ ID NO: 1000:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1441 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1000:

5	AGTGTCTTCTA ATTGCAAAAT TCCAAATACA GTGTCGTTAT TTACTTTTTT CATGCTAAAC	60
	ATCTCCCAAT TAATAATCAC ATATTTAGCT ACTATACAAG TATTAACAAA CATCCCAACA	120
	ATCCCAAAAC TTGTATGTGC TAACTTATGC ATAAAACTAC AAGTGTAAT AAACCTGCTT	180
10	TCTATTCAAT TGTCAAAGTT GAATAAAATT AAATAAGTAT AAAACCAAT AATCAATAGA	240
	GTTAATATTT AGAAATAAGA TAATAGTACA TAAACAAAT CACTCTATTA CATATTAATT	300
	TTATAATTTA TATACAAAGA AAATATATCA ATATTACCAT TAAATTGCAA CAATTCTACA	360
15	TATATTTCTT CAAAAAGAT ACGCATTGAA TTTTATTAC ATATAACTCT CATGAAAATT	420
	GATTAAACTA AATTAATTAG TTAGCCTTAA ATTTAACAAA TTAATCAA CCGAATGAGG	480
	CTAAATTGGT TATTTTTCAA GATATGATT GTATAAATCA AACATTCTCT TACAAATCAA	540
20	ATGTGTATAA AATGTGATAT ACATTTTCAA CCTTAAAAA GTTCTAAAAA AAGATGAGCA	600
	TCTACTGTCT CATCTTTCAG TTCTTTTTTA CGGGTCTGTT TTCTAATTG AGCACAATCT	660
	TCGATTTCTT TATCTAAATG ACTACCAATT AAATCTATTT CTTCTATTGT TAAATCGCTA	720
25	TCTCCATCTT CTTTTATCTC TGGTATTATT TTTTCTTCAA CTAAGTCACG ATATAGTGTT	780
	TTTGAATTCT CGTTCAATTT CGATTCGTGA TTTTGAATAC TTTCTCCGC ACCAATGTAT	840
30	ATCTATTGGC ACTAGCTTCT ACITTTGTAC CATCAATAAA AATTGAATTA TCATCAATAA	900
	GATTTTGCTT TAAACATTGA CTATAGAAGT GAATAAATAA AGATTCAATT AACGCATCAG	960
	TATTAGGATT CACTCTAAAA CGATTAATAG TTTTATAAGA AAGTGTTTGA TCTTGGGCTA	1020
35	ACCACATCAT TCGAATACTG TCATGAAGTA ATTTCTCTAT TCTACGACCA GAAAATACAG	1080
	ATTGAGTATA TGCATATAAG ATGATTTTTA ACATCATTTT TGGATGATAG GATGTTGCAC	1140
	CACGATGATG TCTGAATTCA TCGAATTCGC TATCAGGTAT CGTTTCAACA ATTTATTAA	1200
40	CATATCGCGA AATATCATTG TGAGGAATTC TAACGAAGTT TCTATTGGTA GTGTAAGTTG	1260
	GGGTCATGTT ATAATTTTTA TACATAAGGC ACCTCTTTAA TTTAGTTTAT TAGTATTTAT	1320
	TAAATTATAC GaAGGGACCC AACACAGAAA ATTCATTTTA TTGAATTTA CATTTATGTG	1380
45	CAAGTTGGGG AAAAATGTCT TATTTTTTCA AAGTATTTAA AAGTAAATT ACATGTTAAT	1440
	A	1441

(2) INFORMATION FOR SEQ ID NO: 1001:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 587 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1001:

	TAATAATATC TGTTCATATTA AGTCTCCCTC ATTAGAACTC ATTATAAATG AAGTTATTGT	60
10	GTGTGTCGCC ACTGCCATAA ATTAAATATA AAGTAATAAA TATTGCCAAA TACAATAGTG	120
	TTAATAAATA TGGTTTGAAT GCTTCAACAT ATTTATTAGG TGGCTGTTTA CTTTITAGATT	180
	TCATATTGCA CCTCTTAAAG TTCTTAGTAA AAACGCCCTT ATAAAGACCG TTCAATATAA	240
15	AATACGTTTT AAAATTTGTT TTTTACAATT CATTATATCG ATATTCATAA TGAAATTCAA	300
	TTTTAATTTT ATAGATTCAA CATAGTAATT GGTGTGTCATC ACTCAAKTAT TGTAATGAT	360
	ACACTTTGTT GAGACATCA TTCATTTTAA AGGTTATTTA AACAATAAAC AATTACAGTC	420
20	TATATAACAA TTTTGTATA TACGTCAAAA TCAAATAAAC TCATCACATT AATATGACGA	480
	GTTTATAATG TTATTGAATT ATCATCAGCG CAAATATATA CATTCGCAAG TCAAGCATAA	540
	CATATTTAAC AATTGCTTTG CTTGGTTTTA CCAATGGATT AAAAACC	587

25

(2) INFORMATION FOR SEQ ID NO: 1002:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 522 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1002:

	CAATTGTGTC ATAATTTTGC TGATGATAAC TAATCGTTTT GGAGTAGCTA TTTGATGAAG	60
	TTCGTCCAAT ACCTTTTGTT GAATCGAATT TACCACCTGA GCGTATATGA AAATGTACTG	120
40	TCTACTTTTG CAGTTGAAAT TTTATTTTTC GGCAATTGGT CTAATATTTT AGTTTTACGA	180
	TTTCTTTTTC CTGAAAATC TACATGGTAC TCACTTGGAT ACTTTAACCA ATTTGAATTT	240
	TTTTCTTCTT TATGAGATTC AAACTTTAAA TTTGAATGAA TTGAGCCTTG TTTTTTAACA	300
45	AGTAATACAT TCTTGTGATA AGTTGGATCA TCGATAAAGT CGAATTGTAA ATTTTGGAGT	360
	ATATTGTWT TCTCATCATA TACTGTTTCA GTTCGTTTTG TGATTTTACC GTTTTCCCA	420
	ATATCATCCG GTGCTGTTGA ATTTTATCT TTATTAGTAA CATTACGTTG CTTTTGTGAn	480
50	CTTATCAACA TGTCTTCTAT TTGGCTGAnG CTTATGGnGA TT	522

(2) INFORMATION FOR SEQ ID NO: 1003:

55

- (A) LENGTH: 323 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1003:

10 TCTTTGTTTG AGACATAATG TCATGTAAAG TTTTCACAAT GCGTTCTTCA ACTTGTTGCC 60
 TTGCCTCTCC TTGAAAGGA ACAATTCGAT CTCCGAATAA GTCCTCAGGC TTGTATAGAT 120
 TATCAAATAG ATAGACTGAC TCACCTTCAA ATATTCCAAA ATGCCATTCC TTCAGTCCCT 180
 15 TAAAACGTTG ATACGATTGG WTAGGTGCAA CATTTTCAAG TGTGTCACTT GCGCGTTCTT 240
 GCGTTGATGA TGCATATAAA TCGAAGTTTA TCCCCTTAGT TTCATAATAA CTACGTGCCT 300
 TTnGAGCTTG AGCAATTCCA AGT 323

20

(2) INFORMATION FOR SEQ ID NO: 1004:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 320 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1004:

30 TTCTAACTTG ACGTGACTGG GCATTTCTGT GnAAATnTTC TGACATTCTT TTTAATATCC 60
 GCAACGCAAT TGCTTCAGTT AAAGATTGnA ACTTGCAACA TTATTTGGTA CGAAAAAGTC 120
 35 AGTTTAAATG TCGATATAAG ATTTATTTTT TTTGTTATAT AGTTTCGCAA CTACATTTGG 180
 TTGTCTTACT TGATCATATT TTGCAACCGT ATCGAATGCC GTCTTTTCAA ACAGCTTTAC 240
 GAGATACGTA AACCATGACC CATCCATCGA AGTCCTTGGT AAAAACCCAG GGTTCGGAA 300
 40 GCCGTAGGGT TTGAAAGATA 320

(2) INFORMATION FOR SEQ ID NO: 1005:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1005:

nATTAATGCT AGCCAAAACA GAnTTGCAAn ATGTAAAGAC ATGTGAAACC CTCCTTATAT 60

55

CGATATTTAT AAAAAATAAT ATTTCACTTA ACCAGTTTTT AGTTATCATT GCAATGGTAA 180
 AGGATACGAT GAGTATCACA CCACAAATGA TAATACCAGG CAGGAGCCAA CATAAATCAT 240
 5 CTAAATCTTT ATTATATGTG ATTAAAATAT TAAAGATAAC AAAAGTGGTA GTAATAACTA 300
 TATTG 305

(2) INFORMATION FOR SEQ ID NO: 1006:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 334 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1006:

TTCAAGATAT CAGTATAAAT AAAAAGTTAA TATGCTCATT ATAGACAGCT ATAAGTATAT 60
 TTTCTCGTCC CACTCTATAG AAATCAATTT CATTTATTAC CATTTTCTCT ATTTAAAAAC 120
 ATATTTTTTG TATTAACAAT TAAAAGTGGG TATATATATT TAATGAACCA ATTTTTTAGG 180
 25 AAAGTAGGGC TAAACCTATG GAAAATAAAT TATACCACAT GGTGTCCTCT TTTACCATGA 240
 ACACAGTGAA TTAAAAATA TTAACAAGG TATTGGGGAA GTACACCAG CATTAGGTCC 300
 AATTGGTAGC CACCTTCAA TCCATTGGGG GAAA 334

(2) INFORMATION FOR SEQ ID NO: 1007:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 565 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1007:

AGAAAATTG TTATATTAGA AGCGAATAAT TTTAAACAT ATTATGAAAT GTAAGAAGAA 60
 ATATTATGAA ACTAGGAAAA GTAGATTGTT GAAAATTCCA TACAATAGGn GAAATAAAGT 120
 45 GAATAAATTC TGGTATTACT AAAGCTATTG ATTGTCTTAA TTGTAATAAA AAGTTGGATA 180
 GTGTAGTTGA GGGGAGAAAA ACTAAATGC AAAATATAT TCTCAATGGA TCATTATAAA 240
 TACCAATTAG aTTAAAGTAT TGCGaACATT TACAATTAGa AATCAGTATT AATCTGCCAT 300
 50 AGATAATTGT ATTAATGTAA AAmATAAAAT AGAATGAGAA TATACGGAAA TTTTTTATAC 360
 ATTTATAACC AATTCTGTTA AACTGTGAAG AGTAGAGTTA AAAACGAACT TTTTGTATT 420

AATGATTGAA CGCTATTCTA GAGAAGAAAT GTCTAATATT TGGACTGATC AAAATCGCTA 540
TGAAGCCATG GTTAGAAGTG GAAAT 565

5 (2) INFORMATION FOR SEQ ID NO: 1008:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 400 base pairs
(B) TYPE: nucleic acid
10 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1008:

CAACATTAGC GTCTAAACAA ATCTTTGACT AAACGTTTAC TTGAGCGACC ATCTTGATAT 60
TTAAATGTT TATCTAAGAA TGGCACAACCT TTTTCAACCT CATAATCTTC ATTGTCCAAA 120
GCATCCATTA ATGCATCAAA GGACTGTACA ATTTTACCTG GAACAAATGA TTCAAATGGT 180
20 TCATAGAAAT CACGCGTCGT AATGTAATCT TCTAAGTCAA ATGCATAGAA AATCATCGGC 240
TTTTTAAATA CTGCATATTC ATATATTAAA GATGAATAAT CACTAATCAA CAAGTCTGTA 300
25 ACAAAGAGAA TATCGTTAAC TTCACGATGA TCTGACACAT CGATAAAGTA TyGTsTATGT 360
TCACGTGAAA TAYyAAGTCy ATTtTTTACG AaCGGATGCA 400

(2) INFORMATION FOR SEQ ID NO: 1009:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 454 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1009:

TATTGTAAAA ACTGTGGTTC TTACAATGGC GAAGAAGTAG CAGCTAAATA ATTTTAGTTA 60
40 CTCAATATAA AAaGTCCCGC TTAAATGAT TGTTTTAAGT GGGACTTTTT ATATTGCGAA 120
AAATAATTGG cGAACGAGGT AACTGGATAC CTCATCCGCC AATTAAAATT TGTTAATTTA 180
45 ATAATTAAAT ATAAAGACGA TTTATTAGTT TTTACGTTTT CTAGGTAATA CGAATGCAAC 240
GATGCTACTT AAAGCTAATA ATGCCATTAA TGGAATGTC ATATCTnTAT TTGATTCTTC 300
ACCAGTTTGT GGTAATGATT TTGCTTTATT TTCTTGTA TTTGTATTGT TTTGGCTTTG 360
50 GAGTGTGTCC ATCATTGTG TTTTAAATGT TTGCTTnTnG TAATGGAGCA CTAATCTTTT 420
GCCTCGCTAG AACCTGCCGG AGGTTTGGAA CCAC 454

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1010:

GGTACAACAA CTTTAGACAA TGAAATATTA AAAGATACGG ATGATAAAAA GTCGAGTAAA 60
 ACTTTTGTG GCGGAACAAA AGTTGATGAC CAACATGCTA GTATCGGAAT GGATTTTGAA 120
 AATCAGGACA AAACCTTAAC TGCCAAAAAA TCATATTTCA TATTAAACGA TAAAATTGTC 180
 TTCTTAGGAA CTGGCATTAA AAGTACTGAT TCATCAAAGG ATCCCAGTTA CCAACCGATG 240
 AAAATCCGCA AGCCGATGGG GAATACGGTA TTTTCCGGCC GTTAACCAAC CACCCATTCC 300
 GGTTATCCGG GAAACCATTC CATCCTTTTA GGAGnChCCG GtnCCnAAAA GGACCTCGGG 360
 TATCCTTTTT TTAACCAACC CGAAATTACT GGTAAAAAAG 400

(2) INFORMATION FOR SEQ ID NO: 1011:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1011:

TTAGGTATTC GCAAAAATGA TCTTCAATAT ATAGCTATGG CCAAAAGATG GGCTAAAGCT 60
 TATACAATTA CTGTAGCAGT GGGAGTTGTT ACAGGTACAA TTATAGGACT TCAATTATCA 120
 TGATTTGGCC TACATTTATG GAAATGGGTG GACACGTATT GCACTTCCCT TATTATGGGA 180
 ACATTGCCGT CTCTTGAGC TATTTCTTAA GTATATATTA TATACTGGGA TCGTTTAAAA 240
 TAATGGCACA TTCTTAATAG TATACCAGTA ATTATGGGGC CTTCTCAGCA TCTCATACTC 300
 AGGAATCATT ATGAATACGC CTGCAGGTTT GAGTGAAGAT GGAAGTGGCC ATGTCACCTT 360
 AGAGCGTGTT ACCCACGTTA AGTCGTCATT nCGAATACAC 400

(2) INFORMATION FOR SEQ ID NO: 1012:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1012:

CTTTATGGGA TACACTACAA ATCGAGACTA TAAGGTTTTT TATTTTATTT ATTATTACAT 60
 5 TATCAATAGT TTTATAATCG AGCTTCAAAA CTTTAGAAAA TAGTAGAAAT AGCATTCAAT 120
 ATAGTGCAAA AGTGCAAATT GATAACTTGA CACTTATCTC CTATAAACCG TACAATTAAT 180
 TTGTATGATT TATATATAAT TTCATAAAGT CATATTGAAT TTCATATAAA GAGCAAACCC 240
 10 TAGAAAAAGA GGTGTTGTGG AATTATTAAG CTCCTACTCA TTTGCGGGGT CATCTAGGGT 300
 GTGGAGGATG TGTGCCGAAT TGTGCTCAAA NATCCCAGCA GCAACGCGAG TTGAACAGCC 360
 TAAGACGCAA CAAAGGATCT GnnAAGGACC ATCAGGCACG 400

(2) INFORMATION FOR SEQ ID NO: 1013:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 353 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1013:

GTCGTTGCTG ATACCACGCC AGTATTGGTT TTAAATGATA AAAAAATCTG TGAGATTGCA 60
 TCATTGAGAC AAACATCGCT ATTTGAAATG GCCGAATATA TAGGGTTTAG CGAGCCACAG 120
 30 AAATTAGTAC AATTATTTAT TAACCATGAT AGGAAGGTGA GACGCCAATG AATCAATATA 180
 ATACTATAGG TTTTCACCCG GGAAATAGTC GTATTCATCA ATTAAATGCG ACTGTAAAC 240
 TTTTATnCTT ATTAGTTGTT CTATTCTGCA ATGGTnCTAT GCACAGGTAT TnATTTAATT 300
 35 AGGCTCACAT TTATGGGCAA TTCTCAATGA TGAACAGTCG TTGTGTAATC TCG 353

(2) INFORMATION FOR SEQ ID NO: 1014:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 627 base pairs
 (B) TYPE: nucleic acid
 40 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1014:

CAAAGCCTTC AACAAATAAT CTTTATTAG TACTGATTGG TTGCACAACA TAAGTTACAC 60
 CAGGTTTCATC TTAAATGTA ACTTCCTTAT AGAAGAGACC TTTTITAGCA CTATACTCTG 120
 50 TTTTITCTGA TTTAATTTTA TCTTTTAGAT TTTTCTCTC CAAGTAACTG TCAATTAGTT 180

EP 0 786 519 A2

ATGAAATAAA TAATATAATG ATAACCTTTAA TCGCGTTTTT TTTCTTCTTC ATTATGTAAC 300
 TCCTTTTTGC TTATCTATTG AGTATGAACG AATCAATTTA TTTAATAAAG AGTTTTTACC 360
 5 AGATTGTAAC AAATTTTTTT ATTCTCAATA TTATTATAGA AATCGAATTA CATACACGAT 420
 GCTTTTAAAA AATTTAATAA CAAATTTTAT TTTACTATCT ATCAAAATAA TTGACTACTT 480
 TTAACAATTA AGGATGATAC AATTACATTT TTTTCAATTT AAAAAATCAG TCATATCAGT 540
 10 AATTACCGAT GAAAATACTA AGATATCGAT ATATTTTGCA TTATTTAACA TTAAATTCn 600
 ATTATAGACA ATTCAATAAT TGCGATG 627

(2) INFORMATION FOR SEQ ID NO: 1015:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 462 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1015:

25 TAGTGTGACA TTAACACTGG AACAAAAAGA AGCTGCAATT GCAGAAGTTA ATAAGCTTAA 60
 ACAACAAGCA ATTGATCATG TTAACAATGC ACCTGATGTT CATTCAAGTTG AAGAAATTCA 120
 ACAACAAGAA CAAGCGCATA TTGAACAATT TAATCCAGAA CAATTTACGA TTGAACAAGC 180
 30 AAAATCAAAT GCAATTAAAT CGATTGAAGA TGCAATTCAG CATATGATTG ATGAAATCAA 240
 AGCTCGTACT GATCTAACAG ATAAAGAGAA GCAAGAAGCT ATTGCTAAGT TAAATCAATT 300
 AAAAGAACAA GCAATTCAGG CGATTCAACG TGCGCAAAGC ATCGATGAAA TAAGTGAGCA 360
 35 ATTGGAACAA TTAAAGCTC AAATGAAAGC AGCTAATCCA ACAGCAAAAG AACTAGCTAA 420
 ACGCAAGCAA GAAGCTATTA GTAGAATTAA AGACTTTTCA AA 462

(2) INFORMATION FOR SEQ ID NO: 1016:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 235 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1016:

50 TGATTGGTCC CATCGACTCG TTATAACAAA TTGAAGGTAC AAAAGTATCT TGTGATTTAA 60
 TAAATTCAG TAAATTCGTA CCTGGTTCTA CAAGGTAATT CTTTCCATC AAGTGTAACC 120

ATAAATTGCG CATATCCACC nAAAATAACG ATGCCCCACA CATCTTTCAG ATAGG

235

(2) INFORMATION FOR SEQ ID NO: 1017:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 227 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1017:

GTTGAAACT ATAAGATCA ATAAGGTAA AACAGTGGTT TGTGTCATTA GTGGTGGTAA 60
 TAATGATATT AATCGAATGA AAGAAATTGA AGAACGTTCA TTACTATACG AGGAAATGAA 120
 GCATTACTTT ATCTTAAATT TCCCTCAACG TCCAGGTGCA TTGAGAGAAT TTGTAAATGA 180
 CGTATTAGGA CCTCAAGACG ATATTACTAA ATTTGAATAC TTAAAT 227

(2) INFORMATION FOR SEQ ID NO: 1018:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 284 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1018:

GTCTTACTTG CACAAGGTAT TATTTCTATC TTA CTGTTG CTTTCGCAAT CATGCTATAT 60
 ATCATTAAATA TTTTAGATGC ATATCGTAAT GCTGAACGAT TTAATCGCAA TGAGGAAATA 120
 AAGGATCCGA AGCGCGTATG GTGGCACATG GGACAAGACG TTCCCATACT TACTAATCTC 180
 ACCAGGTACA TnCTTATTGA TATTGTAGTG TAATTCCATT ATATTATGTT GGAGTAGCAT 240
 TACAATACAA TTATACACGC GCCTCGAGAC nACATAGATG GGTG 284

(2) INFORMATION FOR SEQ ID NO: 1019:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 779 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1019:

CAATCATCAC ACATAAATGT TCGAATAGGA TTGTTTTTAA GTCGTTTAGA CTCAGTTGTG 60

CACCTCAATA TGAGTTACAT GGTGATTTCG ATATGAATAA CCATCTTCTT GATTGTAAAT 180
 AAAACTATCT ACACCATTAT CGCTGTAAAG TCGTTTACCG TCTTTTGCAA ATTGGAAAAA 240
 5 TAAATAGGGT AATAGATCTA TCGGTATATC TAAATGATCA TGCTCATTG ATAATCGAAT 300
 AGTAGTTGCA GAGTCATGTG GTTCGGAATG TTTGAAAAAT GGTGTCATAT TAATGACAAA 360
 TGAACCTTCT AACATGGCAC GTTTTTTATA TTTTATTTCT GAATTTAAAG TAGGCGGATT 420
 10 AGTTTGTCCt TctAGGATAG CaCGATTCCa TTCatGaTTA TCTTCAAAGT CGATTGGTTT 480
 TGAACCATCa AATACACCTT TTTCTAAATC TTCGATGCTA ACTTTTCTAT CATCGAAAAT 540
 CCAAGTCGTA CTATCTAATG TTATAGGAAA CTTTACGGCT CCTTTAATTT GTATCATTTT 600
 15 CCCACTCCCT ATCAATGTAT ATAGCATTAT TTTAACACAA ATTGCTAACG ACACATTTTA 660
 AATCTACTTG CTTTTTAATA TTTAGTAAGA TAAACTTTTA GTAAGACTTG AGAATTTATA 720
 20 TAGAGGGGGA GCGTGTCATG GCGAAACAAG CAACAATGAA AAATGCAGCT TTGAAACAA 779

(2) INFORMATION FOR SEQ ID NO: 1020:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 950 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1020:

TTAATGGTCC ATTAAGAATA ATGGCAGAAA ACCGTTTGCA TAAATCAAAC CTTGATTCAC 60
 TAACTTAAGT AATACTCGCT TTTcNTAAAT CCTTTAGAAA GGAACTATnT ATCATGATTA 120
 35 CTTCAATTCAG ACATTCTGAA GATATAGATA AACATATTAT AAAAACACCA TTAGATCATA 180
 CAGCGTCATG GATTAATGTA GTAGAnCCAG ACCGAGAAGA AATTGAAAAT CTTATGGAAC 240
 40 AATATAATAT ACCTGAAGAC TTTATACGTG ACCCTTTGGA CTCAGAAGAA AGTTCCCGTA 300
 TTGAATATGA CGAAGATACT GGTTACTCAT TAATCAITAT TGATTTACCT ATCGTCAATT 360
 CAACTAATCG TAGCGTTCTA TCTTTTGTAa CGATTCCATT AGGCATTATT ATTGGCAACG 420
 45 GTATTATCGT AACAGTTTGT GACGCTGAAA ATGAATTTTT AGAAAATTTA CCTAAGCGTG 480
 ATATTAAATTT AAAATTTTAC AGCAGATTG CATTAGAAAT TTAACTACT ATTGCCGACC 540
 ATTATAATCG TAACTTACGA TTACTIONA AAAGTAGAAT TCGTATTGAG AAAGAACTAA 600
 50 AAAATAACAT TACTAACAAA CAACTTTTCA AATTAATGGA AGTTGAAAAA AGTTTAGTAT 660
 ACTTTTTAGC TGCCTtAAAA GGTAACGATA CAATTATTAA AAAGTTATTC CGTTTACCTG 720

AAGCCATCGA AACAAACAGAA TTACATCAAC GTATCTTAGA AAGTATCACA ACATCATACG 840
 CTTCTTATT ATCTAATGAT ATGAATACGA TTATGAAGAC ATTAACACTT TTCACGGTAC 900
 5 TATTAACgTT ACCAAyActc GTATTTAGTT TCTTCGGTAT GAATGTGTCG 950

(2) INFORMATION FOR SEQ ID NO: 1021:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 528 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1021:

AAATATATAAA ATGAATACAT TGTGAGGCCAA TTATCAAATG AAATTTAATA AAGTAAAACT 60
 20 AGTTATACAT GCGTGTGTAC TATTATTTAT CATTATTTCT ATAGCGTTAA TTTTTCATCG 120
 ATTACAGACG AAGACACATT CTATAGACCC AATACATAAG GAAACAAAAT TATCAGACAA 180
 TGAAAAATAT TTAGTGGATC GTAATAAGGA AAAGGTTGCG CCGTCTAAAC TAAAAGAGGT 240
 25 ATATAATAGC AAGGATCCTA AATATAAGAA AATTGACAAG TATTTACAAA GTTCATTATT 300
 TAACGGTTCA GTAGCTATAT ATGAAAATGG CAAATTGAAA aTGAGTAAaG GTtATGGaTA 360
 TCAAGaTTTT GAAAAAGGTa TTAAAAaCAC ACCGAaTACG aTGTTTTtAA TnGGTTCAGC 420
 30 TCAAAAATTT TCAACAGGGT TACTGTTAAA ACAGTTAGAA GAAGAACATA AAATAAATAT 480
 CAATGATCCA GTAAGTAAAT ACCTTCCATG GTTTAAACA TCTAAGCC 528

(2) INFORMATION FOR SEQ ID NO: 1022:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 713 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1022:

GCATTnCGGT TAAAgTAKGt TTCAGCTTCT CTCaTTTTAG AAGTACCGaA AAtTGGTTGG 60
 45 TTAtCTGGAT TTAATrTcAC GATAAATATT GCTTACTTTG TTACTTTGTt AAATGTACCC 120
 ATTACGAGTC TCAATGTTAT TCCAATAAAT ATCACTTGTT GGTGCATGGT TTGGATATGC 180
 50 ACAATGATTT CTTGAAATTG TTTGAACGAT TTCTAATGGA TCACAGCCAA ATGTACTGTT 240
 TAATACTTCA GAGTCTCTGA ATAATGCACA AAntGAAATA CAAGTTGTCC AGTTTGGTAA 300

AGTATCTTGC GTATAACCAG CCTCTATACG AACCATTTTA AATTTTGTTT GAATTAAATC 420
 TGTAAGACTC TGTCTATCCA TTCTGTTATC TACCTTTCTG TTTGGGGAAT TTTATCCGGA 480
 5 CACAAGAAAT TGCAATAATA CACATTTCTT GAAACACAGA TTACATCTTA ATATATTTTT 540
 AATAAAATGA AAAGAGTCAA TTTCACATTT GTATTAAATT TTGATCAAGT CAACAAGAGT 600
 TAAGTTTACT TTATAATGAA GCGAAAGTAA ATGAAGTAAA ATTTTATGCA TAATCACTAG 660
 10 ATTTGATAAA ACTTACACTT TTATAATGAT ATCAAATTTA CAAAGGAAAA CTG 713

(2) INFORMATION FOR SEQ ID NO: 1023:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 570 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1023:

AATAAAATAG AAATGATTAT GGGTACAATT ATCGCTATCA TATCCTTGTT ATTAAATTATA 60
 25 TTACAAGCAT TTAATATTAC TTGGGGCGTT ATACCAATTA CAAATTTTGG ACATCAATTT 120
 TTCTTTTTCA TTGGTATTAT TTTAGTAATT GCCGGCATAT TTTwAAGCG ACTtGAGTTT 180
 TCGGGaATcG GGTtATTATT TTGTcAAAAA ACCGTCGATG CAATGATTCA TAATCCACAA 240
 30 TCAGCCCAGA TTTTTTCATT AATTATATGG ATATTATTAG TAGTCTAGT TATATATTTT 300
 ACAATTAGAT TATCTTCACG TACAAGATTA TAAATATGAT AAAACTATTC ACTTGaTTAA 360
 TTGTATTAAT TGaGTGaAT AGTTTTTTTA TTGTTGGAAT AACTTTTGGT AATTTATAAA 420
 35 TAATTTAAAA AAATTGTITA TAAAAGGAAG CGTATATAGA ATGAAGGTTG GGTATATAGT 480
 TTATTGAGGG AGGTGTCACA ATGAATAAAG TCACAaTTAA TCCTCAAATC CAATTAACCTT 540
 40 ATCAAATTGA AGGTAAAGGG GATCCTATAN 570

(2) INFORMATION FOR SEQ ID NO: 1024:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 487 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1024:

CTTTGTTGTG CTTGAGCATT ATTAAAAAGA GTATTTAAAT TTAGTTGTTG GTTTGTAATA 60

55

TCAGAATTAT TAACACTATA TTTCCCTTTA AATAATGGTG ATTcGcAATA ATGCTTATCT 180
 TTATCTGCAG CTAActGGAA TTKCCCTAAT GCAGAGTCTG CGATTGTTGG TTCAAGATTA 240
 5 ATCATTGATT TCTCTTTTTT AGGATCATGT CCATATGACA tGTaATTTTC GATGCATTAA 300
 CAACAGATTT AkGAATGCCA AGCCCTTTAA CAATTTCATC TGATGCATCT GCGCTTAATn 360
 CTAATGAAGA TAAAAACGAt TAAtCTTTCAT CkKtTCTTGG AACTTCACTT CATTTTCAAA 420
 10 ACGGTCATTA AAATAATCTT TATACATTTT TGCTGTTTGT TGTTCACTTT TTAGGTATGT 480
 ATTTTTC 487

15 (2) INFORMATION FOR SEQ ID NO: 1025:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 364 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1025:

25 ATATGTTATA AACaTTTTA CAGTTGTATC ATCTACTCGT TCTATTCTCA TATCATCTCA 60
 CTCCTTACAG TCGATGAATA GTATCCACAT TGTATTAAAG CATCAACAA AATACAATTT 120
 ATTTGTTTGA TTTCTTTCTC TAAAmTATTT TGgTAACTTA TAATAACATA TTCGTTGTAG 180
 30 CAAAACTATT TAAATGATTT ATACAGTTAA AGCATTAAAG CACTTTCTTA ATTTAATTTT 240
 ACCTTATCAT AGCAaATGTT TCATCCAAAA CACTCTAATA TAGAGCTATA TTAACAATTG 300
 35 CACATTTGAt TTAAAGATA CTAAGCAATT ATTGtAAAAT TAAAACCAA AAaGCAACAC 360
 TATA 364

(2) INFORMATION FOR SEQ ID NO: 1026:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 333 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1026:

50 TGCCCCAGGG GAGTA(nAAA AAAAAGnGGG GTGGGGAATT GGAAACCGGA ATTAGTGGTT 60
 AAAGGGGTTT ATTGGATGTT GGGTCTAAAG TATTCCAGA AATTTATCCA TcNATGATTG 120
 CTTTAGCCAA GTTTTTTTCC AACCCATCTT TATGGAATTT AGTTATCTCA ATTACATTAA 180

55

CTTATATAAA AATGGCACAA TTATTTTCATG TACCAACAAG AACAAATATTA ATACGTCATT 300
 TAACACCTAA AATTATACCG GCTATTATCG TTG 333

5 (2) INFORMATION FOR SEQ ID NO: 1027:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 450 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1027:

TGTTCCATTC GTTTTAATAA TGGTGATACT GTACCAGTAT CGAGTGCTAA TTCAGTTACG 60
 ACTTTCTTGA CGTTTACAGG AGATTCATCC CATAAAATTG TTAAGACAAG AAATTGTGGG 120
 20 TATGTTAGAT TGTACTTCTT AAAAAGTTT TTAGAGTAGT AGCGATTAAC TTGTCTTTGA 180
 GCATTGTACA AACTAAAGCA TAGCTGTTCT TTAAATTAT GTTGATCAGA CATTAAAGTT 240
 CTCCTCCAGA CATACTATCC GTTTTTTCT CTTTTCGGAT TGGTAATCAT TAAAAAGTTG 300
 25 ATTGTTTATT AATTACAAC TTTCTTTGAT TCAATGCCAT GCAAAAATTA AAGTATGTTT 360
 AAAGTTTAGA AGATATTTTT GATTAAATCA AGCAAAAAGA TAATTTAATA TATATGTGAT 420
 CATTTTTAAA AATAACTGTn ATAGAAAAGA 450

30 (2) INFORMATION FOR SEQ ID NO: 1028:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 535 base pairs
 35 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1028:

TGCATCTATA CCTAGAACTT TAATTTGCGA TTTTATTGCG TCTGTAATAT CTTTAATACC 60
 ACTTTGATA GGGAACTTT CGTCACCTTC AACATATTTA ATACCAAAGC CACCACCAAG 120
 45 GTTTAATAAT TCAACTTGAA TGCCTTGctc TTTaAGCCAA CGTAAAmCAA TTTTAGCaGT 180
 TTCAATAAAc GCTTCTGTAC CTTCAAtCTG TGrACCaATA TGACAATGTA CACCTTTTAA 240
 TTTTAAGTGT TTAGATTGTT GGACTTTGTC AATTGCTTTT TTAGCTAAGC CATATTGaAT 300
 50 TGaTAATCCA AACTTACTAT CTTCTTGCCC AGTTTGaATA AATTCgTGTG TATGTGCTTC 360
 aACACCTGGa TTAACCTGTA AlACAACCTG aACCGTATCA TTAGCATAGC GGTCTATTAA 420

55

AATTTTCATGT TTCGTTTTAT TGTTACCATG GAAATGGATG CGACTCGGTT CAAAA

535

(2) INFORMATION FOR SEQ ID NO: 1029:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 359 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1029:

GAAGATGATT GAATTATGGA ATTTTTTTTCG AGATATGTTA TTTCGGGGAT CAGATTTTGC	60
GATTTTGCCT TGGATAGAGT AGTTGCGATT TGTAATACG TTTTTCAGGT TCGATTCAAT	120
TTCGTTGTTA AGATAATATG GATTAGTTTC ACTGACATGT TGATGCATGT CTTTTTTGTT	180
CTGATAAAGT ATAAATGTT TACTTAACAT TACTTAGTAT GACACCTGCT AATCAAACA	240
TTATTTGAGA CATTCTTTTC AAATTAATTA TAAATTTTAA CCLATAGACT AGTTTGATAT	300
TTATCTACAT CTCAAAATTC TCATCAACAA TCTGTCACAT CCAACATTTT TACTThAGT	359

(2) INFORMATION FOR SEQ ID NO: 1030:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 341 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1030:

TACATTGGTA AATAATTCTC ATAGTTTATG TCTCGTTGAT TAATCCAATC ATTGTTTTTA	60
TCGGGTAAAA CATTCAAATC TTTCAAATTC ACAATTGATT GATGTTTCAT TAAAGTATCT	120
AATTTTTGCT CTCTAGTTAA ATATTCACCA ATGTCATTAT AATAGATATC GTGATTGTAA	180
CTTCCATCTT TAACTAAAAT AGATATAGCT ATCGGTGCTC TACTTCCTGA TCCAAATATT	240
TTTCCACCTT CTTTTCTTga TTTTTCACCC TGTGTTTCGTT GGTCGCCTCT CAAATTAATT	300
ATGtACAGat AGTTAAATTC ATCGaTAAAA CTTTTCTAAA T	341

(2) INFORMATION FOR SEQ ID NO: 1031:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 326 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1031:

5 GTCTAACTTT AATGTTTAAA AAnCTTAGAA ATTTTCTACT CATAAATGGA TGCTTTATGG 60
 AATATTATTA ACAGGCCCCAG CTTCAATGCT AGCTATAGAA TTTGGATGGT TCTTAACAGA 120
 GATGGGTAGA CAGCCTTGGA TTGTTCTGGT TTATATGCCG GTGGCAGAAG CAGCAACACA 180
 10 GCAGCGGAAT AACCTTCGTT ACAATTTTAT TTGGCATATG TACATCATT AAGTATACAG 240
 TGCAACGTAT ATCGTAGTTA AAATAACCGC GTAGAGAGTA ACGTAGCAGA GCAGAGGAAT 300
 GAAAnATTAG CTTATGnAAC GGTAGG 326

(2) INFORMATION FOR SEQ ID NO: 1032:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 263 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1032:

25 ACAATCAGGC TTGTGAAGAA TGAATTGCAT GACGTTTGTG TAATTGAACT GCGAGGTTTA 60
 ATTTCTAAAA TGCTCCAGnC CTCCAATGTA ATCGCATCTT GATATAACAT AGCGATCGCT 120
 TGATTTGTCTG TTGTATCAAC ACCTCTAAAC AACGTTTCAC CTATTCGGCA ATTCATTTAT 180
 30 TTTACCCAAA TCATTATATA AAAGTTGTGG TAACATGCTC GAAATACCAC CTGAAATAAT 240
 nTCAATCGAT ATCCTATTTT TCT 263

(2) INFORMATION FOR SEQ ID NO: 1033:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 514 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1033:

45 TTTCTATAGA GGATCCTGTA GAGATGCAAA TTCCTGGTAT CGTCCAAATT AATGTGAATG 60
 ATAAAGCTGG CATAAACTAT GTAAATTCGT TTAAAGCTAT TTTAAGATGT GATCCTGATG 120
 TTATTTTAAT AGGTGAAATC AGAGATAAAG ATGTTGCCAA GTGTGTTATA CAGGCTAGTT 180
 50 TAAGTGGTCA CCTTGTTCTG ACTACATTGC ATGCAACTGA TTGTAAAGGT GCTATTTTAA 240
 GGCTATTAGA AATGGGCATT TCTGTACAAG AATTGATACA GGCAACTAAC TTAATTATAA 300

55

AGCAACAACCT CCGATATTTTC TTTTCCCaTA ATCATtCATT ACCATCATCA TTTAAGAACT 420
 TAGAAGATAA ACTTGATGAT ATGACAAAAG CAGGTGTCAT TTGTGAACT ACAATGCATA 480
 5 AATACATTTA AACTACATTC TAAGAAGCGA CAAT 514

(2) INFORMATION FOR SEQ ID NO: 1034:

(i) SEQUENCE CHARACTERISTICS:
 10 (A) LENGTH: 598 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1034:

TTTTTCACGA ACATTTTGT TTTAAATTTT CTTTCTTAAA CGTATACTTT CTGGTGTAAC 60
 20 CTCAACAAGT TCATCATCAT TAATGAATTG TAACGCTTCT TCCAATGTGA GAATACGCGG 120
 TCTATTCATT GTTTGTGTTT GGTCTTTCGT TGCAGAACGT ACGTTAGTTT GATGTTTTGT 180
 TTTAGTGATG TTAACAGTTA AATCATTTTC ACGATTATGT TCACCAACAA TCATACCTTC 240
 25 ATAAACTTCA GTACCAGGTT CCATGAAGTT TACACCTCTA TCTTCAAGTC CCAAAATGGc 300
 ATAAGTACTT GCAGAACCTT GATCCATTGA AATTAATGCA CCATTACGAC GACCGCCAAT 360
 TTGTGCTTTA ATACGTGGTC TAAATTCyTC aAATGTATGG TTAATAATAC CGTAACCTCT 420
 30 TGTCATTGAC ATAAATTCAG TCGTATAACC AATCATACCA CGAGCCGGTA CATTAAAGAT 480
 TAAACGTGTA AGTCCATTAT CAGTTGTAGT CaTATCAACC ATTTCACCTT TaCGTGcACC 540
 TAAkGATTCa ATAACAGCAC CTGcATTTtC TTGTGGCACT TCACATTGCA CACGTTCA 598
 35

(2) INFORMATION FOR SEQ ID NO: 1035:

(i) SEQUENCE CHARACTERISTICS:
 40 (A) LENGTH: 313 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1035:

AAACAGTGCT TCATTATCGG TATTGATATT GATCCACAAG CCGTTGACCT AGGGCGCACA 60
 ATCGTTAACG TCTTAGCACC AAATGAAGAT ATAACAATTA CGGATCAAAA GGTATCTGAA 120
 50 CTAAAGATA TCAAAGATGT GACGCATATC ATATTCAAGT CGACAATTCC TTTAAGTACC 180
 AGCATTTAGA AGGAATTATA TGAnTTAACC AATGAAAATG TCCGTAGTGC CATGCGCCTT 240

55

AAGTGGGCCA TGT

313

(2) INFORMATION FOR SEQ ID NO: 1036:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 235 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1036:

15	ATAATGAGT AGGTTTGTAT TACTTACGAG ATAACATGGT GAATTATTCT GAAGATAGAC	60
	TAATGCGTCC ATTACATTTT GCAATCATTG ATGAGGTTGA CTCAATTTTA AATCnACGAG	120
	GnCGACGCCA TTAATTATTT CTGGTGAAC TAAAAGTCAA CGTCACTTTA TACACAAGCA	180
20	AATGTTTTTG CGAAAATGTT AAAACAGGAC GAAGATTATA AATACGATGA AAAAC	235

(2) INFORMATION FOR SEQ ID NO: 1037:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 263 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1037:

30	nCAATATACC TCCTATCATC ACACAAAATG CGTTGTTTCG TATGATTTTC TCAACATCGT	60
35	AGACTAATGC AATCAGACAC AACACCATGC TCTATATCCA ATATTTGCTT TATTGCCTAT	120
	CAATGAGCGT ACTGCGTGCT TTAAATAAT CATCATCAAT TAATGACTGT ACAGGCACCT	180
	CATGAAAATT ATCATCCGCC AAGTATTGCG CACGATCACT ATATGCTAAA TGCATCGCTT	240
40	GTATCAnATG ATGCAAGTAA TCA	263

(2) INFORMATION FOR SEQ ID NO: 1038:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 218 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1038:

55	ATTGTACCTG GCACAATCTC ATTATGTATT GGATTATCAT CATTACGAGC GCCAAATATT	60
----	---	----

ACCGTTATGA CATGTTGATC ATTTTtAGTn AATGTTAAAG CTATAACGTG TGATATTGCG 180
 AGTCTGTCGG CATTCAATAT CTTChAAATG CCCTACAT 218

5 (2) INFORMATION FOR SEQ ID NO: 1039:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 396 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1039:

TTTTAGACGC TAATGnTGGC ATATATTACT TGCCATGCAA ATGAGTCATA AGGCATAGTT 60
 TTCAAGAAGG GTTCTGACAA TGAAGTGAGC AACATGTCAT GATGAGAAGC AGGACTATAC 120
 20 AATGAGAATA ACCTTTTGAT TTTTCATGCA TAAGGCGATT CAATCAAAAG CAATCACCTT 180
 CCAACTGAAT TGTCATTTTG TAAAATAAAA TAATCGATCC AATCGTTATC TAATTCATAA 240
 ATGTGTAAAC ACATACGTTA TGAATGGATA ACGATTTTTT GTTATGTTAA AGTGGTACAT 300
 25 TAATCATGTA TTTCGTATGA TAATTAACGA CAAGTGTAAT GGTAAATGT ATTTTATGGA 360
 TGGAAATGCh ATAATAGGCC TGGGTTACCA TGGnGC 396

(2) INFORMATION FOR SEQ ID NO: 1040:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1040:

40 AGCATTtTGT TTAAAGCTG ATGAAGCAAT AATAAACTCT TCATATCCTC TGTCCATTAC 60
 AATTTGACCT TGATATACAA TTTCTTTAAA GTTTTCGATT TCTTTAAATT CTCTGCTATC 120
 ATTTAAATA GGTGCATTG TAATAACATT CGCTTCCTTC TTATATCCTT TAGATTGATA 180
 45 ATATTCTTTT GCTGCATGAC TTACTGTTAC GAAGGCATTA ACACGATGTT TTACTATGTG 240
 TTTTCTATA CTTTCTACAA ACTTTGAAAT AAGTGGAACT TTATTAATAA AGGCATTTTT 300
 CGCATATATt TCATGCGCAT CATAACAAT ATTAGCTTTT TtATAATTGC TTAAATAGAC 360
 50 CATTAAATAAT ACGTCGAAAT CATTTGCATG AATCACGTCA 400

(2) INFORMATION FOR SEQ ID NO: 1041:

55

- (A) LENGTH: 268 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1041:

10	ACTTGCGTTT CTATTAGCT CAGACAGGAA AAATATTCGC TGGAAATATG TAGGAATTTT	60
	ACTTGCAATT CAGCTTATCT TTGCATTTAT CTTGCTTAAA ACTACAATGG GAATCAAGGT	120
	TATTGGTGGT ATTTCAAACG GTTTTAACTA CTTATTATTA AAAGCAGCAG ATGGTGTTAA	180
15	CTTTGTATTC GGTGGTATTC AATATATCGT CCCTAAGCAC CCACCCATTC CTCCTCCGGG	240
	GAATAATAAC CAACnGAATT AATTCCGn	268

(2) INFORMATION FOR SEQ ID NO: 1042:

20

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 277 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1042:

30	GTGAAGATAA GAAATTAAAG TTAAGACAAG GTTAAAAATA AGCAAAAAAT ATTGAAATTA	60
	CAGATCATTG GAnTGGTATT CCAGAGGAAG ATCAAGATTC ATTTTGTATC GCTTTTATCG	120
	ATGGATTAAA TCTCGTTCAA GAAGTCAAGC GGTAATGGAC TCGATTATCT ATTGCTCAAA	180
35	AAATCATCAA TTAAACGGnG ATCGATTAAA ATTAAAGTGA ATTAACAAGG AGCAACGTTA	240
	AATCTATTTT AATCAGACTG AGACGTCATC AAGTCAT	277

(2) INFORMATION FOR SEQ ID NO: 1043:

40

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 379 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1043:

50	TTATCATCAC AAGTGATTTT GAAAATAAAA AAATCGAAGA TGATGACAAT GATTcAGGTA	60
	GTGTGCCGAT ATTGCCGTTT ATTAAACGAG ACAAAGCATA TCAAATTATC CAGAATTTAG	120
	TACCAGATT ATCTTATCAA CAAGTTAATG AAGGTATGCC TTTGTCAGGn TTCCACAGAC	180

55

AACCATTATT ATTGTTATTA CCGATATAAC CATAGTCATA TGGTTATTGT GATCCAAGGT 300
 ACCTTATGTC AACAGCGGGT TAAAGCAGAG ACAATACATC AGATTCCGGCA TGAACGAAAT 360
 5 TATATTTAAa GCTnCTAAG 379

(2) INFORMATION FOR SEQ ID NO: 1044:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1044:

ATATTGAACT TGAAGACCAT GCAGATATAA ATGAAATAAC ATATCGATTA TCAAAAATTT 60
 20 TCGGTATTAA ATCTATTAGT CCAGTATTAA AAGTAGAAAA AACAAAGAG GCAATAAGTG 120
 CAGGGACAAT TAAATTGCG CCATTnTGAA GAAAACAGCA CATTTAAAT TGATGTGAAG 180
 CGTGCCGATA AAAATTTCCC AATGGATACG TATGAATTAC AGCGTGAATT GGGTGGTGCA 240
 25 GTATTGAAGC ACTTCGACAA TATTTTCAGTG aATGTcaAAC GTCCAGATCA TGgAAATTCg 300
 AGTGGGmAGT TtAGGmTTAG ATGCCAATTT TATATGGTAT GnAAGAAGTT GGTTCGGGG 360
 30 TTCCAGGTGG GATTACCnG TTGGGTACTG GTGGGTnAA 399

(2) INFORMATION FOR SEQ ID NO: 1045:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1045:

AnGATAAACA GAGGAGCACA AAAATGAATn nAAATATAAT CATCAAAAGT ATTGCGGCAT 60
 45 TGACGATTTT AACATCAATA ACTGGTGTGCG GCACAACAGT GGTGATGGT ATTCAACAAA 120
 CAGCCAAAGC AGAAAATAGT GTGAAATTAA TTACCAACAC GAATGTTGCA CCATACAGTG 180
 GTGTTACATG GATGGGCGCT GGAACAGGAT TTGTAGTTGG GAATCATACA ATCATTACCA 240
 50 ATAAACATGT TACTTATCAC ATGAAAGTCG GTGATGAAAT CAAAGCACAT CCTAATGGTT 300
 TTTATAATAA CCGTGGTGGA CTTTATAAAG TTAATAAGAT TGTAAGATTAT CCTGGTAAAG 360
 AAGATATTGC GGTCGTACAA GTTGAAGAAA AATCAACGCA 400

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 626 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1046:

ATAAACTTCT ACACCATCAT CTTTACCTAC AACATTTGTA GCTTTAACTT TTAGTCCAAA	60
GTTATCTTTA AAGAATTGTT CACCTACTTT TTCAAATTCT TTACGATGCT TCTTCACAAA	120
TTCAATCGCA TCTTTTTCTG CAGGTGGTTG AAAGCCTTGT CCTACATATT TTGAAGCTTC	180
CATTTCTTCT GGTACTGATT TTGTTTCTGT ATTTGTTTCT GTATTGTTT CTGTATTGT	240
GTCTTTTTTT GATTCATTAT TCATCGCTGA ACATCCTGAT AACAGTAGCG TTGCTATTAA	300
GATTAATTC GCCTTTTTAA ACATAGCTCA TCACCCATTT ATGTGKTTAT ATAGAATATT	360
AAAAAGCATT TTAATAAAAA TTTATTCATC TTTACGATAC ACTCCATATT CTATTGGATC	420
ATTATCATCA TAAAAGGCTT TTTTAGTATT AATAGTAGGC TTTCTATTT GAATTGTAAT	480
TGTAGATTTT TCTGGCTTAT TTTTAAATT ATATAGTTTA TCACTCAGTT CAATTACATC	540
ATCAACTGTA TTGTCTTTAG TAAAATTTTC CTTTGTACTA AACAATGTTG TTAGTGTATC	600
TGTATTAGCA GCGnAATTCA CTTCTT	626

(2) INFORMATION FOR SEQ ID NO: 1047:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 317 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1047:

GACGGCGTTC AAAATGCAAC GGTCAATTTA ACAACAGAGC AAGCTAAAGT TGACTATTAT	60
CCTGAAGAAA CAGATGCTGA TAACTTGTC ACTCGCATT C AAAAATTAGG TATGACGCGT	120
CTATTAAAGA TAACAATAAA GTCAACGTCA CGCAACTGAA GCGTACAACA TAAATTGATA	180
GnTATCATAT CAGCAGTATA TCTTACCACT ATAATGTAAT GTTGACACTT TCATATGCAT	240
ATACAGACTA TTACGATCAT GTCCATTAT TAGCnCACTG TCATTATATG GTGGCATTAT	300
GAGGGTATAA ACTTAGn	317

(2) INFORMATION FOR SEQ ID NO: 1048:

(A) LENGTH: 223 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1048:

10 ATACGnATTA TGCAAAAAAT AAAGTGAGAT TGGCATATGT GTAAATCTAA AATACTGTTG 60
 AAAAATATTT TTAGTGAAGT AATCAGAAGT TAAAGATTTA ACTGAAGTAA AAATATAATC 120
 AAGATTACGA GTCATTAACA TTTAGCTTTA AAGAGGGAAA CATATCAAAG TAGGTTAGCT 180
 15 AAGnAAAAACA CCGACTAAAT CGGGGATATT TCGTGACATG TTG 223

(2) INFORMATION FOR SEQ ID NO: 1049:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 608 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1049:

AAATATGAAT AATGCAATTG ACCAACGTTT GATAACGTTG CCATACCCAT ATTAAATTTT 60
 30 GTATAACCTT GTTCTTTACT CCAAAGTAAC ATATGCAAGT ATAGACCATC CATTAAATGGT 120
 AAATCTAACT CTGGCAACCA TCTAATTAAA TCGACTGAAA TGGCATCATT AAAGTATGTT 180
 GGCATTAAAC TACAAAATGC AATTACTTCA TTTTCTTCAT TTCGCATTAC ACCAATTGGC 240
 35 GCTTTAGATA AGTATTCTTC ATTAAATnCA CCAACAGAGA AATGCATTTC CyGACGATTA 300
 TCTAGCCaTA AATcACTTAC ATGTTGAAGt TCATTTAtAA ATTCAGTTGr AAACGGtGGt 360
 TCAATAATTT CGAACGAAAT ATTAAGTTCA TCGAATTnAT TTAAAGTCGC TCTAAATCCC 420
 40 ACGGCGTTTT TTACCTGAAG TTGAAAATTG CGTTAAATCA ATAATTGCTT CTTCACCTAA 480
 TTTGAAAAAT TGGTTACCGG nATTATGATA TAAAGGCATG TGTTGATCTG TAACTTGATA 540
 45 GAATATAACA TCATAGCCTA AATACTCAGC GTAATTATAG AATGCTTCTA ACAATTATC 600
 AAAGGCAT 608

(2) INFORMATION FOR SEQ ID NO: 1050:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 292 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1050:

5 TGT TTTTCTA CGCTAAACTA TATGATAAAT TGCCTATGTA TTTACTGGTG TTTATGGCCT 60
 TTACAGCTGT AATTTTGATT ATGATGTACA TACAAGAGAA AAATGAAAAT TACAAAGTTG 120
 AAAAAAGATA TGTGGTTAGA TATCTCACAC TTAACGTTAT TGTGGGATAT ACTTTGCCAT 180
 10 TGCTTTTGT ATCTATTTAC GTTTTGGTG TAGTCGGTTT TGGATTGAT GTTTTCAATT 240
 ATTGTCTAGG TAKTATCTTG ATGTAATWTA TTTCTTGGTT TAGGTnTATT nT 292

(2) INFORMATION FOR SEQ ID NO: 1051:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 734 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1051:

25 TAAATCnTCA AAGATTTCTT CTAAACATT ATAATCTTCA ACGATACTCG AATAATAATA 60
 CAGCTCATCA TTTCTGGGCT TGTATAATA AAATATAACT TGCCACCTC CATCGTCCAC 120
 ATATGCAATT ACATATACCT TTTCCCACTC TACTGGTATC ATTTCAATTGA TCTCATTGTC 180
 30 AATTTTATTG TACATTTTAC TTAATTTTTC TTCGAAATTC ATGTTATCGC CTCTATTGTT 240
 CAGCTTCTTC TTGCTCTTA ACATATTGCT CGATTTCTTT AACTTTATTA ATTTTCATATT 300
 CCGTTTCTGG TAAAAATCCA AATTTTCTAT ACTTATAGTA ATTTTGTGCA CCTACTTGAC 360
 35 CAAATTCTGA ATTTATCCAA TCAATATAAT CAAATGAAAC TTTTAATTCA CCTTCTCTTG 420
 TAAAGTCAAA TTCGCATGAT GTCCATGGTT CATGTCCTTC TTCTTTAAAT AAATTCCTTA 480
 AATTTTGAAA TTGTTTATAC AACTCATACA CTGAATCCAT AAATCCGAT TCCGATATAT 540
 40 CATATTTATT TAATACACTA GTATAGTAGT ATAATTCATT ACTTCCAGGT TCTGTGTAAT 600
 AGTAGAACAC TTCTCCTCCT CCGTCATCTA TATATGCCAT TGCATATACC TTTTCCCACT 660
 CTACCGGTAT CATGCTGCTA ATCTTATTCG CAATCTCGCT ATACATTTCA CTTAATTTTT 720
 45 CTTGAAAGT CATG 734

(2) INFORMATION FOR SEQ ID NO: 1052:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 775 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1052:

5 AATTTTAAACC AAAATTGCTT TTTTATCTGA AAAGAAAGGC AATGTTAATG TTTCTTCAAC 60
 AATTGGTGCA ATCTCTGTTT CGTATAAATT ATATTTACACA AAGTTAAAGT CATCTCTATC 120
 ACTTTTCAAA AATTGTGATA TGATTTCTGC ACTTTGTTTT TCAACCAATT CAGGCACATC 180
 10 TCCATAAATa GCTACAATAT TGTCGCTCAT TCATTACACC CTTTCATTCTG TTTGACACAT 240
 CAATAGATTA TATCATGTTT CTGTATCTAT AAACCACTTG CATTTCATA AGAGCTTGAA 300
 TCAACTTTTA AATTATCATC TAAGTCAATT GTAACCTGAC CGTTTTtGTTG ACTATTGTAA 360
 15 ATGCGtACTG CGAATCCTTT GCAATCGTTt aACAACCTTCT ATATTAGGaa GwTGATACAT 420
 ATTGTTCTTC CCAGAAGAAA TCAAACCTTAT TTTAGGCTTA ATCATCTCTA TAAATTCTTT 480
 AGAACTACTT GTCTTGCTCC CATGATGTCC TACTTTTAAA ATATCAATCT CCGGCAAGTT 540
 20 ATATTTTTTTT AGTAGTAAAG ATTCATTATT TTTACTAGCA TCGCCCATTA ATAAACTTT 600
 TTTATTTTGA TATGTAATCA TAGTAATAAT CGAATACTCA TTTTATCTC GGCTATTTGG 660
 AATAAACTA TCAAAAAAAT AGAAACTAC TATCTCCAAG TTTAAACTA CTAACCTGTC 720
 25 TTACATCCAT AAGTTTAAATG GTTGTACTTA TGGGCTTAAT TTCCGATAAT AACCA 775

(2) INFORMATION FOR SEQ ID NO: 1053:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 234 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1053:

40 GAGACCAATC ATCATAATGA ACGACAATCA TGTATTTAAT ATTTAATAAA ATACTGAATA 60
 ATGCTATAAC TGCCATTAAT TGAAGACCAA TCGCATCTAA TGCGACAACA CGATCGGCAA 120
 GTGATGGGCC TAGCACAACG CGAATGAGCA TAGCTAACAT AGGAnGTGAC AACTATGATT 180
 AATGCAATAA CGATGAATAA CCTTGATGAn TCATTATAAT TCGCCACCT CTCT 234

(2) INFORMATION FOR SEQ ID NO: 1054:

(i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

CTACTGGTAT CATCCACTGA TTTCATTTGC AATTTTCATTG TACATTTGAC TTAGCTTTTC 60
 TTCGAAAGTC ATGTTATCGC CCCTATAGTT CAGCTTCATC TTGCTCTTTA ATATATTGCT 120
 5 CGATTTCTTT AACTTCTTCC ATTTTCGTATT CCAATTTCTGG TAAAACACCA AACTTTTAT 180
 ACATATAATA GTTTTCACGG CCTAATTGAT CAAACTCTGT ATTTATCCAA TCAATATAAT 240
 CAAATGAAAC TTTTAATTGA CCTTCGCTTG TAAAGTCAAA TTCACCTGAT GTCCATGGTT 300
 10 CAAGCCCTTC TTCTTTAgAA AGTTTCTCTT AACTTCTTAA ACAATCTATA TAAATCCGTC 360
 CACAAATCAT AAAATACTTT TTCAGAGACA TTATACTCTC 400

(2) INFORMATION FOR SEQ ID NO: 1055:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 333 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1055:

25 GnGGCATATA TAGCGTTGGC GGTATAACAA AGAAAAATGT GAGATCAGTG TTTGGATTG 60
 TAAGTAATCC AAGTCTACAA GTTAAAAAAG TTGATGCTAA AAATGGCTTT TCGATAAACG 120
 AGTTGTTTTT TATTCAAAAG GAAGAAGTAT CATTGAAGGA ACTGnACTTT AAAATAAGAA 180
 30 AACTCTTAAT CGAAAAATAT AGATTGTATA AAGGAACGTC TGATAAAGGT AGAATTGTTA 240
 TCAATATGAA AGACGAAAAG AAGCATGAAA TTGATTTAAG TGAAAAATTA AGTTTTGAAC 300
 GTATGTTTgA TGTAATGGaT AGTAAGCAAA TTA 333

(2) INFORMATION FOR SEQ ID NO: 1056:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 347 base pairs
 40 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1056:

45 TGGGCAGTTA CATTATGAT TATCAGCACA TGTTTAGCTT ATAAATTTTT ACGAAAGTTT 60
 AGCCGTTTAT AAAGCACATG CATAATGAAA CGAGTATTTG CCACTTGATT AGTACTTCAT 120
 50 TATTATGTCG AAAATAAAAA TAAGTGGTAT TTTTAATATA TTAAGAAGCA CTCATAATCG 180
 GCTGTTAATT AATAATATTT TTCATAAGTA TTGATTCATC ATTTTCTTTA TGTTAAATAT 240

TAATACAGAT ATCAATATGG CnAAAGTGTT nnATGAATCA GAGGTAA

347

(2) INFORMATION FOR SEQ ID NO: 1057:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 312 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1057:

15 TCCAAAACGG AGAGCCAAGA GTTAACTCTA CTTATATAGG ATACGCGCCA ATTGATGATC 60
 CAAAATTAGC GTTTTCAATT GTATATACAA ATCAGCCTGT ACCACCACCA TGGTTAACAG 120
 GTGGAGACTT AGGTAGAGAT GTAATTAACT ACTACTTTAA GCAGTTAGGT AAAGATGATA 180
 20 AAAATAAAGA CAAAGACAAA TAAAATTTAA CCTGACGATT GTGTAGCGCA TGGTTGTAA 240
 ATTTTAACTT TGCnAATATT ATAGATGTTG GTATAATAAT AAAGTCGTAT TcAGaATGA 300
 TwAGGaGGGA AT 312

25 (2) INFORMATION FOR SEQ ID NO: 1058:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 404 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1058:

35 GCCATTTTGC GATGCTAAAT CCATTGAAAA TGCAATACCT GAAAGTCCAC CAAACAAAGA 60
 ATTGTTATAT TGACCACTCT GTAAATATGG TGCTAATTC ATGATGTACT TATGCACAAT 120
 TTGCTCAGTA TTAATATCGA AAACTTTTtG AtAAGCATCT AAAACAAAA TGATACCTGG 180
 40 TATACCATGT GATAAAGTTG AAGGTTCAAA ATAATCGGTT TCAGTAGATG CTTTtGAAAT 240
 GAAATCATCT ACTTCTGAAA TCTCTTGAAA CTTCTTCTTC AATATTAGAT TCATATCAAT 300
 45 CATTTTtGTC ACCACACCAA TGCTTTTGAG TTTTAACAAT TTCTTTTACA ATAGATAAAA 360
 CAAATGTTTC TTGATCCTTA TCAATGCCAA ATATTCGATT GnAG 404

(2) INFORMATION FOR SEQ ID NO: 1059:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 233 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1059:

5 ATGTACATAG GnatGACATA AGTGTGCCTT CTTTTTCAAG ACCCTCCATA GAAATAATAC 60
 ATGGATGTGC ACCACGnTAC TACCTCTTTA ACGTTACCAG AATTGATAAA TTAACCTTCT 120
 CTTGTGTTGC TAAACCTACA ACTGGTGTAC CTTCTTCGAT TAAGGCAATT GTACCATGTT 180
 10 TAAGTTCTCC ACCAGCAAAA CCTCGCTGAA TGTAGAAATT CTTAAGTTTA AGT 233

(2) INFORMATION FOR SEQ ID NO: 1060:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 295 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1060:

GCACCACTAC TGTTAATAAC ACACCTAAAA AAGTCGAAAT ATAAGGGTTA ATCGCTGTTT 60
 25 TAAATGTATA TCCTAGTGCG ATTGCCATTA ATTATCCGAT AGCTTTTTTCT GCTGCTTCCA 120
 TAGTCATCTC ACCTAATGTT GGATGTGCAT GGATTGTAA TGCGATATCT TCAGCATTCA 180
 TACCAGCTTC AATTGCTAAA CCTAATTCAG AGATAATATC TGATGCACCA GTACCAACTA 240
 30 CTTGAGCACC GATTAAAGTA TCATCTTCTT TAGTGTAATA AGTTTAACAA ATCCG 295

(2) INFORMATION FOR SEQ ID NO: 1061:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 413 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1061:

CAATGCCTCT CGTGTTCAG CTTCTCCGG TGAAAAACCA ATATTTTCAA GTGCAATAAT 60
 45 TTGATCTAAG TCCTGCATGA TTAATTTTCT TAATACCATA AAAACACTTC CTTAGCCCTA 120
 TATATCAAAA GTAATACCTT CATCTTTATA TTCATTATTT TCAAACCTACC TTCACTATAA 180
 AAAACGAGTC TTTTATTGA AAGCCATTCG CCTTAATCAA TTAGACAAGT TGTATKATAC 240
 50 GAATTAGTAA ATAATCATGA TTATAATTCA TTTTAAGCAT ACTCAAAAAG ACTGGTACAT 300
 GTATACCACT CTAATTCGAA AAAATATATT TAGTTAAAAC CATTTCAAAA ACGACTTCAC 360

55

(2) INFORMATION FOR SEQ ID NO: 1062:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1062:

AAGCTGCATA CTCTCCAAAC TTAGAACCAC CGAAGCGATT CACAAAATAT TTATTAGCTA 60
 AAAAGTCGGC ACATAAAATT AATATTGTGA GCAATGCCAT AGATACATAA AACACCCATG 120
 ATAAATGCTG ATTATGAAAG CCAAATTGAT AGATTAAAAA GCCAACCCAT AATACTAAAA 180
 CAGAAGGAAT AATCGGCTTA ATCAACCCAA CAAATGCTAA CATGAAGGCA GCGATGATAA 240
 GTAGCCATAA AATAATTGTC ATGTTGATAT CACATCCTCT TTTGTATTTT TTGATTTTTA 300
 GTAAAGAATA TAAGTATCAT ACCTAACATG GTTGATGCTG CTGGATACCA AAAAATACAT 360
 TTTCCCAAAC CTTACCCAGT GGACTCCTAG GCACCACCTT 400

(2) INFORMATION FOR SEQ ID NO: 1063:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1063:

TGGATCAGAA GGAATCTGGA TGCCTATTGA TGACAAAGGA TACTTTAATT TTGACTTCAA 60
 AACGAAACGT TTCGATGATT TAGAGTTAAA GAAAAATGAT GAGATCTCAT TAACATTTGC 120
 ACCTGATGAC GAAGATGAGG CATTGAAGTC ATTAATTTTC AAAACTAAGG TAACGAGTTT 180
 AGAAGATATT GATAAAGCAG AACTAAATA TGACCATACT AAAGTGAAA AAGTAAAAGT 240
 ATTGAAAGAT GTTAAAGAAG ATTTACATGT AGATGAAATT TACGGAAGCT TATATCATAC 300
 AGAAAAAGGT AAAGGTATTC TTGATAAAGA AGGTACTAAA GTaATTaAAG GTAAGACTAA 360
 ATTCGCAAAT GCAGTTGTGA AGGTAGACTC TGAAGTAGGT 400

(2) INFORMATION FOR SEQ ID NO: 1064:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1064:

5 GTGGTCGTGA AGCAACTGAT TTCAAAGAAG AAGATATCGA GCTTTTCAA GACATTGCCG 60
 ACAAAGTAAA ACAAACAAAT AGTTATGATC TAGCGTTTGA TGAATTAGAA AAAGAAAAAG 120
 ACTTCCTGCA AGTCATTGTC AAAAACGATG ACAAAACTT ACCTACTAAT CAAAATGTCTG 180
 10 CTCAACTAGT AGAAGATTTA CGCCTAGAAA TCCAGAAAAT GCGCGAAGAA CGTCACCTAC 240
 TTGGTCAAAT GATGAATCAA GTACATCAGC AACAACAAGA ATTAAAAGAA CTTCAAATC 300
 AACTTACATC TAAAATCGAT TCAAATAGCG AATCCTTAAA AGCCATCCAA ACATCACAAG 360
 15 AGGCTATCCA AGAAGCGCAA GCCTCTCAAG CAAAAGCATT AGCTGAATCC ACCAATAAAG 420
 TTGGAAGAA TGCT 434

(2) INFORMATION FOR SEQ ID NO: 1065:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 239 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1065:

30 TAAACATAT GGCTTGATAT TGCAGAAGAA TTGAAGTTAT CGTACCAGAA ACGGTACATA 60
 ATCATGATTA TATGGAAACG AAAAAATAA AAATGGGTCA TTAAATATAG GAGGATTTAA 120
 CATTGAATTT TGAAGGTAAn TAATTGGAAA AGATTGAAG TTGCATCGTT ATTATCGTTT 180
 35 AATGATTTTA TCACTGGAAG ATTACTTGAA GGTGCAAAAG ATACnTTGAT TCGACATGA 239

(2) INFORMATION FOR SEQ ID NO: 1066:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 202 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1066:

55 AAATATCTAT AGACCTTTAT TCGGAAATGG TTCGTCTCCA GAATGGAAGG ATGAAGTACC 60
 GAGTATTGAA AGGTAGAAGG TCTCCAACT TTACCTGnAC GTGCACGAnA TCTAACTGAA 120
 GATGATTTAG CAATTGAATT GAAACAGCCA ATTGTGCGTT GTAATAACTT AGGAGAGTTA 180

(2) INFORMATION FOR SEQ ID NO: 1067:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 749 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1067:

```

CTACAATTCG ATCAACGCCG AAGTTAGATA CACCGATTGC TTAAATTTTA CCATTTTCTT      60
TTAACTCTTC TAATGCACGC CATGAGCCAT ATACGTCATT GTAAGGTTGG TGAATAAGCA      120
CTAGATCAAG ATAGTCCAAA TTTAATCGTT GTAATGATCG TTCGATTGAT TTGATAGTAT      180
CTTCATAGTT AACATTTTCA ATCCATACTT TAGTCGTAAT GAATAGTTCA CTTCTATCAA      240
TACCTGAAGC TTCAATACCT TGACCTACTT CAGTTTCATT TAAATAACTT TGTGCTGTAT      300
CAATATGTCT ATAACCCGCT TTAATTGCTT CTTTAACCGC TTCTGCGTGT TGTTCCTGAG      360
GAATTTGAAA TACTCCGAAA CCTAATACTG GTATCTTTAC ATCTTTACTT ATTTCAATAT      420
GATTCATAAA ATGAACCTCC TTTATCTTGT ATGTCCACTA TAAACCATGG AGCGAACTCT      480
ATAGCAAGCC ATAGGTTTAC TTTCTCTCCC AAAGACGTTT AACATTGTCA TCATAATTAT      540
CACTAGCATT GATTTTAGCA ATATGATTAT CATAATTATC AATTTTATAT TGTAGTAAAT      600
CTCTCGCATC ATGAATTGTC TTTAAGTTTT CATTTAATTC TTCGAGTTGC TTATTTAAAA      660
TTTGCTTTTG CTGTGCTTGA ATATTTTCAT TTTTAGGCAA nTGCGCTAAC CTGCAAAATT      720
CGATAAGCGA TTCAATACTG ACGCCTGCA                                     749

```

(2) INFORMATION FOR SEQ ID NO: 1068:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 436 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1068:

```

TCCTCGAGAT AGACAAGAAT GGTGGTACAA TGTTTTTTAT GACGATGGCA AATATnATAA      60
GAAnATAAAT AAGAATGACA TGTCAAAAAT TAGGCGTAAT TTTTAGATTT ATCTATATAT      120
TTATGCCACC GCTCAAACCT TTATnAAAAG TAAAAAGACA ATATTTATAT TGAACCTGTG      180
ACGTGAATTT CTAAACACA ACCAACAATC CACAAGGAAA CTAGTAAATA AGTTTAAATA      240

```

TACTGACTAA AAAGATTTTG AAAGTGATTA AGTATTTAAA ACACGCCAAT CTTAAACCGC 360
 TATTGACAAA TATGAAGCGG AACCACCGTT GTnTTTCITT TnCAGGTCCA ATTGAATAAA 420
 5 TGTGATGAA TAGAAT 436

(2) INFORMATION FOR SEQ ID NO: 1069:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1069:

TTCTCAAGTA ACAGGTATTA TTGTTACTGG TACAATCAAT AGAGAAAAAT ATGGCATTAA 60
 20 CTTTAACCAA GCACTTGAAA CTGGTGGCGT AATGCTAGGC AAAGATGTTA AATTCGAAGC 120
 ATCAGCTGAA TTCTCAATCT CAGAATAATC TCATTACACA ATCCTCATTG TTTTAATATA 180
 TATACGAAAT GCCAACTATA TCATCCCTAG GTATAGTTGG CATTTTTCGT TTAACTCATG 240
 25 TGTAACAAGC ACTAACTGTT CACCCTAATT TTAGACGCCT TTCATTCAGG TAGGTCTTAT 300
 TGAAACTGAA AACTTTGATG ACCTTTTGcA AaGCCATTAA CTGTATAAAG CATAGATCCT 360
 CCGCCCATTT CTATATCATT GGAACAAATG ATGAGTTGAT TTGTTCCAGG TATAAATTGC 420
 30 GGATGAGTAG AACGTAACAT ATGCCCTTCA TCTCGGCCTG GTATCAATAT TTGTCCTATT 480
 GGATAACCCC TTTTATTAAA AACTAACACT CGACCTTGGG CCATACATTG CTACGGATAA 540
 ATTATCGGCA C 551

(2) INFORMATION FOR SEQ ID NO: 1070:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 325 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1070:

TATnATGGTA ACGCATGAnC CTGTTGCAGC ACGCTATGCA AATCGAGTAG TGATGCTAnA 60
 AGATGGTCAA ATTTTCACTG AATTATACCA AGGGGATGAC GATAAACATA CCTTTTTCOA 120
 50 AGAAATAATA CGTGACAAA GTGTTTTAGG TGGCGTTAAT TATGACCTTT AACGAGATAA 180
 TATTTAAAAA TTTCCGTCAA AATTATCAC ATTATGCCAT CTATCTTTT TCGTTAATTA 240

CAGAGTCATA TCCAATTATA CnGGC

325

(2) INFORMATION FOR SEQ ID NO: 1071:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 771 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1071:

15	TAAATGCCGC AATTGGATAG CTTTAATATA ATTAATAGTC AAGGTATTnC CAACAAnAAT	60
	ATATTATCCA CCCGAAGGAA GCGGGGTATn TTTGTTACTG ACGTCCCACA AGAAGCACAn	120
	ATATTAGATA GTGCTTACTC TGAAATTA AAAATTAAGAG ATTATGATGA AAACTAATT	180
20	ATCCCAGGAT TTTTGGCGT ATCTCATGAA GGATATATAG TTACATTTC ACGTGGCGGA	240
	TCAGATATAA CTGGCGCCAT CATATCAAGT GGTGTTAGAG CCACAATTTA TGAAAATTTTC	300
	ACCGATGTGT CTGGTATTTA TAAAGCTAAT CCAATATCA TCAAAGATCC TGAATTAATC	360
25	GAAGAAATCA CTTATCGAGA AATGCGTGAA TTATCTTATG CTGGTTTTAG TGTTTTCCAT	420
	GATGAAGCCT TACAACCACT TTATAAAGAT CGCATCCCTG TGGTGATTAA AAATACGAAC	480
	CGCCCTCAAG ATAAGGGTAC ATTTATAGTG CATGATAGAG AAATAAATGC GAAAAATGTG	540
30	ATTAGCGGTA TAAGTTGTGA TAAAGATTTT ACCGTTATTA ATATCAAAAA GTATTTAATG	600
	AATAGACAAG TCGGCTTTAC TAGAAAGATA TTAGGTGTCT TAGAGGATAA TAATATTTCA	660
35	TTTGATCATA TGCCTTCAGG TATTGATACT ATCAGTATCG TTATGCGTTC AAAACAAATT	720
	CAAAATAAAG AAACAAAAGT ATTAAATGAA ATTCGTCAAA AATGTGATGT T	771

(2) INFORMATION FOR SEQ ID NO: 1072:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 237 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1072:

50	CGGTGGATCC GTTTGATCTT GAGTGAGGTC CCAATGAAAG TTATGAGCCA CGTTGTCGCG	60
	CGCACCATAT CGTAGCACCT AGTGATAATA ATAAGGAGGA ATTATAAGTG TTTGATCAAT	120
	TAGATATTGT AGAAGAAAGA TACGAACAGT TAAATGAACT GTTAATGACC CAGATGTTGn	180

(2) INFORMATION FOR SEQ ID NO: 1073:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1073:

```

CTTCTTGTA GGGCGTTACA TCACCGGATG GATGATTATG AACTGCGATG ATTGCATTGG      60
CATTCTCTCT CACCGCAATA CTAAAAATTT CACGTGGATG TACAATCGAA CTATTTAATG      120
TACCTTTAAA AACACAGGTT TCTTTAATCA CTACATTTTT TGAATTTAAC AATAAAATGA      180
CAAAATGTTT TTGTGTTAAA TCTTTCATTG TTGGAATCAT ATAATCAGCA ACATCACTTG      240
GTTGCGTTAT TTKTATACGA TTATTTTCAG CTCTTCTCCC CATCCTTTCC CCTAACTCAA      300
ATGCTGCTTT TAAAGTAATT GCTTTTGTGA ATCCAATCCC TTAACTTTT ATCAAATCGT      360
TAATTGAAGA TTTTTTCAAT TCATTCAGAT TcGAAGCAGA TTTAAGCAGT TCATTACTAA      420
TGTCTATGCT CGAGAATCCT TTTCTTCCGG TGTTAATTAA TATAGCTAAT AATTCTGTAT      480
TCGAAAGACT TTTGCAACCA TGGCTTAACA AACGTTCTCT TGGCATTCTT GAAGTTACCA      540
TTTCTTTAAT TTTCAAAAAT ATACGCCTCC TAAAAATTGA TGGATATCAT TATAAAAAAG      600
TGAATTGATA AAAAAGGAaA TAAATATAAA TGGAACAAGG GGTAATAGTT TAATCGGCTT      660
AAATATCATG GTAATTAAAG CAACTAAACC AGCAATGACA AATGTAAATA AnATGACATA      720
AATAGTGAAT TGG                                         733

```

(2) INFORMATION FOR SEQ ID NO: 1074:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 768 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1074:

```

ATGGTTATTT TTTATGCTCT AAATCGTATG TGTTTATAAG GTTCAATTTC TTGCCATTAT      60
TAATTTTGCC GAGCAGTAAT GGGTTGTAA AGTATCCATC AATTATGCTT AGCAATATTG      120
AAAAATTTAA ATGAAAAAGC GGTTC AATTC TTGAAAAATA GTGTATATTT ACATTTTGGT      180
AATAACTAAC TATAATTTAC ATAAGGAGGT TGTCATTGGG TAGTATTGTT ATCTTATAGT      240

```

TATTTACATA GAAGATCGAA ATGTGATTAT TGTAATTCGT CACTCAAAGW GTATGAATTA 360
 ATGCCGATTA TTAGTTTTTT ATTATIAAAA GGGCGATGTC GAAACTGTCG AAAGCGTATT 420
 5 TCCCTAACAC ATTTCTTAGG GGAAACCTTT GCTTTAATAC CTATCGTCTT TATTAAGTAT 480
 GATTTACAT ACGTAAATGt ACGCTATTTA TAACTACATA TGTTtTTCTG CTTATTTTTA 540
 CTATGACCGA TATCACTTCT TTAATGTTAG ATTGTCGCTT AATTATAATT TATTGTATCG 600
 10 TTTCTCTCTC GTTAAGTATG ATTTATCCAG TAGCTTTTAT CATTATTAGT ATGACCACGC 660
 ATATATTCTA CTTTTTATTT CGGGCATATA TTGGTTATGG TGACGTTTTA CTAATATCTG 720
 15 CACTTTCITT GTTTTTCCCC TCTCCAATTC ACTATTTATG TCATTTTA 768

(2) INFORMATION FOR SEQ ID NO: 1075:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1075:

TCATATTCAT TAAAATACC ATTCACTCA GCTTTTCTCA TTTTCTCTT ATAAAGTTGT 60
 ATTTCTGTG TTAATCTGG CTCAGACATC TCACTCAATT TCTTCTGTT CATCGGCAAT 120
 30 ACCACTTTCT TCTAATTTAG CTTTAATTTT ATCATATTTA TATCCTTTTC TCATAAGGCC 180
 TTCGATAGTT TTTGAAATTA ATTTCTGTTG CGGTACTTTC TTTGATTTT TATTATAAA 240
 35 TTTTCTAAA TCTCGTTGTA ATAAATCGTC TAAAACCGCT TCATCTGTG TAAAATCCAT 300
 TTCATTCAAT ACAGCATGAA TGGTTTCCAT TTCAAACCT TTTGAATTA AAGATTGCAT 360
 TACTTTTGCA TAACTnTAT TTTGTGGCCC nTTTTTTGTT 400

(2) INFORMATION FOR SEQ ID NO: 1076:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 753 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1076:

TTAGGTCAAA ATATGTTTTA TCCAAAAGGC ATTTTATCTC AATCTGCTGA AGCAAAGACT 60
 ACAACATATA ATGCAACTAT AGGTATGGCG ACAACAAAG ACGGAAAAAT GTTTGCATCA 120

CAAGGCATCG AAGAATTACG TGATTTATGG CAACAAAAAA TGTTGCGTGA CAATCCAGAG 240
 CTATCAATCG ACAACATGTC ACTACCAATT GTTACGAATG CATTACACA TGGTTTATCT 300
 5 TTAGTTGGCG ATTTATTTGT AAATCAAGGT GACACTATCT TGTTACCAGA GCATAATTGG 360
 GGTAATTACA AACTTGTTTT CAATACGAGA AATGGTGCAA ACCTTCAAAC ATATCCTATC 420
 TTTGATAAAG ACGGGCATT TACTACTGAT TCACTTGTAG AAGCTTTACA ATCATACAAT 480
 10 AAAGATAAAG TCATTATGAT TTTAAATTAT CCTAATAaTC CGACAGGTTA CACACCTACG 540
 CATAAAGAAG TGACTACCAT CGTCGATGCA ATTAAAGCAT TAGCTGATAA AGGTACAAAA 600
 15 GTTATAGCTG TTGTGGATGA CGCATACTAT GGTTTATTCT ATGAAGATGT GTATACTCAA 660
 TCATTATTTA CTGCATTATC TAATTTAAAT TCAAATGCAA TATTACCTGT TCGTTTAGAT 720
 GGTGCAACAA AAGAATTTTT CGCATGGGGA TTC 753

20 (2) INFORMATION FOR SEQ ID NO: 1077:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 405 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1077:

AGAAGCGATT TGACaATAAA ACAATTGAAA AACTTTTGGA AAGCAAGTGG TGGGAGAAAA 60
 CGCCTGACAA ACTAAAAGGA TTTTCGGTTG AATATTTAAA TAAAAAGGAT ACTTAATGAT 120
 35 ATGAGAATTT TAAATATTGT ATCGAGTAAT ATTGTTCAAG ACCCAAGGGT ACTTAAACAA 180
 ATAGAAACAA TTAAAGGCGT TACGGATGAT TATAAAATTG TTGGAATGAA TAATTCACAA 240
 TCTACTAATA AGCGATTGGA AAATTTAGAT TGTAATTATC GTTTGTTAGG TAGCAAGGTA 300
 40 GATCCAAAAA ATATTCTTTC TAAATTAATT AAGCGTATAA GATTTGCaAC AGGTGtTATC 360
 CGAGAAATTA AAGCTTATAA AcCTGACGTG AtTCATGCaA ATGAT 405

45 (2) INFORMATION FOR SEQ ID NO: 1078:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 504 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1078:

CATATTCACA CCTCAATATT CCATAACTTT TACTTTAATA TCAATCCACT TAAATCCATT 120
 CACATCATTG AGTATAAAAT AATATGGTGA TATACGTTGA AGTTGATTAT CTTATATATA 180
 5 AGTATACTTC ACAGAATTTA TAAAATTGTT TAGTCATCAT CTTCAGTGTT TAATTCAAAA 240
 ATATAGAATC TACTGTATTG CATmATAaAT TGAAATGCCA CAAATTGTGC TCCAGCGATA 300
 ACCAACATGC TCACAATAAT TATAAACTTA GATTCTACGA CATTAGGmAC CaAtATATATT 360
 10 CGATAATTTT TTCAATATAT ATAAATGAGC TCAACGCAGT TAATAATAAA CCGAAATGTG 420
 TTTTAGTTTT ACCACCCCAA CGTTTTGTGA CTTTAGGTAA TTTTAATAAC GTGAACATTC 480
 CGCCAATTAC TAATAACAAA TAAC 504
 15

(2) INFORMATION FOR SEQ ID NO: 1079:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 645 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1079:

CATTCTTAAT CAACTTCTTC ATAATGGGAA TTTGGCATGG TAnCGAAGTG TATTACATTG 60
 TTTATGGTTT ATACCATGCA GCATTGTTTA TAGGTTATGG CTATTATGAA CGTTGGCGTA 120
 30 AGAmACATCC GCCACGTTGG CAAAATGGTT TCACAACAGC ACTTAGCATT GTGATTACAT 180
 TCCACTTTGT AACATTTGGC TTTTAAATCT TCTCAGGTAA ACTTATATAA TAAAGGAGAA 240
 35 TTTAATTATG GAATTTAGAG AACAAGTATT AAATTTATTA GCAGAAGTAG CAGAAAATGA 300
 TATTGTAAAA GAAAATCCAG ACGTAGAAAT TTTTGAAGAA GGTATTATTG ATTCTTTCCa 360
 AACAGTTGGA TTATTATTAG AGATTCAAAA TAACTTGaT ATCGAAGTAT CTATTATGGa 420
 40 CTTTGaTAGA GATGAGTGGG caACACCAAA TAAAATCGTT GAAGCATTAG AAGAGTTACG 480
 ATGAAATTAA AACCTTTTTT ACCCATTTTA ATTAGTGGAG CGGTATTCAT TGTCTTTCTA 540
 TTATTACCTG CTAGTTGGTT TACAGGATTA GTAAATGAAA AGACTGTAGA AGATAATAGA 600
 45 ACTTCATTGA CAGATCAAGT ACTAAAAGGC AACTCATTTC AAGAT 645

(2) INFORMATION FOR SEQ ID NO: 1080:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 267 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1080:

5 AGATGCCAAA ATCAAAACCA ATAATTTGGC AACAAACATAT AAGTCCACTT TACCGGACCT 60
 nTAAACTTAT CGTACTGTTA AGAACCAAAC GACTGCCATT ATTTGGTCCT AAAGCCACGA 120
 TGGTGGTAGC CnTTGTAATA ATGCTGGCGT AGAAATTAAC CGCCACCCCT AACACCGAAC 180
 10 TATAACGCGG ATAAACCAAT AAAATTATGA TATAATCATC CGTTAAGTCA ATCCATAGTT 240
 TAACATCCCT TATGTTTAAT TTACTCT 267

(2) INFORMATION FOR SEQ ID NO: 1081:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 441 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1081:

25 ATATTTGTCC ACTATTAGAA GATTGTAGAG AAGGACAAAA ACGTTATAAA GCTAGTTTGA 60
 AAGAAGCGTG ATAACATGAC AAATATTACA AAAGAGGTAT TTGATAATTT AGAACAAGAG 120
 ATTGATTTAT TTGCCAAAAA TAAGACATTA GGTTCTAGTG AAGCAAAGCC CTACTTGGAT 180
 30 GAATACCATA GTAAAATTAT TGATTATTTT AAGCAGGTTA ATGACATAAC TGGCAATATT 240
 GATTTTGATA AATTAAATCA ATATCCTGTT GTGCCAATGA ATTTTAAAGA AAGATATGAT 300
 TATATGATTG AACGTAAATA TCATTTTATG GGCTATCGAC AAATGAAGAC CTTTAAAACT 360
 35 GAATTGATTA AAATGnATGC TAGTTATCAA ACAAGGTTAA AGAATAAGCA GGTATnGACA 420
 AGATTGATCA TTTCTAGTG T 441

(2) INFORMATION FOR SEQ ID NO: 1082:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 410 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1082:

50 AAAATTTCTT AGtTGATGGT TATTGTTGTT GATTTGAATA ATTTGTAGCG AATATTTTGT 60
 AACAGGGCTA CTAAGAATTT TTATAACAAG AAGTGATTTT AATCGTACTA ATTTTCAGCG 120
 TTTGATGACA AAAGTGTATG TCCAACTTA AATAATGGCG ATAAAATTTA AGTCATGGGT 180

GTTGAATATC AATAAAAGAT TAGCAAGTTA GTTCAAAATG GCTATGTATT GTTTTTTGAA 300
 ATACATAGGT TTGAAATCTA TATTAATTAT GCACCAAATA ATTTAATTAG ACTCTATCGA 360
 5 AAATTTCCAA ATTTTGCTTA CTATCTTTCT AATTTTACTT TTATATTGTG 410

(2) INFORMATION FOR SEQ ID NO: 1083:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 586 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1083:

CAGATATTCC CTCTATTTA CAAGTTGGTA ACCCATATTT ATCAGACAGC GTAGATAATC 60
 20 ATACCGAAAA GTTGTTAGAA CGCTACGAGC AGTTGGTGA TTTAGTTATG yAAAGTAATG 120
 ATATGAATCA CGTTTATGTG TTACCACAAC TTCATACATT ACTTTGGAGC AATAAAAAAG 180
 GTGTATAAAG GATAACTTAT TTTACTTTTT GCAAGTCATA TCTAATGAAT AACAATATAT 240
 25 TAATTGATAT AATAGTTGAT TGAAGTATAA CGAACAGGAG TCCATCATGA TTGTATATAT 300
 ATTAATTAAT ATTGCTATCG TCGTACTTAT AACTGGATTC AATTTATATA GGCATCAAAT 360
 GCAACATTTA TCATTAAGTG CAATGTTGTt ATCTATTACG ATaAACGCCT TTATAAATAC 420
 30 ATTCATAATT GATAAATATA ACTTTATTAC TCTGTGTACT ATAACGATGT TTATCATTTG 480
 GACGATACTA CAGTTTTACA TTGATAAAAA ATTAAACCT GTATACATTA CAGACCAAAA 540
 ATTTATTGCT ATCATATTAA CGATTGTTGT CAGTTTAACG CAACGT 586
 35

(2) INFORMATION FOR SEQ ID NO: 1084:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 638 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1084:

AACACAAGGT AGCGTCATTA GCATTTAGAT TTTATTGCCA AATTATATAA CGATTTATAC 60
 ATCAATAAGA AAACAATAGA TGATGTATCT GCAGTTTCTG AAGTGAAATA TGATATTAAA 120
 50 TCACAAATGA GTGATGATGA GATTAAGCGT TTAGAGGAAC AAGGACTTCA AGCTATTAAA 180
 GAAGGACAAT TTGCAGTACT TTTGATGGCA GGTGGTCAAG GAACAAGACT TGGTTACAAG 240

TTAAAAACAT TGAATCATCA ATCTGGACAT ACGATTCAAT GGTATATTAT GACAAGTGAT 360
 ATCAATCATG AAGAACTTT AGCTTATTTT GAAGCACATA GTTATTTTGG ATATGATCAA 420
 5 GAAGCAATCC ATTTCTTTAA ACAAGATAAT ATTGTGGCGC TTAGTGAAGA AGGGAAGTTA 480
 ATTTTAAATC AACAGGGTCG TATTATGGAA ACGCCGAATG GTAATGGTGG CGTGTTTAAA 540
 TCTTTAGACA AGGCAGGATA CTTAGAAGAA ATGTCTAATA ATGGCGTTAA GTCATATTTT 600
 10 CTTGAATAAT CATCGACAAT GTTTTCASTA ACATGTTG 638

(2) INFORMATION FOR SEQ ID NO: 1085:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 618 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1085:

ATTGCAATAT TATCCGCTTG GTTCAACACG TCATCTTTat tAaTAcCATT TATTGgAAaT 60
 25 tCcACCCcATT TCCcGGgCTt CTtACTTCCc TTTATAAAAC TGcCCATTTA ATGAATACTG 120
 TTTAGCTTCA TCTATTGTA CTTCCACTAG TTTACCAATC ATTTCTTTAG GCGCTTTGAA 180
 ATTAAGTAGC TTATTTTAT CAGTGTAGCC AGCAAGAACC TGATCATCTT TTTTACTACT 240
 30 TCCTTCACAA AGTACTGTTA CAGTTTGTC TCGTACTTA CTcATAGCTA TTTGTGAATA 300
 ATGACCAACT TTTTATTCA AACGTTGCAA TCGTTCCTTT TTGACATTTA AAGGTACATT 360
 ATCTTTCATT TTAGCAGCAG GCGTACCATC ACGTTGTGAA TACAAGTACG TATATGCATG 420
 35 TTCAAAACCA ACTTCATCAT ACAGAGTTAA AGTTTCTTCA AATTGTTCTT CTGATTCATT 480
 TGGATACCCT ACAATAATAT CTGTAGTTAA TGCTACATTA GGAATTCTAT CTTTGATTGC 540
 40 TTTTACTAAA TCCAAATAAC TTTCTCGTGT ATATTTTCTA CCCATTATnt TTAATACTGC 600
 ATTATTTCCA GATTGAAC 618

(2) INFORMATION FOR SEQ ID NO: 1086:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 502 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1086:

55

CAGCAATATC ATGATCAGTT TAATTCATCA AGAAAGTCTA TCTATACTTT TACAGAAACA 120
 GTTCAACCAA ATCGTGTTAT TTATGATTTT TATGATTATG ATCCGTATCA ATTAGCAGCA 180
 5 AATAATGCGA AAGCATTAAA GGATCATATT GAACAAAAC TTAACTTTAA AGTCCAAAGT 240
 ACCGGTGTTA KTTACTTTAG TGATGGTACA GTTAATATCA TTCAAGGTTT TGAAGAAAGA 300
 GATAAATATG TTGATACAGT TTCAACAAAA TCTTCGTTAC GACGTATTAT TAGCGAACCA 360
 10 nAGAGCTATC TAAACACCCT TTAAATAAAG AACAGGTTGA nCAAATCACA GCTATATTCA 420
 AGTAAATCA AAGGCATAAT AATCCATTAC GCTCGATACG CACTATGGTT ATTATGCCTT 480
 TAAATTATTT CTATATATAA nT 502

(2) INFORMATION FOR SEQ ID NO: 1087:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 429 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1087:

TTGCCTGATG GACGATGGCC ATCAGAGACT GAGTTTAGAT TGTCTTTAAT GCAACAACCTT 60
 GCTGTAAACC AAATTACGAG TGGAATGAA AGAATAAGTT CAGTTAATGG GCCACCAGGT 120
 30 ACAGGTAAGA CTACTTTATT AAAAGATATA TTTGCTCATC TAGTAGTTGA nAGAGGTAAA 180
 GAGTTAGCTA AACTAAATAA TCCTAAAGAT GCATTGTGCA nACAAAAAT TCATGAAACG 240
 GATGATAAAT ACGTATACTT ACTAAAGGAn nTATTGCCAA ATATAAGATG GTAGTCGCAT 300
 35 CTAGTAATAA TGGAGCTGTT GAAAATATAT CTAAAGATTT ACCGAAAATT GAAGAAATTA 360
 TAAGAnATCC CGAAAAATGT AAATCCCTA AATATGAACA GAATTATGCA AATTTAGCAC 420
 ATGAATTAA 429

(2) INFORMATION FOR SEQ ID NO: 1088:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 611 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1088:

TAAGACCTAA TATTTTGGnC CAATACTATG GTGCAGGTnA TTTATATATG ACACCAACAG 60

ATCCATTATT ACATGAGTTT GGTACGAAAA AGTATCCAGA TGAATATCGA TATGGTTTCT 180
 ATGCTAAGCC AACATTGAAT AGACTTAACG GGGGATTCTT TGGACAAGTC TTTACTGTTT 240
 5 ACTATAATGA TAAGTATGTA GTTGTAAGTT CATTAAATGT AAAAGGAAAC AATGAAGTTC 300
 GAATCAAACA TATTTATAAT GATATTTTAA AACAAAATAA ACCTTACAAT ACGAAGGGTG 360
 TTATTGTTCA ATAATTAATA TAGAAGATAT AACATGTATA TGGCATTAG GCATCGACCT 420
 10 TATCTGACCA GTATACGAGT TATATCTTCT TTTTATAGT GGTAAGAGT TTAAAGTATA 480
 AGGTTGAAGA AGGATGAGTT TAAAAATATG TGTTAACTGA TAAAGGGGA ATCATTG 540
 15 TGAGTTGGCA TCAGCTATAA nTGAATGAAG ACGAATACGT TGGTCCATGC GTGGTGTGAn 600
 TGTCTTCTAA A 611

(2) INFORMATION FOR SEQ ID NO: 1089:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 644 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1089:

CTGGATAATC GCTTTAATAT GATCAATTTG CTTTCACTC ATATTAAGCT TATCTAGCCC 60
 30 TTTCAATTTA TTCATACCTG GAATCATTTT CATAATATCA TCCAGTGGTC CTAGATTTTT 120
 CACCyGwTCA AGTTGTTCTA AAAATCATC TAAAGTAAAC GATGACTCAC GCATCTTTTT 180
 CTCTAAATCT TTTGCTTTTT CTGTATCCAC ATCTTGTTGC GCTTTTTCAT TTAAACTTAA 240
 35 CACATCACCC ATACCTAAAA TACGTGATGC CATACTTCA GGATGGAATA GCTCTAAACC 300
 ATCTAACTTT TCACTCATAC CAACAAATTT AATTGGTTTT TGTGTCACCG AACGAATAGA 360
 40 TAAAGCTGCA CCACCACGTG tATCACCATC TAATTTAGTT AAGGLAACAC CTGTGACATC 420
 AAGTTGATCG TCAAAGATT CTGCAACATT GACAGCATCT TGACCCGTCA TTGAATCGAC 480
 AACTAACATA ATTCGTTTG GTTTAGCAAT GTCTTTTACT TCTTTTAATT CGTTCATCAA 540
 45 TGCTTCATCG ATGTGTAATC GACCTGCTGG TATCAATGGA TTACAAAGTC TAAATGTTCT 600
 TCTTTAGCAN GTTTTAATGC ATTAGTTACA ATTTGGTGTG GGCC 644

(2) INFORMATION FOR SEQ ID NO: 1090:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 207 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1090:

5 GATCCTATCG TTTTAGTGCA TGGTTTCAAT GGGTTTACAG ATGATATTAA TCCTTCAGTG 60
 TTAACCTATT ATTAGGGCGG TAATAAAATG AACATTGCGC AAGATTTAGA AAAAATGGTT 120
 ACAAAGCTAT GAAGCAAGTA TAAGTGCTTT TGAAGTAAC TATGACCGCG CATTGAACTT 180
 10 TATTATTATA TCAAAGGCnG TCGnGTA 207

(2) INFORMATION FOR SEQ ID NO: 1091:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 212 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1091:

20 GACTCTAGAG GATCCCCGCT CAAATGCCAT TTGGTGATAT TCGTGAAGCT GCTGGCGAAA 60
 AAACATTTAA TCCAGTGGA GATCGTTTCG TAATAAAAAA ATAATCAATA TATTGTTAGG 120
 25 AATGGGACAG AAATGATAAA GATTCATTAG TTATTCATTA TGTGTAGTT CTTACACATT 180
 AGCCGCTGCT AATnGCAACT TAAGAnTAGG TT 212

(2) INFORMATION FOR SEQ ID NO: 1092:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 208 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1092:

40 TGGACCTGCA ATTGGTTGCC AGCAGCAACT TTAGCTTCAT AGCTAGCTTT GTGAGCAAGT 60
 GGTAACCTG GAACGATATC ACCAATTGCA TAGATATTGC TGATAGACGT ACGGCTTTGT 120
 TTATCAACnn CTAATAATCC ACGTCAGCGA ATTTAACACC TAATTCTTCT AGGCCTAATT 180
 45 CGTCGTGTTT GGACGACGAC CTACATTA 208

(2) INFORMATION FOR SEQ ID NO: 1093:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 902 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1093:

5 CTTAAAAGTT GTnTCAATAA TGATTGTATT AGTTGCATGC TTATTATATA TGAATCATAG 60
 TATTGATTTA CCTACCCTTA TTATGATTTC AATTTTTTCA TTTGTGATAT TTGATAGTGT 120
 TGAAAATATT AATAGTGCTG CACACGTACT TGAAATGATA GATATGACGA TAGATGATAT 180
 10 TGmAAAGATA AAAAATGCTC CAGAACTGGA TGAGAATGGA AAAAATTGTA CGATTAAAAA 240
 TGAAAATATC GCTTTTCAAA ACGTGAACCT TTCATATGAT GATAACAAG TGATAAAGAA 300
 TGTGAATTTT GAGATACCTA CACAAACATC AACAGCAATA ATTGGACCTT CAGGAAGTGG 360
 15 GAAATCTACA TTATGTCACT TACTCTTGCG CTTTTATGAT ATCGATGATG GAAATATTCG 420
 CATCGATGGT GTTGATATTA AAGATATGAC ATTAAGTACG TTAATGTCGA AAATTAGTGC 480
 AGTATTTCAA AAGGTGTATT TATTTAATGA TACGATTGAA AATAACATAT TGTTTGGCAA 540
 20 TCCAGGTGCA ACGAAAGAAG AAATTATTCG TGCCGCGAAG CAAGCATGTT GCCACGACTT 600
 CATCATGTCA TTACCTGAGG GATATCAAAC AATGCTAAAT GAAAAAGGTA GTAATTTATC 660
 TGGCGGAGAA AAGCAAAGGA TTTCTATTGC TAGGGCGATA TTAAAAGATG CACCAATAAT 720
 25 TATTTTAGAT GAAGCAACTG cAAGTATTGA CCCTGAAAAT GAACAGCTGA TTCAAACGGC 780
 AATTAATGAA TTAAGTAAAG GCAAAACAGT AATTACAATT GCACATAAAC TTGAAACTAT 840
 30 TAAAnATGCA GATCAGATTA TAGTGCTCAA TGnAGGTGAA ATAATTCAA AAGGTAGTCC 900
 TG 902

(2) INFORMATION FOR SEQ ID NO: 1094:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 632 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1094:

45 ACGCAAGGGA TATGATTAAA CACCATAATA TACAATCACC TTTTTTATAC ATGAAAGCAC 60
 TAATAGAATC GATTCATTTA AACATAAAAC ATGATTTTAA CCAGCAAGAT TTGATTGAAA 120
 TACCAATTGT GTATGGTTCG AAATATGGTC CGGATTTAGA ATCACTTTTA AAACATTACA 180
 50 AAATCAAGCT AGAACTTTT ATTGAATTAC ATTCTAAGGC GCAATATTTT GTTTCGATGA 240
 TGGGATATTC ACCTGGGTTt CTTATTTTAA CTGGATTAAA TAAGAAATTG TATATTAATC 300

55

GCGGTATTGT AACTACGGAT ACAATTAATG ATTGGTTAGT TATTGGTTAT ACACCATTAT 420
 CACTTTTTTAA TCCGAAAGAA TCAGATTTTCG CACGCTTAAA GTTAGGCGAT AATATTAAAT 480
 5 TTAGACCTAT CAATGAAAAT GAATTAGAAG TAGGAGCGTT TAAAGATGTC AATCATAATT 540
 GAAAAAAGTG GCTTATTCAG TAGCTTTCAG aCTTTGGCAG AAGGGGATAT GAACATGATG 600
 GTGTAATTCC ATGTGGTGCA CTTGATACTT TA 632

10

(2) INFORMATION FOR SEQ ID NO: 1095:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 212 base pairs
 15 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1095:

GCAACACGAT TGCTAAAAAA TGATATACTT ATTAATTGTT GCAACACTTG TTTATGTGAT 60
 ATACTGATCA TGGCATTGTG GCGGATTGAT GACCATACAC ATGACGTGAC AATGATAAAT 120
 25 CATTGCCACT, ACTTGTTAAA ACAGCAATTT CTCAAGTGGA AGCTGGTGCh ATATATTnCG 180
 CCAATAATAT GATGGATGGT TTGTGCTGA AA 212

30

(2) INFORMATION FOR SEQ ID NO: 1096:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 466 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1096:

40 AATGAAGTGG GTATTnCTTT TGCCATGCTT GCCCACTGCC ATGCATTGTC TGAAGAATGT 60
 TTTTGTAGAA CCGgCTTAAG AATTGAAGT TTGTAGTAGT ATCTAAAAAG ATTACTTGGT 120
 GATACTCAGG ATGAATTCCT TGTTcATTAT TTTCAGCTCC TTTGCCCTGA ACCATCTGGA 180
 45 ACAGAGTTAT TTGTGAGTTT TTACCCAATA CTGTGTAATT ATAAAGGTAA TTAATTCAAA 240
 ACGCAAGCCC CTATATAATG TTTATTAAAT TATAGGTCGA CCCGTTTTAG TACTTTCTTC 300
 TGCAGACTTT TGTAGCTGCT TGAAGAAATC TtCATTATTC TTAGACCTTT TAAGTTTGCG 360
 50 AATAAATCTT TCAGTAAAGT CAGTTGAGTC AGTGAATAGA TTTCTTAATT GCCATAATGT 420
 GTCTAATTCA GATTACTTA TCAACAATnC TnCTTTACGC GTTGAA. 466

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1097:

TTTTGATTCA ATAGGCGGTT CCGTGTTATC ACTGACAACT TTAGTTGTAG CTTCATCTTT 60
 ATGTATTTCT TCGTTAAATC CTTCAGTTT TTAGTCGTGG GATTTTAACC TCAGGATGTT 120
 CCATCATGTC TTGACTATCA AGTCCTTnTA CACGTGTCTT TATGTGATGC TTGATTGCGT 180
 nCCCTTACTT TTTGAATAGG GGTAGTAA 208

(2) INFORMATION FOR SEQ ID NO: 1098:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 571 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1098:

CCTTCAACGA AAATAGCATA TTGACCATAA ATATTATGAA TTGTTTGTTT AACATCTTCA 60
 GTTATTGGAT CTTGACTCAA AAATACATTA AATACTTCAA TACCAAATTT ACGTGACATT 120
 TCTACAGCTT CATACGTATC AATAATACCA TCTTGACTAT AATTAAATGC AGACGGTTCG 180
 CCGTCTGAAA ATAnGATTAA AAATCGTTGA TGTTGATTTT GACGCATTAA TCGTTCACTT 240
 GCAACTCTAA TAGCAACACC ATCAGGATTA TCATCTTGAG GTTCAAGTGC CATAATACGT 300
 GGGCCATCTT TTTCAAAGGT TGAGTAATCA TAGTTAATAA TTTCATTAAT GATATTTGGT 360
 TGTGCATGCT CGTCTGAATC AAAGGCATCC TCACTGAATG ATAAAATTTT ATGTTTAATG 420
 TTCAATGCTT TTAACGTCTC GTGGAATAAC ACTACACCTT TCTTCGTTTC AGCCATTTTA 480
 TCATGCATAC TTGCTGgACG CATCAATTAA CAATGTAAAT GTnCGCATCA AATGATTnAC 540
 TTAAATCTTG TTTTTTGTA AACAATTTAT A 571

(2) INFORMATION FOR SEQ ID NO: 1099:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 214 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1099:

5 TGCAAGTGAA CAAGTCATTA ACTTGTTTAA TACACGTGGG GCACAAGTCG AAGAAGTTTG 60
 GGTGAAGGGC CATGAAATTA CAGAACTGG ATTAACGGCT GGTCAACAAA TACTTGGGAA 120
 ATAACAGTTC TATTAAGAAG CGGACAGATG GAAAAGATTT TACCTTTCAT CTGCCGCTTT 180
 10 TTGATTTGAA GGCCGGACcA AATTTACCAcA GGAA 214

(2) INFORMATION FOR SEQ ID NO: 1100:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 1158 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1100:

20 CACCGTATTG TACAAAGAAT TTAAAGTTTT CGATTCTTn TCTTAATTIA TCATCAGCTA 60
 TCGGTTTGGT TGGGATAATT TTATTATGTT CCATTTTAC TGGATATTCT TTTTCTTTAC 120
 25 TACGTGAGTA TCCTnTTTTA TCTTCCCATATA ATTCCCTCAC AATAAATTCA CCTTTTGTCG 180
 TTCTGGTATT GCGGTCTACA TACAGGACCA ATCCTCTGGA TTCCaTATTc TTTCCATTTG 240
 TTTCTATTGT CATTTTAGAG TGAATTATCC aAGTGCCTTT GTCGCCCTTT TCAAACCTCTT 300
 30 GATCTCGAAA GCCTTCCTTA TCATAAAAGT CTTCAAGATT TTTAGTTGGA TACAGGCTCA 360
 ACGTTTTGTT GAAGTTTTCT TTAATTTGTT TATCTTTACT GCTTTCTTCC TTTTTCATTC 420
 35 CGCAAGCCCC TATAAAACA GCTAATAATA CTAAATACAG CCAGAGCCTC AACCTTTTTG 480
 AGTGAATCAT CATTTACAT CCCCATTTTT ATTTTGGATG CAAAGTTAAG TtCAAATTCC 540
 AATATTAATA ATATTAGAGG aAAtTTtATG ATTAAaTCCT CGTGGATTtT AATTGTTAGT 600
 40 TGTATTTTTT ACATATCAAC ATCACAAATT ATTTAAAAAG CGCAAATATC TTTATAATTT 660
 TTATTGGCCT AACCAACTAA TTAATTAAGA TAAATTGCGC TTATATTAA AATAATAACA 720
 CTGAACTCA ATGTATTTAC TTATTAAAT TGATGTTTAA ATACAACTTT ACTAACATTC 780
 45 ATTTTTCGGT TTACATTAAT TkGTTAGATA ACGATATATA TCAtCTCTTA CAGCTTTATC 840
 CAGTGCTAAA TCCATCGTTA CTACGTTTGA ACCATTCGGC ATTTTATCTT GCTTTTCTGA 900
 50 ACCTTCTATA TATCCAGCAG TTCCTAAATA ATAAAAATAT ATACCATCAT CATCTTTTTT 960
 CTGTACAAAA ATATACATTT TAATTCCTTT TGCTCTATGG GACAAAATTT TCTGAACTTC 1020
 TTTTGATTCT AGCGTTCAT TAGATTTAGT AAACCATTTT AATTCATCTT GACTTAAAAA 1080

TGGCATTCTCT TGGCTTTT

1158

(2) INFORMATION FOR SEQ ID NO: 1101:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 353 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1101:

CCAAATGCTA TAAATGCAAA AATnGCTAAT GCGATATTTT TATAGAAAAT ATCTGGTCCT 60
 AAATTTTGTA AATACGTGGT AAACGTTGCG TATGACAATA AGCCAATCAC TACAGCGTAA 120
 ATATGTGAAA TAATTAAACC GAACCGTCTC GCTCTTTCAA AAACAAGTGT TGCTAAAATA 180
 ATAATCAGTA ATCCAAATGA CAAGGGTTGT CGCCAACCTTA ATGGTAAAAA CTGACCTAGA 240
 TAGCATCCTA TGCCAAATAT AATCCAATAA TACATAAAAT AAAGCCACAC TTTACTTAaT 300
 GATTGAGCAT GCTGGTTTTG ATGATAATAT GAATGATTTG TATTTTGCGA CAA 353

(2) INFORMATION FOR SEQ ID NO: 1102:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 194 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1102:

CCCACACGTC TTCCACCGGA TAGGTTTCGA GCTCGGGCGA ACCGGGCAGG GCGGTGACCA 60
 CCCCCTGCGG TGTCTTCTTG GCCAGCTTGT TTGCGAGCTT GGCCAGCGTC TTCGTGGGGC 120
 CAATGCCCAC GCAACACGGG ATGCCACCCC ATTGGAGGAT CTGGGCACGG ACCTGGCGGG 180
 CCATCGCTTC GnTG 194

(2) INFORMATION FOR SEQ ID NO: 1103:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 220 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1103:

TCGCTTTTGC TTTTGCATCA GGTATTGCTG CTATTTTCAGC ATACTTTTGT ACTTTCAAAT 120
 CTGGTGATCA TGTCATCTTA CCCGATGGAT GTATATGGGC GGTTACTTTT CGTCTCACTG 180
 5 TGGCAAATTT TGGATCGATT TAATATTGAA TTTTACAACC 220

(2) INFORMATION FOR SEQ ID NO: 1104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 235 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1104:

GGACATTGTA TCGACATTTT AAAGATAAAA GCGATTTATG CTACTACGTC ATACAAAGAG 60
 20 ACCTTGATAT TTTTATCACT CATTTTAAAGC AAATTAAAGA TGACTATCAT TCTAATTATG 120
 AGGTTATGCA AGTGTCGTCA TCGCTTTTAA ATAAGTCATA ATAAAAATCA AATAATTCTT 180
 GATAAAATGC GCTTTGGTAA AAACGTAATT TATTGTTGCC TGCTTCAATA CATTG 235

(2) INFORMATION FOR SEQ ID NO: 1105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1105:

GCGGGGGCAG GTAAATCCAC TTTGACACAT CTTATTGCAG GTGTTTATCA GCCAACAATA 60
 GGTACTATAA GTACAAACCA GCGTGATTTT AATATTGGGA TACTTAGTCC AnAGCCCTAT 120
 40 ATTTCAAGCT CTATAAAGA GGATATTACG ATGTTTAAAG nTATAGAAA TAATACTAAT 180
 GAAGAAGTGC TAGACGAAGT AGGGTTTATT AGACCAAGTG CCATCCTTC 229

(2) INFORMATION FOR SEQ ID NO: 1106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 596 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1106:

5
 10
 15
 20
 25
 30
 35
 40
 45
 50

FAATTAAAAT CAAAATTTGA TGAAGTAAAT GTAATCGCGT CAACAAACGG TAGAAAGTTT	120
ATTAATGGTG AAATATTAAA GCAATTTTGC GATAACTATT ATGACGAATT TGaAGACCCT	180
TTTTTAAATC ATGTAGATAT AGCAAATAAA CATGACAAGA TTATTATTTT ACCTGCGACT	240
TCTAATACGA TTAATAAAAT TGCmAAATGGT ATATGTGaCA ATTTATTATT AACTATTTGT	300
CATACAGCTT TTGrAAAAC TCTATATTT CCAAATATGA ATTTACGAAT GTGGGAAAAT	360
CCAGTTACTC AAAATAATAT TCGATTATTA AAAGATTATG GTGTATCAAT ATATCCAGCA	420
AATATTTTCA AAAGTTATGA ATTAGCGTCA AAAACATTTA AAAAGAATGT TGTCGCACCT	480
GGAACCATAT AAAGTTCTGG gAATTCATTT GaGATTAGAA TATGGAAAAT CaTAAAGCGC	540
GCKATTGATT AGTTTAATGA ATGATTGGAG TcTTTTGAAG ATnCGATTTA CAATGG	596

(2) INFORMATION FOR SEQ ID NO: 1107:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 673 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1107:

30
 35
 40
 45
 50

ACTAGCTAAA AATAATATTC TATCTTTTTT AACATGGGCA GTTCCAGCGG CAATTAGTGG	60
TATTTATATT AAATATATAA ACAAGGCTAC GGTAGAAAGA TTTTTTAAAT TAGTATTTTT	120
CATATTTTCT GTTTCATTTA TTTTGTAAT TTTAATACCA AAACCTACAG GTGAGATACC	180
TAGCTATATC AATTTTGGAC TTATGAAC TA AAAACGCT TCGTACCTTT CAGCATTTAC	240
TGCCGGATT A GGCATTTATT TCATTATGAA AGGTTTCAGTT AAACATAAGT GGATATATGT	300
TCTATTTACA ATAATTGATA TCCCTATTGT GTTTATACCA GGAGGGCGTG GAGGTGCTAT	360
TTTATTAATT CTTTACGGCT TATTTCATT TATACTTATT ACGTTTAAAA GAGGAATACC	420
TATCGCAGTA AAAAGCATT TGTATATTTT TGCATTAAGC ATATCTAGTG TATTGATTTA	480
CTTCTTTTTT ACAAAGGTT CGAATACTAG AACATTTTCA TATCTACAAG GTGGAACACT	540
TAATTTAGAA GGTACTTCTG GgAAGAGGAC CGATTTATGA AAAAGGTATT TACTTtATTC	600
AACAAAGTTC GTTATTAGGC TATGGGCCAT TTAACATATcA TAAACTAATC GGAAATATAC	660
CACATAACAG CAT	673

(2) INFORMATION FOR SEQ ID NO: 1108:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 506 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1108:

	TCACAGTGGC GGCACAAAAA GTAAGTGGAG ATTATTTTAA TTTAATTGAC CATAACGATG	60
10	GCACAATGAG CTTTGCTGTT GCAGATGTCA TTGGAAAAGG TATACCAGCT GCTTTAGCAA	120
	TGAGTATGAT AAAGTTTGGC ATGGATTCTT ATGGACACTC ACAATTACCG AGTGATGGTT	180
	TAAAACGTTT AAATCGTGTT GTTGAAAAGA ATATTAATCA AAATATGTTT GTCACAATGT	240
15	TTTATGGTTT ATATGAAGAA ATGAACCATT TATTGTATTG TAGTTCAGCT GGTTCATGAGC	300
	CTGGATATAT TTATCGCGCT GAAAAAGAAG AATTTGAAGA AATTCAGTT AGAGGTAGAG	360
	TGTTAGGAAT CAGTTCACAA ACACGATATC AACAACAAGA AATCCAATA TACCTTGATG	420
20	ATTTAATTAT CATTTTAAACG GATGGTGTGA CTGAAGCTAG AAAAGTGAAG GTACCTTTAT	480
	AGATAAACAA AAACTTTGTAG AATATA	506

(2) INFORMATION FOR SEQ ID NO: 1109:

25

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 552 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1109:

35	TGATATTGAA AAAGATGTCA TTGTTAAAGC ACAAGAAAGT AAAATTGCTC AAGTTATCAC	60
	GAATTTGTTA ACGAATGCAA TTAACATTC TTATGAAGAT GGAGATATTA ACGTTCGTGT	120
	GTATCGAGAT GACTTTCGTG TCATTTTCGA AGTACAAGAT TTTGGTATAG GTATTAAATT	180
40	AGAAGACCAA CAACGTATTT TTGAAAGATT TTATAGAGTA GATAAAGCAC GAAGTAGnAG	240
	ATTCTGGTGG GACAGGTCTT GGATTATCAA TTACAAAGCA TATAGTAGAs GCCCACCAAG	300
	GCAATATTGA AGTGAATAGT CAAGTTGGCA AAGGCTCAAC GTTCAAAGTT ATTCTAAAAG	360
45	ATTATAAAGA ATAADAATTA AAGTGGTAAC AGCGCGTGTA TTTTTCACGA AGTTGTGTCT	420
	GCTTTTTTAT TTCGTTCAAA TCTCTATTGG TATTAAATTA GAGTTAATGA AGTGAATAA	480
50	ATTGCGTTGT CAATTTTGTG ATTTGTATAT GTCGATTGT AAGTTATAAG TAATAGATTG	540
	TTGTGAATCA nG	552

(2) INFORMATION FOR SEQ ID NO: 1110:

55

- (A) LENGTH: 244 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1110:

10 AACAGATCAA CAATTGCAGA TCAATTACTC GTCATATTGA TCAATTATCT CTGACAACGA 60
 GTGATGATTT ATTACGTCAT TGATTGATCA TCACAAGATA AGTCGCTATT GATTTCTCAA 120
 TTTTACAACG AATTAGGAA AGCTGnAGCA GATAATTGGC TAAGTTGGAC GATAAAGGTT 180
 15 ATCAATCGCC AATCGTTGCC ATTGAAGAAC TTTTGCTCAC TGGCGCACGC TTCAGATGAT 240
 ATTT 244

(2) INFORMATION FOR SEQ ID NO: 1111:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 825 base-pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1111:

30 AGTATTGAGT GTTTTGATTT AAGAGACTTG GAAAGTGAGT CAATAATCAA AATCTTACTC 60
 AATACTTTTA TTTTGGTACT CGCAATTTAG ATTTGTTAAC TAATTTTTTC GCTTATTATA 120
 AAACGAAACA ATCAAGTAAA CGATAAGGCC TACAAAGATA CCCAATAAAA TAGATAGTAC 180
 35 TGCCGTCACT ATTAATGGTA ATTTGAAAAA TATTTGTAGG AAAATmCCAA TGATAATTGC 240
 GATAATTACT GCAATaATG TGA CTGTATT TTCATTGAA ATGTTCATtT ATttTCACTC 300
 CTtAACAAATA ACATTATATC ATGCTATAGC TTTCCAAAAT ATTGAAATAT GTAGATATGG 360
 40 CTATTGACGA TATTTCTTAA CTTTTATATG ATTAATCGGA ATGAAAAAAG AGAAGTAGGT 420
 GGCAATATGA AGTCAAATAA ATCGCTTGCT ATGATTGTGG TAGCCATCAT TATTGTAGGT 480
 GTATTAGCAT TTCAATTTAT GAATCATACG GGTCTTTTCA AAAAGGGGAC GAATCATGAA 540
 45 ACTGTACAAG ATTTAAATGG TAAAGATAAA GTACATGTTT AAAGAGTTGT GGATGGTGAT 600
 ACATTTATTG CAAATCAAAA TGGTAAAGAA ATTAAAGTTA GGCTTATAGG GGTTGATACG 660
 CCAGAAACGG TGAAACCGAA TACGCCTGTA CAACCATTTG GCAAAGAAGC ATCAAATTAT 720
 50 AGTAAGAAGA CATTAAACAAA TCAAGATGTT TATTTAGAAT ATGATAAAGA AAACCAaNTC 780
 GCTATGGTAG ACCATTGGCG TATGTATGGA TAAGTAAAGA TCGTA 825

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1112:

```

10  TCCCTATAAA TCCCCACAAA AAAAGGGTAA ATCCGAACGT CAAATTTTCAT CTCTTTATGC      60
    CAATCAACGA TCGATTTTACC TTTTAAAAAT TAAAATGTAT CTGTAATACT CTTAGCAATA      120
15  TTTGTATATG CTGCATCTTC CATAGGTGGA TTGTGTAATC CTGCACGCAT GTCACGATAG      180
    TAACGTTGTA ATGGTCGATT CATTTCTAAG CTTTTAGCAC CAACTATTCT CATCGCTAAA      240
    TCAATCACTT CAAGCCCTTG ATTCATTACC ATCACTTTAC TCGCACTTGT TGGATTTCTT      300
20  ATCTGACTGT CATCTTTATA TTGTTGATAC CCTTTTGCTG TACTCCATAA AAATTGTCTC      360
    GCACATAACA ATAGCGTTTC CATTTTCCCT AAATTTTGTT GCACAGTTGG CAATGTTGCA      420
    ATCGTTCCTT CAATACTATT AGGGCTATGT TGGATTGCAA AGTCTACAGC ATAATTTCTA      480
25  GCAGCTTGTG CTATACCTAA ATAACAGCTT GGTATATGCA ATATCCAACC ATTTGGCGCT      540
    TTACTTTTCT CTGTTTCCAC TAAATGTTTT AATGGTACTT TTACATCATT TAATATTAAA      600
    TCATGACTTT CTGTCGCTCG C                                          621

```

(2) INFORMATION FOR SEQ ID NO: 1113:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 619 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1113:

```

40  GTACCTTGTT TTTCAGTGTT AATTTCCCAA ATATAACCTG GTGTAATACC GTATTGTTTT      60
    GAGAAGTACT TGAAGGCACC TTCACTGTa ATCATGGCAC GTTGTTCTTC TGGAAATGTCA      120
45  TTAAATTTGT CTTTACTGTC TTTACTGTCA TTATTTAATT TTTCCAATTG AGCAATGTAT      180
    TTGTTACCTT GCTTTTCATA ATCTGCTTTA TGTTTTTTGT CGTTATCGAT AAATGTTTGT      240
    TGAATTGTTT TTACGTAATT AATACCGTTA TCTAACTTA ACCATGCGTG TGGATCTTGT      300
50  TTATCTTTGT TGCTTCTTC ACCGTTTAAA TAGATAGGTT TAACATCTTT TGATACTGCG      360
    ATAACCTTTT TATCTTTTAA TGATTTACCA GCCTGTTCTA AGGCTTTTTC AAACCAACCG      420

```

ATATCyTTAG GTTtAACyTC aTawTCaTGA GGATCTkGAC CAACAGGTAC AATACnATGA 540
 ATATCGACGT TGTCTCCACC AACATTTTTTA GCCATATCAG ATAGAATTGA ATTCTGTCGnT 600
 ACTACTTTnA ATTTGCCAT 619

(2) INFORMATION FOR SEQ ID NO: 1114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 549 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1114:

GAAAACCGCA CGACAATGAA AAAAGATATC CGGTTAAAT GGTGATAAT AAAATCATTC 60
 CAACAAAAGA AATTAAAGAT GAAAAAATAA AAAAAGAAAT CGAAAACTTT AAGTTCTTTG 120
 TTCAATATGG TGACTTTAAA AATTGAAAA ATTATAAAGA CGGAGATATT TCATATAATC 180
 CAGAGGTGCC GAGTTATTCG GCTAAATATC AATTAATAA TGATGATTAT AATGTAAAC 240
 AATTACGCAA AAGATATGAT ATACCGACGA GTAAAGCTCC AAAGTTATTG TTAAAGGTT 300
 CAGGGAATTT AAAAGGCTCA TCAGTTGGAT ATAAAGATAT TGAATTTACG TTTGTAGAGA 360
 AAAAAGAgGA AAATATATAC TTTAGTGaTA GCTTAGATTA TAAAAAAGC GGAGATGTAT 420
 AATCATGGCT CAATCAGAAT ATGAAATCAA TCCCGGAAAA AGAGAGTGAT GAAATGATAA 480
 AACGTGTAAA TAAATTAGTG CTTGGTATTA GTCTTCTGTT TTTAGTCATT AGTATGCACT 540
 GCTGTTTGT 549

(2) INFORMATION FOR SEQ ID NO: 1115:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 317 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1115:

TGTTATTGCA ACTATTAAAA ATATnAAAA CAAAGACTTA ATTGATTATT CTTTATATAT 60
 TAAAAAAGGT TTAGTGGAGC CATTTATTGT ATTCTATAAT CAAGAAAATA TATTCTACAT 120
 TAGCAATGCT GTTATAGATA TCATCAGTGT ATCCAAAGAT ACTCTTGATT ATATAAAGTC 180
 AGAATTCAAT GATCACTGTG tTCGCCATAC TGATTTTATA GCACAGAGAA AAGCCGTAGT 240

TTATCATCTT CGATGAA

317

(2) INFORMATION FOR SEQ ID NO: 1116:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 400 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1116:

15	ACTATGACTG CAATGCCAAC AATGAATGTT ATTGCATAAA GCATATGCAT TTAATCAAGC	60
	TTATGTTATT TAGCAAGGGC AAGCCACTTC TCGATTCATG GCTTGCCTTA TTTTATTGTA	120
	TATTTTTTAAT TATCGTTGCT GGGCCCTTGA TTGACAATCG TATTAAAATG GCCTTATTAA	180
20	GTCAACTTTG TCTATACGGT TTGGAATCTT CTACCCAATG TCTTATAAAA GACAATCCCG	240
	CACCTGAAAC ATAACATCATG AAnATAAGAA nATGGTATAC CATTAAATnTG AATCCATTTC	300
	CAnTTTATAA CGTTGTAAAT GACACATAAT TAGAATCATA AAGTTTTTCG TGGGATATTG	360
25	AAACTTTACC CATACGAGAC ATCATGGATA AAAGCGAGGC	400

(2) INFORMATION FOR SEQ ID NO: 1117:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 696 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1117:

	TGCATATAAT GTTCAAACGG CAATGATTCC ACTGTATAAA CCTCTGTGTCG TGGATGATCT	60
40	TTAATTTGTT GTTTTAATAA CTGAACTCTT TTTTCATGTT CATAATGAAC ATAAATAAAT	120
	GCACTAAGAT ATATCACAGC TAAAACTAGT GATGCACCTT TAATAAAATT AACATTGATT	180
	GACTTATATT TCCGAAATTC TTTTAATAAA ATAATTAAAA TGATTACATG TATCGTATAA	240
45	ACAATCAAAA AATTACCTGG TTCTATTGGA GTAACAAATGA CTAGTGTCGA CGCCGAAACA	300
	CATATTGCAA TAAGTAAGGA ATATAAAGTG ATTTGTGTTT TACGATCATT AATAGACAAA	360
	TAAATACCTA CAAATATCGA AAACGCAAAG TAACCACATA CAATTACGTT CACAAAACCA	420
50	ACCAATCCAA TATCTGTATT TTTATTTAAT AAAAAGTATG TGTAAGATAG TAAATAATAA	480
	AGCGGTAAAG TGATAAATCC TATCATAATG ATACGACGCT TTAATATTGT TAAATGTACA	540

ATCTGACTGA AAATGACACC ATATGGAAGT GACGTTGATA TCATTCTGC AAATTTTGAA 660
 AAGATACCTT GATTATTGA AACTTGTTGA TATTCT 696

5 (2) INFORMATION FOR SEQ ID NO: 1118:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 223 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1118:

GCAGTGACAA TCAATAATTT GTAAC TAGAA GATAATAAAG AGAACGCTCT ATAGAGACGA 60
 ATTGAAGGTT TGATTTTAAT GTCTGTTAGT AAGAATCATA TCAATGAGAT GCCTATAGTA 120
 20 CTCAGANTAT ATTAAATTAA AACCGTCATT AATTGTTTTT CCGAAAACCA TTTGTAACCA 180
 TTTnAATGTA GTTGGACCTA CCTACGTTCT CCAATAATCC ATT 223

(2) INFORMATION FOR SEQ ID NO: 1119:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 265 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1119:

35 CTGCTTTACC TTTTCCAAC TTAGAAATTG CTACATAGGG GCCTTCTGCT TTACCGCCCC 60
 AATTCCAATT CCACATTTTC AAGAAATAAG GAGGCAAAAA TGTT CATATA TGAATTGGAA 120
 TTATTTGTTT TCTTATTAGG CCGAGATGCG CCGCGTGCGG CTGCTGGAGA TGGCGGACGC 180
 40 GATGGATATG TTCTGCCAAG GGTGGTTTG CGCATTCACA GTTCTCCGCA AGAATTGATT 240
 GGCTCCAATT CTGGAGTGG TGAAT 265

(2) INFORMATION FOR SEQ ID NO: 1120:

- 45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 329 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1120:

55

CATGACATTA CATGCTTATT TAGGTAGAGC GGGAACAGGT AAGTCTACGA AAATGTTGAC 120
 CGAAATAAAA CAAAAaATGa AAGCAGATCC GCTTGGAGAT CCAATCATT TAATTGCGCa 180
 5 ACTCAAAGTA CATTTCAATT AGAAcAAGCC TTTGTCaATG ATCCGGAATT AAATGGTAGT 240
 TTAAGAACAG AAGTGTGCA TTTTGAACGA TTAAGTCATC GTATTTTcCA AGAAGTTGGT 300
 AGTTATAGCG AACAAAAGTh ATCTAAAGC 329

10

(2) INFORMATION FOR SEQ ID NO: 1121:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 252 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1121:

ACACCTCTTT TTTAATATAT GTGAACATA TATAGTGAT AGAGAGGTTA CTTGTTACTC 60
 AATATAACA AAAATCAACT TTGTCAAAAT AAATGTGACA AAATTAAATA AAGTGCATC 120
 25 AATGTGACAG TATAGATCAT TTTGCAAAAA GTCAAAACAA AAAAATTGTT TTAGGGATTT 180
 TTCAAAATTT TCATTGTGGA AAATGATTG nCAAAACAA ACAACACnTT GTGAGCAATG 240
 AAGCTTCAAT GG 252

30

(2) INFORMATION FOR SEQ ID NO: 1122:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 base pairs
 35 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1122:

nAATAATAGG CGCCACCTAA TAAACCAGCT GGAATGCCTA TCATTGGTGT TGTGAATGAG 60
 CTTAATACAA TAACAAGTAT TGTAAAGCA ATGACGTTAT ACCAAGTTAC AGTCAAATTT 120
 45 TTCAAATCCT CATATGATTG TTTTACTAAT TCTCTAAATT TCATGATTCA ATCTCTCCTT 180
 TTCTTAAATC AATTAATAGT TATACTGCAT TGCTTAAAT CAAnATATCT AATGAAGCAT 240
 CTTTTAATCG TAAGTGTTAA ATGGTTTTCT TAATCAATTT GTTTGAATGT ATAAAnAGCAA 300
 50 TTAAAGCAAA AGTAATGTAA ATGAGGGTGT ATACAACAAT TACTACAATA ACCGGCATGA 360
 AGGAAATATT ACCCATCAAT TTCATAAATG CGATTGCGGC 400

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 270 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1123:

AATACAACCT TTTATCATAG TGTAATGTA TTCTACCAGT ATTGAGAAGT TTTTCATATTA 60
 TTCAATACCT GAAATCGCCA TAGTAATATT ACTAAATGCA CACTGCATAT GTTGTTTAAAC 120
 AACACAACCT TAAAAATATA TTCTAACTCT ATCTACGAAT GTACTTAAAT ATCATAACAA 180
 TCTATTCCAA ATCAATTAC ATTATTAGCT ACCTTAAAAA CCAAACCGAG GCCTTAGGCC 240
 TCGGTTTAAA TATATATAAC GAGCGACATG 270

(2) INFORMATION FOR SEQ ID NO: 1124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1124:

GGTCTTCTGT ATCAAATTAT ATTATGAGAG TTAGACTCAG TATTAATGAA ATCGTTGATA 60
 AAGATCAAGT GCTTGCATCA CAAATAGGTG GATTACTATG ATGTTTAATC AAATTAATAA 120
 TAAAAATGAA TTAGAAGAAT CATATGAATC TGAGAAAAA CGTATAGAGA ATGAACTGCA 180
 AAATTTAAAT GAACCTTAGGC ATAGAAGTGG AAAAGAAAAT GAACGTAGTT ATGATGTTTT 240
 TCAATATTTG AAGCACGAAA TGAATTATAG TGAAGATGCC CAAAGGAAA TGACGAGAAA 300
 TATAGAAGCG TATGAGCAAG AAATCAATGA GATAATTAGA AAGCAAGAAT GGAAATTAGG 360
 AAGAATATAA AGAAGACTTA AAAAAGTCTT ATGAAAAGCA 400

(2) INFORMATION FOR SEQ ID NO: 1125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1125:

AAATATATC GATCATTGAT AATATGTGCT TTACTATTCC AAATAACTTT CAAATTTTGA 120
 TAATATTCAT CAATACTCTC AATATAACTA ACCAGTTTAT ACAACCATTC GCCATCGACT 180
 5 TTCACATATT TTTGAATACT ATTTCCAAyT TtAAGACGTG TTGGTTCATT TACAAAATGC 240
 CCCAAATTAA TACCACTTAA TAATCCAAAT GGAGTTGGAC GTGTTGACAT TCTAATTAAA 300
 TATTTTAATA AACTTTCTTT AGCATCTCTA ACCTTCTTGT TATCGCCATC AAAATTTATT 360
 10 nCnTGTAAG TACAATA 377

(2) INFORMATION FOR SEQ ID NO: 1126:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 227 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1126:

GTTACATTAC TGTTTCATGTC CACAACGTTT GTCAGCGATT ATCATCATTG TGTATATCT 60
 25 TATTTATAGC AAGTGCATGG TCTGGATAAG TAAAGGTATT CATACTTAAG TAACTTGAAC 120
 ATTGGTTTAG GACTATTTTA ATGGTnCTGC TtNAATTGTT GGCCAACTGG TTCTTATTTT 180
 AAATATGTAA CTAGCTCTAC GGGTAGTTAC TAAACACATC CTGGTTT 227

30

(2) INFORMATION FOR SEQ ID NO: 1127:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 717 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1127:

CCAATGATAC CGTTATAGCA TGTTAATGTT TTTGTACGAA TATAATCTAA AACTTCGAGT 60
 TCATCACTTA ATTTCCACGT TTCTGCTTTG TTTAACCACA TTAAAGGAGT ATGAATGnCA 120
 45 AAATCTTTGT CCATAGCTAG GCTTAATGTT ACGTTCATTG ATTTTATAAA ACTATCGCGA 180
 CAGTCTGGGT AGCCTGAAAA GTCTGTtTCA CATACGCCTG TAATAATATG CTTAGCCCCA 240
 ATTTGATAAG CTAGAGCGCC TGCAAACGAC AAGAAAAGTA AATTTCTAGC TGGAAACAAAT 300
 50 GTATTAGGTA TACCATCTTC ATTATTAGTA ATTTCCATAT CATGTTGTGT TAATGCGTTT 360
 GGAGTAAGTT GTGATAATAA TGACATATCT AAAACGTGAT GTTTCATTCC TTGATCTTGT 420

55

AGTTCAACTT CTTTGAAATG TTTTTTTGCA TAAAAGAGAC ATGTTGTACT GTCTTGACCA 540
 CCACTAAAGA CAACGATGGC TTTTTCATTA TTTAATACAC TTTCCATTTT GTAATTGCTC 600
 5 CTATCATTAA TAATATTAAT AAAGAGGTTA ATGGCATTGA TAAGCCCGTT TTTAATTTAT 660
 AAAATAAAAA AAGCCnATCT CCATAAAAGA TAGACGAnAG AAATGGGTTG CTCCTAT 717

(2) INFORMATION FOR SEQ ID NO: 1128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 595 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1128:

20 GATTATTGCA TTGACAACTT ATATCTATCT AACGTTAGTA GCATTTAATT TCATTTCCGG 60
 GGTTCCTAACT TTACCAGGAT TAGCGGCGTT GGTATTAGGT GTAGGTATGG CTGTAGATGC 120
 CAATATTATC ATGTATGAGC GTATTAAGGA TGAACCTCGA ATAGGTAGAA CGATAAAGCA 180
 25 AGCCTTTTCT AAAGCAAACA AAAGTTCATT CCTAACAATT TTTGATTCTA ACTTAACAAC 240
 AGTTATCGCC GCAGCAGTAT TATTCTTCTT CGGTGAAAGT TCAGTTAAAG GTTTCGCGAC 300
 AATGTTATTA TTAGGTATTC TAATGATCTT TGTTACAGCC GTGTTCTTAT CAAGATTCTT 360
 30 ATTATCATT A CTGTTTCAT CAAATATATT CAAAATCAA TTTTGTTTAT nTGGTGTTAA 420
 AAAGAATAAA CGACATGATA TTAATGAAGG TGTAGATGTT CATGACCTTA AAACCTTCATT 480
 35 CGAGAAATGG AATTTTGTTA AATTAGCTAA GCCATTAATT GGAGTAAGTA TTTTAATTGT 540
 AGTGTCGGTT AGTTATTCnT TATATCTCAA GTAAACTTAG GAnTGAnTCT CAAGT 595

(2) INFORMATION FOR SEQ ID NO: 1129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 371 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1129:

50 ATTTTAACTT TCTAAGAACA AGTTGTCAAG ATTTTTTCTT AACATATCAT GTTCTGTTAC 60
 AATATTAATA ATGAATACCG AAAGGACGAA TTGCAAATGA AAAGCGTTGG CTTAATCACA 120
 GAATATAATC CCTTTCATAA TGGGCATCAA TATCATATTA ATCAATCTAA AAAACTTACA 180

ATCTATAATA AGTTACTCGT GCAAAATGGC TTTCCACCGC TGTTTAGTTT CGGACTACCA 300
 GCAACTGCAG TTATCACTGG CGnCATTTC GACTGCCGTT AAGCGCGGTT TTGGTGCGTA 360
 5 CATGCnTGGn G 371

(2) INFORMATION FOR SEQ ID NO: 1130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1130:

ATGATGATGA GATTATGACT TTGCGAAAGA TATCGGTGCG CCTTATGAAT TTAAACAnA 60
 nTAAAGACCA TGGTCGTTAC CGGTAGTTAA CTTTCAGCT GGTGGCGTTG CGACTCCTCC 120
 AAGATGCTGC TTTAATGATG GAATTAGGTG CTGACGGTGT ATCCGTTGGA TCCGGGAATT 180
 TTAAACCAG AAGATCCCGG 200

(2) INFORMATION FOR SEQ ID NO: 1131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 203 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1131:

ATGAGCTAGT CGATTAAAGT TAATATTTGA AAATGAAGCA GGGCATTAAA TGCAATAAAT 60
 TAAATAAGTT GTCATTAAAG CATTATAAT AGAAATGATT TTAACAGGAA AAAAGTGATG 120
 AATATTTGGA AAAGATATAT ATCGTGCAC TCTGAGAG ATTAGATTG GGAACCAAT 180
 TTATCCnTGA ATCGAGATGT nGC 203

(2) INFORMATION FOR SEQ ID NO: 1132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1132:

ATCTAACAGT ATTGAAATCT TTGTTAGCGT CAATCGACCA TGTATTTTTA AGTATTGTCT 120
 CTAACAACCTT AGCACGACGT CTTAACCCCTT TAGTAAACAA TCCTTCTTTA GntTTTAGTAC 180
 GT 182

(2) INFORMATION FOR SEQ ID NO: 1133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1133:

CAAATTTAGC CAGGnATACA AATACATATA ATAAACTGT TTGTAAGCTT ACGTTGACAA 60
 TCTGGCGTAA TTGGAAACT AATGAATTTT CTCCAAGTAG GTTTTACCCT GTAAACmAAA 120
 TAACAATTcm AATAATATGr AATCACAaaG CGACTAGAAA TCCGGTAATA TGACTAATCA 180
 TATATTCAAT GTGTaATAAT TTTAACAGCA ATAAATAGAC AACATAATAA TTTAACGTAT 240
 TAATGCCGCC AACAATGATA AATTTTAAAA TTTCAGCATG CGTTTGTGTT AGTTTCATAT 300
 GTGTACTCCT CAACATCana ATATATGCAT AACTACGTTT TCGAACATAC TCGAATATGC 360
 GAGCCAATCC GCTTCACTTC AAATATGCTT ATTTCAATCT TTATACCCTT TCACAGCAAT 420
 TTAGTCTCTT TnCCCTCATC CTTATAGGCC nTATATGTAA CCGATTATCG GGTGGACTCA 480
 TTGGC 485

(2) INFORMATION FOR SEQ ID NO: 1134:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1134:

CTTGTTTTAT CGACAGATGC TGAACACGCA GTCaAAATAA ATAAATTACA TAATAAAATT 60
 CCAATTAATC TTTTCATAAT ATCATCCTTT TGTTTTTCAA TTGATATTCA TTTTCAATTA 120
 TAAAATATTA rAAGAATTAG TCAACGCCTG TGAGTAATAC ACATCAGTAA CATTTCATT 180
 TTCATTTATG ATATTATCTA ATTATTAATT TAT 213

(2) INFORMATION FOR SEQ ID NO: 1135:

(A) LENGTH: 526 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1135:

10	TGTGACTCCA TGTATTCATG TATCATTGTA CATTGTGTGAA TTGGTTGAAA ACTGATAGTG	60
	CCCGTTTAAT TACAGGTA CT GAGGATGCTT TCAGGACATG GATTGTGTCG AAAATATGAC	120
	AATATTTTAC TATGGCTCTT CACTTCTCTA GGGAAATTGC TATCTTGATT AATCCAATCA	180
15	ACCAATTCAC CTAATGGTGT GTCATCACCA AGAAAAC TTT GCATAAATTC GTAAAAGCTC	240
	AATACCTCAC CTCGATTaAT ATACTTGAGC CATTaATATA ATACCATAAA AGAATGTGTA	300
	TTTTaCATAA AACTTTTTTaT AAAGCAATTT TATAAACTTA AATTTATTGA ATGTTTAAGA	360
20	AGTTGTGTAT AATGAAATAA ACGTGTGTA ATAGTTAGGA TGATGTTAGT GGATATCAAA	420
	CATATGAAAT ATnTATTGA AGTCGTTAAG CAnGGAGGCA TGACTAATGC TTCCAAATCA	480
25	TTATATATTG CACAACCTAC AATTAGGAAA GChATTAAAG ATATTG	526

(2) INFORMATION FOR SEQ ID NO: 1136:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 321 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1136:

35	AGCAAAAGCA ATGATGACAA AAGAACAAGC TATGTTAGCA GATGGAAGTA TTAACAAGA	60
	TCAATATGAT AAACAAC TGT TATCGAnAAT CGGAAAATCA CAATTAGATG AATTGTCTTC	120
40	TAAAGATTTA CAAGTTTTAG CTATTTTTCC GAGAGATGAA TGCAGGAACA GTTTTAGATC	180
	CACAAATGAT AAAAnATGGA AGATGTT CAG TGCAAAAAGA GTATGGCAGC AGTTTTCTCC	240
45	AGCAACTTTT CCAaATTTAC CCAGGTGTTT AACACGTCCT ATTGGATTG GGCTTAGGAC	300
	AATTATTCCA TATGGGCGGA T	321

(2) INFORMATION FOR SEQ ID NO: 1137:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 330 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1137:

ATAAATCCTT ATAAAATCAT GAGTTGGACG ACTTATCCAC ACAATTATAC TTTTTCAGTT 60
 5 CTCTATATGT CAAATACCTG TGAATATGTT GCTAAAAATA GTATAACTTT GTGTGTAATT 120
 TCTAATTATC CACAATTCTG AAAACTATAA ATGTGCATAA GTGGATAmCT TTcCTTCTAT 180
 AGAGTATCTG TTAGTGAGTG TATCAAAaCA GTTTGGGAAa TAATTtATAA AGTnTGTATA 240
 10 AGAnCTGTAT AAGGTGTTcG AACATTGTAA ACACTCATGC TTCGGACCAA ACTCATGGTG 300
 ATGTTATGAA ATTTGATTGC TCGCATCGnG 330

(2) INFORMATION FOR SEQ ID NO: 1138:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 459 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1138:

25 CTTGATAAGG GAAGATTTGC CAGCACCATT CGGGCCCATG ATACCAATTA TTTGCGCCGCG 60
 TACTGGTATC GATAAGGAAA TGTTTTTAAG TACATGCTTA TTACCTAAAA ACAGATTTAA 120
 ATCTTTTGTT TCTAACAAAC GTTTATACCT CCTAATTAAA AGTTTAGGCT AACCTAATTA 180
 30 ATTGTATAAT AAACTGAGAA TATTTATCAT GTCAAGTAAA TTCGTGATAT AATATAGACA 240
 ATGTATGTGA GGTGAAAGTA TGTTAACTGA AGAAAAAGAG GACTATTTAA AGGCAATCCT 300
 TACGAATAAT GCGGATAAAA ACTTTGTGAC AAATAAAATC TTATCTCAAT TTTTAAATAT 360
 35 TAAGCCTCCA TCTGTAAGTG AAATGGTAGG ACGTCTTGAA AAAGCAGGCT aTGTTGAAAC 420
 AAAACCATAC AAAGGTGTTA GATTAACAGA GGATGGTTT 459

(2) INFORMATION FOR SEQ ID NO: 1139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 165 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1139:

50 TAAATACTCT CGCCGGTTAT TTCGCTTCCT CTATACGGAT TGGCGAAAGA TACGGTGCAT 60
 CGGTTTCAAC TAGCAAACGC TCCATTGACA CATGCTTAGC AACTTCTTTA GGCTGTTTAG 120

(2) INFORMATION FOR SEQ ID NO: 1140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 214 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1140:

ACCTGAATGC CCATCAATGA CAGACTGGAT AAAGAAAATA TGGTACATAT aCACTATGGA 60
 ATACTATGCA GCCAGAAAAA AATGAGGTCA TATCTTTTGT GGGACATAGA TGCAGCTGGA 120
 GGCCATTATA TTTAGTAAAC TAATGCAGGA ACAGAAAAAC AAATACTGCA TGTTCCTACA 180
 GCAGGAGCTG AATGATGACA ATTCATGAAC ATAA 214

(2) INFORMATION FOR SEQ ID NO: 1141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1141:

GTTTCTAGTT GTTTATAGTT ATCGCATCTC TTGATGTAAA ATCTAAGATG TTTATGCATT 60
 TCTAGATGTT AATTCTTCGC TGCTTAAGAA TTATCTTGAT GGTAATCTA GTTGTGCCAT 120
 GCATTTTCTC ATGATGAATC TAAGGTGTTA ATTCATTGAT GTGTGCTGTT TCTTGnnGCT 180
 TCATTATCTT GATGGTGAAT TTCGTTGTCT AATGCACTTT TTCAAATGA 229

(2) INFORMATION FOR SEQ ID NO: 1142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1142:

TTATGAATCG TAAACACTAT TATGAAATTG CGAGAGTAGT ATGGACAATT AGAATGATGG 60
 TTTGAGATna ATGAGTGTTA TTAACCAAGT CGTCATTTTC AAAATTAGAT GCCATTTGCT 120
 GTTTATAATA AGAAGATTTG ATAATTAAGT ATGATAAGAA AAAATAATAA TCCCCTAAT 180

GGTAAATGTG TGAATAGGAA TATCGTTaAA CTAGTTGTGT TCATGCTAAT TTTAGTTGTA 300
GCAGTAGCCG G 311

(2) INFORMATION FOR SEQ ID NO: 1143:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 403 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1143:

GATCATGCTG CTTATGAATT ATTAAAAGAG CAATTAGAAG ATGTGCTTGA TACATTAACT 60
GATAGAGrAG AAAATGTATT ACGATTAAAGA TTTGGTCTTG ATGACGGCAG AACAGAACA 120
CTTGAAGAAG TTGGTAAAGT TTTCGGTGTT ACACGTGAAC GTATTCGACA AATTGAAGCA 180
AAAGCACTTA GAAAATTAAG ACATCCAAGT CGTAGTAAAC GTTTGAAAGA CTTTATGGAT 240
TAATTTATAT CAATTTTAAT GATTGATACC AAGACATGAA TAAGCGCTTA TTTATTTTTA 300
GATTAAGATT TAACTTGAGT GAGGTTGGTA CATTGAATAA TGTTCACCT CTTTTTTCAT 360
AAAGGAGAAT AAAATGATT CGTTAAATAA CCGATTAACG ACn 403

(2) INFORMATION FOR SEQ ID NO: 1144:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 612 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1144:

nTTTTTATTG ACGAATnGAT TTGGATGCCA GTGGTTTGGT CCTAAATCTT GAATGGTAAT 60
CGGTATTTTG AAGTTGTCAT GTGCTACACG GTAACCGATA CChATAAGTA CGAGTGCGAT 120
GACAATGGTT GTTACGAGCA AGATGTATTG TAACCATTGC TTGAACACAA CAAGTTGTAT 180
ATAAGGCTTC ATTGACGATA CCTCCAAACC AATACAGCTA AATTAAATTAT CAAAAGTGCG 240
ATGAAsCTAA GATAGAACT AGGGTGCACT TCTAAAATGT AGTTGTTTAA AATAATTTCT 300
AACAAATTGAT TTGTTACAAC TGCGAACGGT TGAATATTGA AAACACCATT TGCTATATGT 360
TGTAaaaaaa TCGTAGGTAT TGTTAAACCA GATAACACCA GGATGACAAT AGCTAATATG 420
ACTTTACTAA TACTATTCAA CAAGCCTGTT GTTAAAAGTT CGATGAGTAA TAACCACAGT 480

AAGATATTCG GAATACTGAA CACAATCCAA ACTACACCAA CGATACTCCA TAACATAGTA 600
TAAAACCATG Tn 612

5 (2) INFORMATION FOR SEQ ID NO: 1145:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 223 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1145:

GAATTAATTA CGCAATATCC GCATCTTAAA TCAGTAATCG AAAAATATGC TTTAAAATAA 60
 AAGTGATCAA TGAAGTGGTT TGAAGGTTGT NAATAAACCT TTGAGTCACT TCCATTTTTTA 120
 20 AAnGTATCCT GGATGGAACC AGAATAGATT TGAAGCTTCA GCGGTTTTTA ATGAAAATAG 180
 CCATTAAATG ATTTGGAAAC GGTAAGAGGG GGTATTTAAA ATG 223

(2) INFORMATION FOR SEQ ID NO: 1146:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 598 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1146:

35 CGCTACGGTT GTCACGTAGA GATTAAAGAC ATTCGTTTTA GCAGTCTTTT TTATTGAAGT 60
 AGACAATGCA AGATTAAATG GTAGTGTCA TTTATTTTAT TCAACTGCTA GTAATACAAT 120
 ACTCAGTGTT ATTTCTGAAG ACTGGTTGAT AATGAGACAT GCATATTTAC AAAACTGTGT 180
 40 ATATTGTGTA TATTGTATAT ATACAGAAGT TGATAGGGGG ACGTTGATGA AAATAATTTT 240
 AAAAAACAAT AGTGATTTTC CGATTTATGA ACAGATTAAG CAACAAGTAA AACAAAATAT 300
 TTTAAAGGGA CATGTTGCTC CTGGAGAGCA TTTGCCGTCA ATGAGAGAAC TTGCCAAAGA 360
 45 TCTTCAAGTA AGTTTGATTA CTACCAAACG TGCTTATGAA GATTTAGAGA AAGACGGTTT 420
 TGTTACAACA ATTAGAGGAA AAGGGACCTT TGTTAAGGAG CAAGATAGTT CTATTTTAAA 480
 AGAGAAACAA TTTTTTACCA TGAAAATTG GGTAAAGGA ATTGGGTTAA TTAAGCnCAA 540
 50 GCCCATCCGA ATTGCCACTT GAGGGAnTTC CAGGATATTT nACCGTCATT TATTGAGG 598

(2) INFORMATION FOR SEQ ID NO: 1147:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1147:

10 CAAATCCAAG TGCAACAGCT TTATCAACCG CTAAAGCTAA ATCCGTATCA GCTTTTTCAG 60
 CTTGAACTGG TTTGATTGT AACTGTTCTG TTAGAAGTTG GCGTTCTTCT TTACTGACTG 120
 AATCAAAGTC TCCCACTGAG AAAAAAGGGA TAATTTGATG CTTCAATAAA ATCAAAGCAC 180
 15 CTCTATCAAC GCCGCCCAT TTACCTTCAT TACTTTTGGC CCAAATATCT TGCGGCAAGT 240
 GTCGATCAGA ACATAATAAA TTTATATGCA TATACACTCA ACCTTTCAAT GCTTGTGTGA 300
 CTTTTTAAAA TCCCCTGGTT AAAGGAAAT GAACnGGTTA CCAGCATTGT AGGCACCATT 360
 20 TCAACACCAA CTTCCGGCGGA TcGGATTAC GCCTCCATCC 400

(2) INFORMATION FOR SEQ ID NO: 1148:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 352 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1148:

TTTTATGGGT CGTATTTTTT AAAGAACAGT CATCTCAATC AAAAGAATAG AATTATGATT 60
 35 AAAATTGATG TrGTGTAATT AAAGTAAGAA ATrrKAAATT TCCTAGATAC TAGATAAATC 120
 GTTCAATATT TTATTAATAA GGAGGCCAAT GAYTGTGAAT TATCTTAAAC ATAAGTTTAA 180
 TAACTTGTTA ACTACAATGA TTGTTCTCTT TATTTTTGTA CTTTCTGGTG CGATTTTTTT 240
 40 AACATTTCTA GGGTTTGGTT TATATGGATT AAGTAGAATA CTTATTTATT TTAGGTTGGG 300
 TGACTTTACA TATAATAGAA GTATGTACGA TAACCTATna TATTATGGCA GT 352

(2) INFORMATION FOR SEQ ID NO: 1149:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 916 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1149:

GGCAAGTGTC GATCAGAACA TAATAAATTT ATATGCATAT AACTCAACC TTTCAATGCT 120
 TGTGTGACTT TTTTATAATC CTCTTGTTTA AAGAAAAATG AACCTGTTAC TAGCATTGTA 180
 5 GCACCATTTT CAACACAAAC TTTCGCTGTA TCGGTATTTA CGCCTCCATC AACTTCAATA 240
 TCAAAGTTTA ATTGACGTTT CATTTTAATA GCATTAAGAC CCGCTATTTT TTCTACGCAT 300
 TGATCAATAA ATGATTGACC ACCAAACCCT GGGkTAACTG TCATCACTAG TACATAATCA 360
 10 ACAATGTCTA AAATAGGkTC AATTTGTGAT ATkGGkGTAC CAGGATTAAT TACTACACCA 420
 GCTTTTTTAT CTAAATGTTT AATCATTGTA ATAGCACGAT GAATATGAGG CGTTGATTGG 480
 15 ACATGAATTG AAATCATATC GGCACCATGT TCTGCAAATG ATGCAATATA CTTTTCTGGA 540
 TTTTCAATCA TCAAATGTAC GTCTATAGGT AATGTTGTGC CTTTTCTTAC TGCATCTAAT 600
 ATTGGTAAAC CAATAGATAT ATTAGGCACA AATTGACCAT CCATAACATC AAAATGAACT 660
 20 CCGTCGACGC CTGCTTCTTC AAGTCGTTTT AATTCATGTT GTAAATCCAA AAAATCAACA 720
 GATAATAATG ATGGATATAG TTTTGTCAAT TAATATCTAA CCTTTCTATT TGAAATTTCA 780
 TTAAATAGTT GTAAATAATG GTCGTATCTA AATTGCGCAA TATCCCTAT CTCTAATTGA 840
 25 TGCnTAACCA TTACCATTAG GGTCTTGGn AAGGnTACCA ATCCCAAACC TTACCATGGT 900
 TCCCCATAA CGGTTT 916

(2) INFORMATION FOR SEQ ID NO: 1150:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1150:

40 AATTGGTTGT GGCTTCATAA TGTGCAGTTG CTAGACTGAC AACAACATCG TTTTCTAAGC 60
 CTACGATAAA GACATAATCG TTTTCGGTTT TGTGTTTTTC TAATGAGCGT TTGATGATAC 120
 GTTTATCTTC AGTTAAAGAT AAATTTAACT TAGTATCGTA AAGTTTAAGT GCTTCGTTGT 180
 45 AATGTGGATC TTTGACAGAT TGAATGGTTT TAAATTCCAT AAGAACACCT CCCCATTATA 240
 AATAATATTA TAGCATAATC GCCTGCTGTA AAAGACTGTT CATAACTTT TAAATGGTAT 300
 AAAAAACTGT ACTATCTTAA ATTAGACAGT ACAGTAATCT CATTTTGAAT TCAGTGTGAT 360
 50 AACTAAGCtT TGGGACCTTT AGATGChTTC ACAAATGTG 400

(2) INFORMATION FOR SEQ ID NO: 1151:

55

(A) LENGTH: 156 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1151:

10 CACCGTAGTT GCTACCGAAT ACCACCATGT CGCCAGGTTG TGCTAAGAAG TCCGGTGTAT 60
 TTTGGTATAC AGTTAGCTAA TCCGTCGAAG TTGTTAGCGA CGGnATATCT TTTGACCTAA 120
 ACCTTTTAGA GTAATCCAAA CAAACTTTC CAACCA 156

15 (2) INFORMATION FOR SEQ ID NO: 1152:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 233 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1152:

25 TCAATATATT TATTTTCTAA TGTTAGCATT AGACTGTCAA TTAAGTTGCT AAGTTGTTCT 60
 TTTTGTAGTAA TTAAATAACG CATATAATCA TCTAAATTTT TCATTTTTC ATCTAGCATT 120
 30 AAAAGGTAGT GCTTGATATT ATTTCTACTG TTGCTCCAAA AACTGnTCAA ATATACGTCA 180
 CATAGTCATA CTACACCTTT ATATAGTTTA AATACCAATn GCATGACCTC GTG 233

(2) INFORMATION FOR SEQ ID NO: 1153:

35

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1153:

45 TATATTAnCT ATCGTTTAGT GAATCGTTTT TTAAAGTAA AGAGAATTC ATATATTACT 60
 GGTACAATA CAAGTGTTAA TAAAGTTGAT GAAATTmAAC CACCAATAAC TGTAGCAGCT 120
 AATCCTTTTCG AAATAAGAAT CGnGCTATCT TGACCAAATA ACAnAGGAAC TAATGCACCA 180
 50 ATTGTTGCAA TCGCCGTCAT TAATATCGGT CTAATTCTAG TACCGCCTGC TTCGATTAAT 240
 GCTTCTTTCA TCTCCATGCC CTGTTGCTCA TTATTAATAA CACGGTCTAT TAACACAATG 300
 GCATTTGTTA CTACGATnCC AATTCAACAT TAGCATACC 339

55

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 817 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1154:

10 GTCACCCGAG GAGATATnCA ACTATTACCA TCAAGATGnG AAAAATGGTg CCTTGGGCAA 60
 TTGGTTTTTTT CATTTTAATC TTAATTATTA TATTGTTCTT TTTACTAAGA AACTTCAATT 120
 15 CACCTGAAGC GCAAATAAA ATATTAGTGA ATGCGATTGA GAATAATGAT AAACAAAAAG 180
 TAGCAACATT ATTAAGTACT AAAGATAATA AAGTAGATTC TGAAGAAGCA AAAGTATACA 240
 TTAACATAT CAAAGATGAA GTTGGGCTTA AGCAATTTGT CAGCGACCTT AAAAATACGG 300
 20 TACATAAATT GAATAAGAGT AAGACCAGCG TAGCTTCTTA TATTCAAACC AGATCTGGTC 360
 AAAATATATT ACGTGTAAGT AAAAATGGCA CACGTTATAT CTTTTTCGAT AATATGAGCT 420
 TTACTGCACC TACCAAGCAA CCAATTGTTA AACCGAAAGA AAAAACAAAA TATGAGTTTA 480
 25 AATCTGGTGG TAAGAAAAAG ATGGTTATAG CTGAAGCAAA TAAAGTGACG CCAATAGGTA 540
 ATTTTATACC GGGGACATAT AGAATTCAG CTATGAAATC AACTGAGAAC GGTGATTTTG 600
 CAGGCCATTT AAAATTTGAT TTTAGACAAA GTAATTCTGA AACGGTAGAT GTTACTGAAG 660
 30 ATTTTGAAGA aGCAAATATA TCTGTAACCT TAAAAGGCGA TACAAAATTA AATGaTAGTT 720
 CTAAAAAAGT AACTATAAAT GaCCATGAAA TGGCATTTC AAGTTCCAAA ACGTATGGTC 780
 35 CATATCCACA AAATAAGAT ATTACCATTT CCAGCTT 817

(2) INFORMATION FOR SEQ ID NO: 1155:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 322 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1155:

40 AnAGAAAATC TACTCTTTTA TTTCACCGGT CCTTCTnTAT CTCTCATTTG GCATACTCCT 60
 50 TTTTCCATAT TTCCACTTT CanTACAGTG AATTTACATG TAATTTTACA AATACAATAG 120
 GAGTGACAA TGGAATATAA CACTTATCTT TATAAAATTA TTTTATATAT TGaCGCACTT 180
 AAAACAATCT ACAAATAAAA TTATTATTTT TAGTTTTCAA TGAATAATTC ATTCTTATCT 240

ATTTAAAAAT ACATTTAACT GC

322

(2) INFORMATION FOR SEQ ID NO: 1156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 777 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1156:

5	GTTC	CAAnG	GAAnGnGAAA	AGCGTCAACC	AATGTA	ACTA	TTTAAAGTCA	AAGTGTTTGA	60
15	CCAAATTTGA	CTTAATATGT	AAAATAATGA	GTAACAGTTA	TTACAAGGAG	GAAATATAGA		120	
	TGAATTTAAT	TCCTACAGTT	ATTGAAACAA	CAAACCGCGG	TGAACGTGCA	TATGATATAT		180	
20	ACTCACGTTT	ATTAAAAGAC	CGTATTATTA	TGTTAGGTTT	ACAAATTGAT	GACAACGTAG		240	
	CAAATTC	CGTATCACAG	TTATTATTCT	TACAAGCGCA	AGACTCAGAG	AAAGATATTT		300	
	ATTTATACAT	TAATTCACCA	GGTGGAAGTG	TAACAGCTGG	TTTTGCGATT	TATGATACAA		360	
25	TTCAACACAT	TAAACCTGAT	GTTCAAACAA	TTTGTATCGG	TATGGCTGCA	TCAATGGGAT		420	
	CATTCTTATT	AGCAGCTGGT	GCAAAAGGTA	AACGTTTCGC	GTTACCAAAT	GCAGAAGTAA		480	
	TGATTCACCA	ACCATTAGGT	GGTGCTCAAG	GACAAGCAAC	TGAAATCGAA	ATTGCTGCAA		540	
30	ATCACATTTT	AAAAACACGT	GAAAAATTAA	ACCGCATTTT	ATCAGAGCGT	ACTGGTCAAA		600	
	GTATTGAAAA	AATACAAAAA	gACACaGATC	GTGaTaaCTT	CTTAACTGCA	GAAGAAGCTA		660	
	AAGAATATGG	CTTAATTGAT	GAAGTGATGG	TACCTGAAAC	AAAATAATTC	AAAGTAAAGA		720	
35	GTAGACTAAG	CnGTCTGCnC	TTTTTGTATG	AGTAAACCAA	GGTGTCAATA	ATTTGTn		777	

(2) INFORMATION FOR SEQ ID NO: 1157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1157:

50	TAATGTTGGT	AAATGGGTTA	AGGCTCAAGG	TCGTATTGAA	GAAGATACAT	TTATTAGAGA	60
	TTTAGTTATG	ATGATGTCTG	ATATTGAAGA	GATTAAAAAA	GCGACAAAAA	AAGATnAGGC	120
	TGAAGAAAGC	GTGTAGAATT	CCACTTGCAT	ACTGC			155

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1158:

10 GACTGGTAAA ATTTACTTGC AAGTAAAGCT AAAAGGTCAA ATAGATAAAG AACAACTTGT 60
 TTTTCAAAAT GACAAAAATG AAGAATTTCC TTTTGTATA AAAGATGAAA AGGATGACAC 120
 15 AATAGTAAGA ATTTTAATTG AACAGCATAT GGATAAAATC AATATGCATG TTAAAACGTT 180
 GGCTGAAAAG AAAAATCTAG ATAACAAAGA AATGGTGTAT TCTATTCATT TTAAAGAGAA 240
 AAAAGTACAA CATGATGATG CAAAAGAAGT GCCTTCAAAA CATCAAAATC AAGAAAATAA 300
 20 TCAAGATCAG CTTAAAAAG ATATTGATGA CAAAAGAT AGTCAAAAAT CAGATACTAA 360
 GGGAAAGACG TACTAGCCTT TTTACTGGAA AAAGGGTTTA 400

(2) INFORMATION FOR SEQ ID NO: 1159:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 347 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1159:

35 GTTGGTACAG TGTTGAGTGG CTTTGAATAT CGAACACAAA AAGAAAAGTA TGACAATTTA 60
 TATAAATTTT TCAAAGATAA TGAAGAGAAA TACCAATATA CAGGCTTTAC AAAAGAAGCA 120
 ATAAACAAGA CGCAAATGT CGSATATAAA AATGAATATT TTTATATCAC TTAATCGTCA 180
 40 AGAAGCTTAA AAGAATATCG TAAGTATTAC GAACCATTSa TTCATAAAAA TGATAAAGAA 240
 TTTAAAGAGG ATGGACCAGC CAGAAAGAGT GATTACGCGC TATCAGTCAG TACACATGTT 300
 AGTCCAGGAA TTTCTAGCAT CnGTGTGTGA ATGAnGGGTA CTTTATn 347

45 (2) INFORMATION FOR SEQ ID NO: 1160:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 203 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

ATAGTTATTA ATTAGTAGAA TGAGGATATT TAAATATAAA GGATAGTGTT GAACTTATGG 60
 ACATGAATTT CGATTATACA TGAACGGTGT TGTAGAACAA GCAAGGAATG AAATTGAATC 120
 5 TCGGGGATAT GAGCAATTAA CTA CTGTCAGA AGATGTTGnA AAGTCCAAGA TGGnCACCTA 180
 GTTATGATCA TCTGTATGTG GTG 203

(2) INFORMATION FOR SEQ ID NO: 1161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 203 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1161:

20 AATTTTCcNT TTCCATTGAC ATAACAATAT TTAATGGGGC CCAGTTACAG GAAATATCAG 60
 TTTAATTTTCG TCTTCAATAC CATACTCATA ATATTGAGTT TCTTGTGTTG GAAAATTTCC 120
 25 TACATAGTTA CTCATTTTAA TTTGTnCAAT GTTTGAATTC GGAGGCACTC TGTTAGCATA 180
 TCTTTAAACA TTAAATATGG ATA 203

(2) INFORMATION FOR SEQ ID NO: 1162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1162:

40 GCCATTTTGT TGATAAAGGT AACTAAAGG TTATCGTTTT GAAATTTTAA GTAAC TAGGA 60
 TATGTTTCGT GTTATAGGAC CGGAATTGT GGTATACGGT AAAATTTTAA TGCTATTGGA 120
 ATTTTAAAA ATGGAAAAAC ATGGACATn AAATTGGAAT TTCATAATAT GTCCTAATTG 180
 45 GACTTAACTT GTTGGGAGTT CATTTTACTA TTTnTATGGT 220

(2) INFORMATION FOR SEQ ID NO: 1163:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 409 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55 (D) TOPOLOGY: linear

TTCACGTGCT TTGCCATCTT TTAAGAAAAA CACAGAAGGC TTTTGCAATT GATTGTAATT 60
 GGATGCCATA GAGTAATGAT ATGCACCACT TGATAATATA GCAAGATAGT CTCCACGTTT 120
 5 GACTGGATGA AGGTAATTGA GCGTCTTTAA TAATGATATC ACCAGACTCA CATAATTTTC 180
 CAGCTATAGT TACACTGtCA TCTGCTTCTT CATTCTATT AACAGCAAT GCTTGATACT 240
 TTGCGTCATA AAGTGCAGTT CTGATATGAT CACTCATACC GCCATCGATT GAAACATATT 300
 10 TATTATCTC TGGAATTTCT TTAATGGtTC CAACTTCATA TAAAGTAACG CCAGctTCAC 360
 CTACAATTGA TCGTCCCGGT TCAATACCTA TTTCTGGTGC ATCTATACC 409

(2) INFORMATION FOR SEQ ID NO: 1164:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1164:

CTAAATAATT GTTTAAATT TTCTTTACTA TTTAGACTCT TTCGTTTGCC ATGACTTTGT 60
 TGACTTGCAA TGATATGAAC TTAATTTTTG CGAAAAAGG AATCTAGATT CATATATTCG 120
 GTAAATCTAC CTTCTTCATC TTTTGAAC TGTAACTCTA GCAATTCACG TTGTATCAAA 180
 30 TTTTGAATAA CCATTGTAAT ATCACGTGGT TGCATAGTTG AGCCCTTCTG AAGCAATTCA 240
 ATTGAAGGCG TTTATTTGAG TTTCGAAGCA TAAATCAATT TAAGCAAAAG ACTAAATCTG 300
 35 CTCATCTAAA CCTAAGTCAC GTAAGGTCTA ATAATcCTC GTATCACTAC GGTCTGCTTT 360
 AATGATATTA nCCAGTnCGT GCCCCTTTAA ATTAACCAa 400

(2) INFORMATION FOR SEQ ID NO: 1165:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 327 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1165:

TCGTGATTAT TATTAACATT TTTTAGTAAT TGATAAAAAT CTTTCTCGC TTCTGTAGGG 60
 CTAGTGATAA TCATATAAAC CCTCCTTTGT TTTCTGTTAC GCTAATTATA GCGTACGCAC 120
 TTGAGTACGT CATATTACCA ACGAGTTCAT CGTATnCTTA ATACCTATAT GCTATCGTTC 180

GGACACACTA CCGTnCCCTA CCGCTTAACC ATTTTCGGGC TTAAATATTT TnAAAATTTT 300

GTGAGGATCC GCTTAAAAAA CCCATA 327

5 (2) INFORMATION FOR SEQ ID NO: 1166:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 148 base pairs
 (B) TYPE: nucleic acid
 10 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1166:

TTAGGATTTT AAAAATATGT CAGGTTTTCT AACATAAAGC TTAAACATT GATATAACTG 60

CGATTTTCGAC TTCTAAAACG aCATrAAACA ATCTACAGCT AAAATATTTT TTCAATTTTT 120

20 AAAAATAATC nGATAAATCT CCGCAAAA 148

(2) INFORMATION FOR SEQ ID NO: 1167:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 243 base pairs
 25 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1167:

nAAGGTTATC GCTCGTAAAG ATGTTGTTGA AGAAGTATTT GACGACAGAG ATGAAGTGGA 60

35 TTAAAGTACA GCGCTTGTTA AAACCCTGCA TATGAAATTG GTGTATTACG AGAAGAGTAC 120

ACTAAAGTTT TGGTCGTGTA GGTCTCAGCA GCGAACAGCA GTATGCACGT CTnCGTGTGC 180

TGACGTGAAT TTTATTTGAG GATTATAGAC AAGAGAGACA ACTTACTGGA TTTTGCCCTG 240

40 TTG 243

(2) INFORMATION FOR SEQ ID NO: 1168:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 321 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1168:

TTATATACTA CTATAAAGTA ATCCAACCTCT CAAGAAGGTA AATGGGCAAT CAAAAGCTAA 60

55

CTACGATATG ACGCACAAAT GTTTAGGGTA GGTGTTGACA TGTGTACAAA AACCAGCAGA 180
 AAATGGTTGT ATTCTTAGGG AAGCATTAT CCTGTTTGT TATGGATTCA TGTTACGTTA 240
 5 CGCAATATGT TTTTATGTAG CACAGCGAAG TTACAAGTTT ATACGCTATA TATCCnTTGA 300
 TnATAAGCAG ATGAAAnCGG C 321

(2) INFORMATION FOR SEQ ID NO: 1169:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 191 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1169:

20 TCTACATATA AAATTACTCT AAAAATATGT ATAGTCATAA ATTGTTGGTT GATTTAATTA 60
 AAAGTATGGA AATTAAGGGG CTCTTATGTA TATAAAAAA TGAATTATGA TAAAATGTAA 120
 GAnAATATTT AGGGTCGATT GGAGAGATAC CAGTGGTACC AATTAGAAGA CGACAGTTTA 180
 25 ATGTTACCAT A 191

(2) INFORMATION FOR SEQ ID NO: 1170:

- (i) SEQUENCE CHARACTERISTICS:
 30 (A) LENGTH: 303 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1170:

AATGTGGCAA GGATGTGTGA TATGACTACA CATTTAATAT CATCCTACAT AATAATTTTG 60
 40 CGCGCCAAAT TATCACACAT ACCACACCTA CGATAACAAC CAGCAAATAA GGAGATAGCT 120
 ACACGTATTT ATCGTCCAAC GCACTACTAT ATGTTGTTGn ATGGGCACAA GTTACAAGTG 180
 CTGCGCGATA nCAAAAGGTA TACCCCTAAA ACGTCCCAAA AnAATCGGTT GCCCTCTAAA 240
 45 TTGTTTGTGT GCCATGATTA GTAAGCTCCT ATGTTAGAAA CAAATTTTnTA ATATGTATAT 300
 TAG 303

(2) INFORMATION FOR SEQ ID NO: 1171:

- (i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 299 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1171:

5 AAAAAGGTGA GAAGACAATA ACGnACGCCA AACTAAAAA ATCCATTAC TGGAGAAATT 60
 ATTAGTAAAG GTGAATCGAA AGAAGAWATC ACAAAGATC CAGTTAATGA ATTAACAGAA 120
 10 TTCGGTGGCG AGAAAATACC GCAAGtCATA AAGATATCTT TGATCCAAAC TTACCAACAG 180
 ATCAAACGGA AAAAGTACCA GGTAAACCAG GAATCAAGAA TCCAGACACA GGAAAAGTGA 240
 TCGAaGAGCC AGTGGATGAT GTGATTAAAC ACGGACCAA AACGGGTACA CCAGAAACA 299

15 (2) INFORMATION FOR SEQ ID NO: 1172:

(i) SEQUENCE CHARACTERISTICS:

 (A) LENGTH: 821 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1172:

 CCATAGCACC ATACCCATAA CCAACCCGGA TCCACTAACA AATAAGTAAT AATTnACACG 60
 TAAAGAAGTA AATAGTTTGT TTAGTTTTTT ATCAACATTA TCAATGTTTT CCTCGACAGA 120
 30 CACTGGAATT TTAACGATAG CTGTACCATT TGTAATTGAG TAGTTCGTTA TTCACCCAT 180
 ACTACTATTT GGTAAAATTG TTAATTCTCC TGAAATTGTA TTGATTCGTG TTGAACGAAG 240
 TCCAATTGAT TTAAGTGTTC CCTCTGCTAC AGTTGTACCA CCGTTATTTA TCTTAACATA 300
 35 ATCACCTACA TCAAATTGGC TTTCAAATAT AATAAAGAAT CCTGTAAtTA CGTCTTTAAC 360
 TATAGTTTGA GCACCAAAAC CTAAGCTTAA GCCTACGACA CCAGCACTGG CAATTACCCC 420
 TTCAACACTA ATGCCAAATT TACTTAAAT CGTCGTAATA ACTATAAACC AAACGATATA 480
 40 CTTCACTACA TTTTGAACAA GAGATATTAA AGTTTTAGAG CGTTTTTTGT TACTCTTTTT 540
 ACTTTTATTT TGAATCTTAA ATCCCTGTTT AATCATTTTA TTCAGTATTT TAATAACTAT 600
 GAGGGCTACG ATAATATAAA TAACAATCAT AGCGATTTTA GTTGCAATGT TTCATATGT 660
 45 TTCTATTTTT GTTAATGGCT CAAATAGAGA TGAAATAATA TTCATGACTT GATTCATGAT 720
 AAGCGTIACT CCCTTCTATA TGATATATCA ATATTGtATT CATATTTtga TATACTTCCC 780
 50 AACCTACGAA AGTCTCAATA TATCTAGAAT CAAATGGTCA C 821

(2) INFORMATION FOR SEQ ID NO: 1173:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 318 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1173:

CCCCGAAATTA AAATAATAAT GTATGAATTT TTAAATATGA TTTAAACGTT TTCAGTTTTT	60
ATGAAAACGC ATGCATTTTA CAAATAAAAA TGGTACGATG GCACTGGTTA AAACGTTTTA	120
CTAAAAACAA ATCATGAGGT GTATAACATG GCTGTCTCTT GATCAGATCT TGATCCCCTG	180
CGCCATCAGA TCCTTGGCGG CAAGAAAGCC ATCCAGTTTA CTTTGCAGGG CTTCCCAACC	240
TTCCCAGAGG nCGCCCCAGC TGGnCAATTC CGGTCGCTT GCTGTCCATA AAACCGCCCA	300
GTCTAGCTAT CGCCATGT	318

(2) INFORMATION FOR SEQ ID NO: 1174:

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 206 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1174:

GTTTTAAAGC GGTATAAAG GTAAATATGC AATATTATAA ATATCTTCTT CAGAACTAAC	60
CGAATCTCTT AAACATCATCT GTTTAGATAG CGTTTCAAAT TGATAAGTTT TAATTTTAAC	120
CGTTACAGTT TTAGCTGACT GCTGTAATTT ATTTGACGTG CAGCTGTTTT CCTGACAATn	180
CCATACTGTC TTAAAnCTCT GCATCA	206

(2) INFORMATION FOR SEQ ID NO: 1175:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 653 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

40

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1175:

TCTGCATCAT GATATTTTAC ATCATCACTT ATAAGAACGT CACCTACATT TAACTTTCA	60
TCTAAAGCTC CAGCAGAACC TGTATTAATA ATGACGTCCG GTTTAACTT ATTAATTAAT	120
AATGTCGTAG AAATTGCAGC ATTAACTTTT CCAATGCCAC TTTGGGTAAT CACTACTTCT	180
CTATCTTTTA AAATGCCAGT ATAAATTTA ACATGTGCAA CTGaAATTC GCTTAATTGT	240

55

TTTTGATTCA CCTCTTTTAA AAAATCCTAG CATTTGTTAT TTTATCACAT TTTAATCACT 360
 ACAACGACAA GTATCATGAT TTTCATTACT GTATAAACT TTGCTTTAT TTTATAAAA 420
 5 AGCTCTTTCA ATAAGTTGGT ATCATAAATA ATTATTTTTG AACTCATAAC AGTGAATTCG 480
 ATGCTTCTAA TTGTCTATT TTTCAAAAAC TAATCATATC AATTACATCA ACCAATCGTG 540
 10 TCAAGAACT TTATTAAATA ACAGGAAAAT GATATGTTTA AAATAATAAT GATGAAATAA 600
 GAGGATTTTA CAATGGCTGA AAATAATCAA AATAGTCTCG TAACAAAGAT AGC 653

(2) INFORMATION FOR SEQ ID NO: 1176:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 410 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1176:

ACATGGCCGT TnGnATCAAG ATCGTCAACG ATTGATTGGT ACAATTGAAG TGCCTATGAC 60
 25 ATTGGCAACC GTTGGCGGTG GTACAAAAGT ATTACCAATT TGCTAAAGCT TCATTAGAGC 120
 TACTAAATGT AGAGTCAGCA CAAGAATTAG GTCATGTAGT TGCTGCCGTT GGTTTAGCGC 180
 30 AAAACTTTGC AGCATGTCCG GCGCTTGTGT CAGAAGGTAT TCAACAAGGT CATATGAGTT 240
 TACAATATAA ATCATTAGCT ATCGTTGTAG GGGCAAAAGG TGATGAAATT GCTAAAGTAG 300
 tGAAGCTTgA AAAAAGAACC CCGTGCAAAT ACACAAGCAG CGGAACATAT TTTACAAGAA 360
 35 ATTAGACAcA ATAGTAAAgG CGAtAGCGAT TaaAGAAAGT TAAGTCAGTC 410

(2) INFORMATION FOR SEQ ID NO: 1177:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 146 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1177:

CATCAAGATA ACTACGGTTA TAATCCAAAC GACCCAACAT CATATAACTA TACTTACACT 60
 50 ATTGATGGCA CAAGGTAAGT ACCATTACAC ATGGGAAAGG TAACTnGCAT CCAAGTCAAT 120
 TAAACCAAGA TAATGGCTAC TACAGC 146

(2) INFORMATION FOR SEQ ID NO: 1178:

(A) LENGTH: 225 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1178:

10 TGTTCCTGGT GGAACTTTGG TTTCTTAATG TCATTCCAAA TTGCTGTATA TTCATTTCATT 60
 GGTATTGAAC TTATAGGTGT ACTGCTGGTG AACGAAAGAT CTGAAAAAAC CTTACCGAAG 120
 CAATTAATAA TGTACTATCC GTATTTATAT CTATATCGTG GThATATAAT ATGTCATCAA 180
 15 CTGGAThTAC GTCAATAGAC CTTGTAACC TTAATATCGC GACAT 225

(2) INFORMATION FOR SEQ ID NO: 1179:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 319 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1179:

TGTACGCATA TGGTATCACT TGATGACCAA TTCACTTTGT CGAGAATAAC GTTATTTAAA 60
 30 GACATTGATT TAAAGAATTG GTATTnCGTT AATTGTATTA GACTTTTCAC AGAAGTAGCG 120
 ACAACAATGA TTTTCATCGTA ACACGTGCGC ATAThAAGCA TGGGAAAAGA CATGGnCTAT 180
 TGACCTGGCC TTTTGTGCAC TCGTACGATT GGTCAAACGT GGCTATATTG AAAGTTTGAA 240
 35 ACAAGATGCG ACGGACACAC ATGCACGGGT GGGGATAGAT CTTGAGTCTT ATTGAGACGC 300
 TGTGACTGTG GCTGACATC 319

(2) INFORMATION FOR SEQ ID NO: 1180:

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 654 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1180:

50 TTTTATCTTG GTTTTGAAC AGATATGTGT AGTCAGGTAT AACATAATCA ATCCCTAATA 60
 AGTTATCATC AGGGAATTG ATACCTGCAT CGATAATGAC AATTTCGTCT TTATACTCAA 120
 CTGCATAAGT ATTTTACCG ATTTACCTA GACCTCCAAG TGCATATACA CCTACTTCAT 180

55

5 TTCGTAATCT AAATGTGCGC CCTCTAATTT AGTGATAAAT TCGATATTAA AATTACGATC 300
 TTTCAAGTAA CGACGTACTT GTTCTTCTGT TTGAGCTTCA ACATAAAGTG ATTGTGTATT 360
 10 TTCACGCACA ATTACCTCGT CTCTGTTATG TTGATAaAAA ACTTTAAATA CTGCCATGTT 420
 AAAAATTCCT CCTAAGAATG TTTGTyTAAT TTATTGTTAA CCCTAGtAAA ATCGTATTGG 480
 AGTATATATC GATAAATTCA TTCCAATCAT CTCTATAATG TAACTTAATA ACGATTTGGT 540
 15 TAATAACTAG GTTCATCATG TCGTTCATTT TAAAAAATTA GTGAAATAAC ACTAAAATTT 600
 CAGTTAAATC TTATTTTACA TGATGAATGA TAATAAATAA AGCAGTTTAT CTCA 654

(2) INFORMATION FOR SEQ ID NO: 1181:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 231 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1181:

25 CGGGTATATT ATAGACAACG TTTAAGGACT TCTTTTAAAC TAACTTGTTA AATGTTGCTG 60
 TATTTTTGTA TTCAACACTG TAATATAGTT CTTCTGACGC TGATGCATTA CTCATTGTAA 120
 TAGATATTAA AAGACTCAAT ATTATTAAaAC TAATAATAGC GCGCTTTATG ATTTTCATAT 180
 30 TCTAATCTCA AATGAATTCC AGAACTTTAT ATGGTTCAGG TGCGACAACA T 231

(2) INFORMATION FOR SEQ ID NO: 1182:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 245 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1182:

40 TAAGTTTATT AAGCTCTTGT CCGTTAGCAA TCACGCTACC GCTAGAAATT TGGTCAATTG 60
 45 AACTTAGTAC ATTTAATAAG GTTGTCTTAC CTGATCCAGA AGGCCCCATA ATCGCACGAA 120
 TTCGCCTTTT TGTATGTCAA GTAATATCTT AGGCTGAATG GTGCTTTACG AGTTTGAACA 180
 GTGCCGAAAA CGCAAGCCCC TTGATATCTT AAATGTGGAG ACTACAACnC ATCAGTGnCA 240
 50 CGGGA 245

(2) INFORMATION FOR SEQ ID NO: 1183:

55

- (A) LENGTH: 490 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1183:

10 AGGTAAAGAT GATGGTGTAG AAGTTTATGT GCATTGCGAA GATCATGGCA TTGTATTTAA 60
 TGCAAGTCTA CCTTTGTACA AAGATGCCAT CCATCAAAAA GGATCAATGC GTAGTAATGA 120
 CAACGGTGAT GATATGAGTA TGATGGTGGG TACAGTGCTG AGTGGCTTTG AATATCGAGC 180
 15 GCAAAAAGAA AAGTATGATA ATTTATATAA ATTTTAAAA GAAATGAAA AGCAATATCA 240
 ATATACAGGT TTTACTAAAG AAGCAATTAA CAAGACGCAA AACGTCGGGT ATCAAAATGA 300
 ATATTTTTAT ATTACATATT TATCAAGAAA TTTAAAAGAA TATAGAAAAT ATTACGAACC 360
 20 ATTGATACAT AARCAATGAT AAAGAGTTTA AAGAAGGTAT GCAACGAGCT AGAANAGAGC 420
 TAAACTATnC TGCTAATACA AATACTGTAG CAACGTTGTT TAGTACGAAT GATGAAAGGA 480
 ATAGAAAAGA 490

25

(2) INFORMATION FOR SEQ ID NO: 1184:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 153 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1184:

35 CTCGCATTTT AATAAATCAT ACTCTTCTTT TCTTGCATAC GAATTAGAAT AACTCGCGAG 60
 ACCTATAAGT CTCTTTCCTC ACTAGATAGT TTATACTTTT GGTCnGTTGA AGTCAATAAT 120
 40 TTTATCTAAA GCTATAAAAA ATCTTTTGAT AGC 153

40

(2) INFORMATION FOR SEQ ID NO: 1185:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 164 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1185:

50 TCATTTATAT ACTGTATCAC GCTATCTTTA GTATGTGTGC ATnATCATTC GTTAGTGCCT 60

55

AATGAACGTT ATATTTTAAAT TCATGATTTA CTAAGACCAA CAAT

164

(2) INFORMATION FOR SEQ ID NO: 1186:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 143 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1186:

CACAAGTGAC GTCATATTAT AAAAGTCACT CGGTTTGCGA TACGTTCTGT CTAAGAAATA 60

GCGACGTGCA ATTTTCATATT TTnTATAAAC ATCCGTTGAA AAAGGACATA AAACCATGCG 120

TTGAACCAGG GTCTATACTT TCT 143

(2) INFORMATION FOR SEQ ID NO: 1187:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 209 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1187:

CGTACAAGTT AAAACGTTTA TACGATTACG AAAAACCACC GAAGAGTGAG AAAGATCCAA 60

TGGTTGTATA TGCAGTAAGT GCAGGATACA AATGGTTCTT TGCTTATCCA GATGAACATA 120

TAGAGACTGT TAATACATTA nCAATCCCTA AAGATGTCTG TTTATTAGCT CAGCATGAnC 180

ATACAGTTCG GTCACATAGT GTCAATTCA 209

(2) INFORMATION FOR SEQ ID NO: 1188:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 243 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1188:

TTGGAATCGC TGTCGACGGA TAGCAACCTG GATTGTAAAA ATAATGTTGG CTTTAAAAAT 60

ACCGTGGCAA TATTAAATGA ATAATCGCTG CTGAATATAC ACCCAATGCA CATTGAACTA 120

ACATAAnCCC ATTTGACTCA TCTACGCCCA CTATTGGCTT TATAGTCACT GAACTAAGCT 180

TAT

243

(2) INFORMATION FOR SEQ ID NO: 1189:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 527 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1189:

15 CACCAGGCAT TACCATACCA GCACTCATAC CACCAATCAC kCTCGATAAC ATCAATACCG 60
 AAAAATTGTG GCCAACTGCA AACATAAATT CTGACACTGA AAACAAAATT AATCCTATAC 120
 ATATAATTAA TTTCTTCCCT AATTTGTCTAG CTAGCGTACC ACCAAACGGC GATATAATCA 180
 20 TTTGAGATAA CGCAAAAGCA GCAACTAGTA ATCCTAAATC ACTACCAGTT AATCCCAAAT 240
 CTTTTAAATA AACAGGCAAG ACTGGTATTA CTAACCGaT ACCTAAAAAA ATCAAGAAAA 300
 TATTAAATA TAAGACAAAA ATCTGTTTAT TCATATGCTC ACCTCTTTTT TCTCCATGTC 360
 25 ATGCTTAAAG CTGTTATTCT TCTTTCCAC GACAGATTGC AATTCATGCT AAATATTCGT 420
 AAAAAGTTTC TATATTGTAT TACAACATTG CTACATTTGA CAATATTTTt TCTTTGTAAT 480
 TATCACTATn TTCCATTTAA TTGTATAATT AATGACATAT TnnAATA 527

30 (2) INFORMATION FOR SEQ ID NO: 1190:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 889 base pairs
 35 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1190:

40 AAAACGTGGC AGTAATTCAG TGTATGTTCA ATATGATGAT ATTATGTTTT TTGAATCATC 60
 AACAAAATCT CACAGACTCA TTGCCCATTT AGATAACCGT CAAATTGAAT TTTATGGTAA 120
 45 TTTAAAAGAA CTGAGTCAAT TAGATGATCG TTTCTTTAGA TGTCATAATA GCTTTGTCGT 180
 CAATCGCCAT AATATTGAAT CTATAGATTC GAAAGAGCGA ATTGTCTATT TTAAAAATAA 240
 AGAACTGTC TATGCATCGG TGaGAAACGT TAAAAAATA TAATAAGATA ATAAAGTCAG 300
 50 TTAACGGCGT ATTCAATTGT AAATCTTGTT GGATTTTAAAC AAGATAACTA GCAAATGCAC 360
 TGTATAGCTG GCTTTTTAAT TTTATTAACA AAATTAAATA TGACGCGTGA ATTAAAAAAT 420

55

CTCTACAATC CTATGGCGCA ATTTTAAACA AAATATACTA AATtTGAAAT TGACAGTTTC 540
 CATAAATAAA ATGCTATAAT TTCCTACTCC GTGAATCTTA TTTTGTTTTA AAAACAGTG 600
 5 TAGCACAAGm TTCAACTTGT TTCCTACACT GTTTTATAAA TGAATTATTT AAAATGCTTG 660
 TTCAACCTCT TGTAATGTTG GCAAACCTATT AATAGCACCA TACTTCGTTG TCACTATTGC 720
 AGCAACGCGA TTGCTAAACG CTAAAATTTT CTCACCTTCA TTTTCAAATA ACTGAGTTAA 780
 10 ATTTAACACA TCGGTGCTA AAATCCTGCT AATAACTGCA CCAATAAATG CATCCCCTGC 840
 CCCAGTTGTA TCAACAGGTT TTACCTTATA GCCACTATGA TAATGATTA 889

15 (2) INFORMATION FOR SEQ ID NO: 1191:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 254 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1191:

25 AGTTAAAGGC GTTGAATTAA TGGCCTTCCC ACACATGACA TATAACAAG CGTGGCCTTA 60
 TGGTCTGATA ACCAGATACA CGTTTGAATG GATTAATTGC GTTCTCAATT AGGCGTGATA 120
 TGGCTTTAAA GATTThAGTC TGTGAAATG ATGGTGAATT AAAGCATTGC GCTAAAGGTG 180
 30 CACTGACAAT ATCTCGTAAA GATATGGTGC TTTAACAGAT TTGTAACATC TATGGTCTAA 240
 nGATACGGGG TTAA 254

35 (2) INFORMATION FOR SEQ ID NO: 1192:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 189 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1192:

45 GGCTCATAAG TTACACCTAA TGCATCTTCA TCTGGAAGCT GTGGTTTATC ATCTTCTAAA 60
 TCAGCAGTTG GCGTTTTTTC ATATAATTCC TTTGGCGCAC CAAGATACGC TAATAATTGA 120
 GCTGTGTATn CTAGGCATCC ATACCGTATT CGTTAAACCA GTATATTTC A GCTGAATGAT 180
 50 CTGTCCTAC 189

(2) INFORMATION FOR SEQ ID NO: 1193:

55

- (A) LENGTH: 215 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1193:

10 TTGATATTGA TCGATGACCG GTGATGTGAA ACTGAACTG ACCAATGAAT GAAAAAATGG 60
 GCAATGTTTT TGACAACAAT GGTCAATTTAC CAAATAAAAA TATTGnCACCT ATTAATACCT 120
 GnCAAACGTG TTTGTTCCCTA AACCTAAAAT GATAAATCCT ATTGAGATTA TGGACTTGGT 180
 15 AAGCTGCAAT CTTTTTAATA TCTTTATAAG CAATG 215

(2) INFORMATION FOR SEQ ID NO: 1194:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 158 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1194:

AGAATTTGCh ACTGGTTTAC CAACGCATCT ATGTTCCGGTC GCTATACCGC ATACAGATGT 60
 30 CGAACATATT AACCATAGAA CGTAGGTGTG GTGTTCTAGA AAAAGAAGTG CCGTTATTGA 120
 AATGGGACAC TTGATCAACA GCCGAGTGAA ATCGTTTT 158

(2) INFORMATION FOR SEQ ID NO: 1195:

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1195:

45 ACTGATTGCC GTGTCTGTGG TTGATCATTG ACTAATTGTT GGCCATCATT GgnTTCTTGG 60
 TCTAATAGTC CTCTTGCTCC TGCCTCGTAA CATTGATGCC AATCATGCAC CACAGCCTTT 120
 CGGTTTCATCA TAACATTTCT GCCAATGTCT CTACATGACA CGGGCTCATA TTTCTTTCTG 180

50

(2) INFORMATION FOR SEQ ID NO: 1196:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 274 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1196:

5 GTTAAAGCGA ATTATTCACA ATATAAAAAA TCATCTGATC AAGTCGTATT AGAATTAGGT 60
 ACAAATGGCG ACTTTACTGT CAAACAGCTC GACGATTTAC TTAATCAATT TGGaAAAGCC 120
 AAGATTTATT TAGTTAATAC ACGTGTTCa AGAaTTTmTG AGGgCAAATG TAAATCGATT 180
 10 ATTaGCTGAC GCGGgCGAAA CGGAAGTCCC AATGGTCACA TTAATTGGAT TGGGTATTAG 240
 GCGCTTCACA AGGGACCTTG GTGGAATATT TTTG 274

(2) INFORMATION FOR SEQ ID NO: 1197:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 154 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1197:

25 ACGTAACCTA TATTCATCTC TGTCCCAAT AATAATAACA ATACATAACT GATATTAGCT 60
 GTCTTTTCATC TCTAATTAAT ATGCGTCATG AGCTAATATC GGTTATTTTT TTGTGTCAAA 120
 nAATTTTTTAT TTTATCTCTC CATCGCATTA ATTG 154

30 (2) INFORMATION FOR SEQ ID NO: 1198:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 249 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1198:

AATcCTCACC CGTAATAAAC TCGTTTTGAA TGCgTTGTTC GAACTCACGG TATATTAAAG 60
 CAATATCCTC TAACTTATTT TTAGTTCGAG TTTGCATATT TTTATCAGCA ATAAAGTGTT 120
 45 CTAAATGTTT AGGCGTTACT GCATATTTTT TAAATCTTG AATTGTTCT GTTAATTTTT 180
 CACTAAACCC ATAATATTTT GCTTGTGATT GATAAAGTTT TAAATACTTT TGTTGTTcT 240
 GAACAATGn 249

50 (2) INFORMATION FOR SEQ ID NO: 1199:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 165 base pairs
 55

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1199:

TTGGTAAAAA TCGACATCAA ATACTAAATC TGATCCATCT GAATCCCATG TTTAAAACCG 60
AACTTGTGCA TTAATTGTAA AGTTCAAACCT CTCATCTTTC ATAGATCGCG ATGCGTCTTT 120
TAATTCTTTA GGTGATACTT TGTAGATCTA TAACAAAATT AnCTA 165

(2) INFORMATION FOR SEQ ID NO: 1200:

15

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 236 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1200:

TTATTAGCGA TTA AAAATAT GTCCTTTTTT AATAAAGGAG ATATATGTTT AATTGGATGT 60
CAACATCGGA GTCATTGGCG AAAAATAGGG CTGCGGTCGA TGGAGTTCAA AACTATTTTA 120
ATGCATTACA TGTGTGAGT nAAGAGCAGG CGTACTTTTC TTTTCAAGGA TGTGACATAT 180
CACAGAGCGT ACAATAGAAA CTCAGATACC ATTACATGGA nAGTCCGCGT CCAATG 236

(2) INFORMATION FOR SEQ ID NO: 1201:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 146 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1201:

TAATGCTAAA CAGTATTATG GAGATAATGG CTTAGTTCAA ATTTCTGATG nAAGTCAACA 60
CTTCTAAAAT TTGTAAATGA AGCATTAGAC AATAACGAAC AATCAGTTGA AGATTACAAA 120
AATGGTAAAG GCAAAGCTAT GGGCTT 146

(2) INFORMATION FOR SEQ ID NO: 1202:

50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1202:

ATTGATGATA GTATTGATGG TAATACCATA AACATACAGC ATTATCACAA GTAAAAGTAA 60
 5 GCCTAAATGT CCATAAAATC TACTTGTCAC AATATATGTC GGTATTATCG TAACGGGAAG 120
 TCATTTGTA CTTGATTAAA CTTTGTGTA ATTGCTTTAG TACCTTCTAA ATACCTGGTT 180
 GATGAGACCA CATACTGATA CCACCATAAC CATAAACAAA AGGTACACCA GAATTGGGCA 240
 10 TACTTCTTAT TCCATCCAAA CCTCCGTAAC CTAATTGCTA CGGGTAATAT TCCAGCACAC 300
 TAATAGTATT GATGTAATTA ACTGAACGGC CGCTnGATAA TAAGTCTTAC AGTCTTAAAC 360
 15 GTCCATGCGC ATCnTnACCC ATTCA 385

(2) INFORMATION FOR SEQ ID NO: 1203:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 580 base pairs
 20 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1203:

ACATTGAACG AGTACAGAAA ACTTATTGnA AATAAAGGCT GTTAAAGATA AAGAAGAAGA 60
 TACTTACAGA GGTAAGTATT TTGCGGAAGA AAGAAAAAAC GAAAAATTGG AAAAAGAAAA 120
 30 TATAAACTA AAAAACAAAA TTTATGAATT ACAAACGAA GAAGATAACG AGGAGGACGA 180
 AGAAGACAAG GAGGACGAGA ACGATGTATT ACAAATTGG TGAGATAAAA AACAAAATTA 240
 35 TAAGCTTTAA CGGGTTTGAA TTTAAAGTGT CTGTGATGAA GAGACATGAC GGTATCAGTA 300
 TACAAATCAA GGATATGAAT AATGTTCCAC TTAAATCGTT TCATGTCATA GATTTAAGCG 360
 AACTATATAT TGCGACGGAT GCAATGCGTG AyGTTATAAA CGAATGGATT GAAAATAACA 420
 40 CAGATGAACA GGACAACTA ATTAACCTAG TCATGAAATG GTAGGAGGTA TGAnAAGTGA 480
 ATGATTTACA AGAGAGAGAA TTAGAAACAT TCGAACAAGA CGACCGATTC AAAGTAACTG 540
 ATCTAGACAG TGCTAACTGG GTTTTTTnAG AACTnGGAT 580

(2) INFORMATION FOR SEQ ID NO: 1204:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 base pairs
 50 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

ATCGTATGGT TGAGGCGAAT TTAAGTGAAG AATTAGACAA CGAATCTTGA TGAAAAGaNa 60
 ATGATTGGCT AAAATAGCAG CTAATCAAAT CAACGATAAT GAATGCTTAT TTATCGTGCT 120
 5 GGTTCATCTA CATTGGAGC 139

(2) INFORMATION FOR SEQ ID NO: 1205:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 229 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1205:

AGGAGGATAT CAGaCTTACA rACTATTCTT TATCATACTG rAATTCAAAT AGACACCTTA 60
 20 ATTGaAGaAG GCGTTGACGC GCTACTTTTC rAAACGTATT ACGACCTnGG AAGnGTTAAC 120
 AAATGTCATT TCCACGGAAC GGGGAAAGGA AATACGGCGT TGCCAATCCA TTGCTCCAAT 180
 TGAACCGGTT GCAAACACAA ATTGACTTGG TTGAAGGGGC AGGCATGCA 229

(2) INFORMATION FOR SEQ ID NO: 1206:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1206:

TCGACAAGTT TTTAACAGTT CGTTATTATA TGAATGTAAG TAAAAATTTT TTAGCTACAA 60
 CTTACATATT ATAAATGCAT AAATTAAACA AAAAGGGGCG AAAAAAGTTG ACTCATTAT 120
 40 CAGATTTAGA TATTGCGAAT CAATCAACAC TACAACCAAT TAAGGATATT GCTGCATCAG 180
 TAGGTATTTT AGAGGATGCA TTAGAACCTT ATGGTCATTA CAAAGCTAAA ATCGACATTA 240
 ATTAAAATTa CGCCAAGAGa AAACAAAGGG AAgTTGTTTT aGTAACTGc GATGaGCCCCA 300
 45 ACACCAGCTG GTGaAGGtAA tCmACGGTTA CAGTTGGtTT AGCTGaTGcA TTCCmTGaGT 360
 TaATgAAAAC GTTATGGTTG ccTTaGGGGG CCTGCCTTT 399

(2) INFORMATION FOR SEQ ID NO: 1207:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1207:

5 ATAAACAATA CGATTGTAT GCTACTTGTC TGAAGACAAC ATAAACAAT ATGTTAACTG 60
 ATATTGAGGA TATGGGCTTT CAAAATCCAn TTCAAAAAAG GTTATGTTGA AATATGACGC 120
 GGTGTTGAA AGAGACCC 138

(2) INFORMATION FOR SEQ ID NO: 1208:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 655 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1208:

20 AGTCGCCTAA CTTTTATTAT CAGTCCAATC AGTTTCATCA AGTATAAATG TTATTGCTTA 60
 TTTTGTATT CTGGCGTAT TTTTGTCT TCTGCATAGC GCTCTGGATT TTTCTTATAA 120
 AAATCTGGT GATAGCTTC GGCTTGTA AATTGTGACG CTGGTA-TAT TTTGTTGCA 180
 25 ATTGCCTTAT CAGCATTAAAT CGTATTTTTA AGCTGCTCGA TATAAGTCTC AGCGAGTTCT 240
 TTTTGATGAT CATTAGTGTA GAAAATAGCT GTTTnATATT GAGGACCACG GTCTTGATAT 300
 TGACCACCTG TATCTAATGG GTCAATGACT GAGAAAAATA TTTCTAATAA CTTATTGTAT 360
 30 GAGAATAATG CAACATCATA TTGAATTTCA ACAGTTTCTA AATGACCACT CGTACCTGAT 420
 TTTACTTGTT CGTAAGTAGG ATTTTCAATA TGTCGCCCA TATATCCAGA AGTTACTTTT 480
 35 TCTATGCCGT CAAAGGTGTC AAATGGTTTC GTCATACACC AAAAGCAACC TCCGGCAAAA 540
 TAAGCTGTAT TAATATTCAT TTTTGACATC CTTTCCATTA GACCTTAGTA CGATTTATTA 600
 AGAAATCACT TGCTTTTTGA ATTGTTTTTA TATAACGTTA ATAnGnGATT ATnAT 655

(2) INFORMATION FOR SEQ ID NO: 1209:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 213 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1209:

50 CCGnCCCCAA TGGAAATTAC CTTTGGTnCC CCATTTTAAT TGTAAATTA TTGGAGTGTT 60

55

AATCTCACTA TCTTCAAATA AGCTTGTGTA TAGTTTAACA GCTTCTTCAG CTGGGTTATT 180

AAACATTAAA AATGTCGTGA TTTTGGGAAT ATC 213

5

(2) INFORMATION FOR SEQ ID NO: 1210:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 149 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1210:

TTGATTGCAA CTTTGATAGT GAAAACATTT ACATGATGTT GTCAGTGGTG CACGGGCGTC 60

GAGAACAATC ATTGCGTTAT TGACGGTGCT GCCCTGCTGT CTCAGTTTGT TGATGCAAGC 120

ACTACACGAA ACCCAATAGC nTACCAGG 149

20

(2) INFORMATION FOR SEQ ID NO: 1211:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 142 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1211:

AATGTTTCTC CTAATTTTAA ATGTTTAGGA ATATTAACTA GGATAATAAA TCCAGCGACA 60

TCCCAAATGG GATTCATATA AGAATGTTGG ATGATAATnT TGGCCGTTAA TATACATATT 120

35

TTCTATTATA AAATTGGGCA AA 142

(2) INFORMATION FOR SEQ ID NO: 1212:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 163 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1212:

TTTAAATTTA GTCAC TTCAC CTTTAAAGC ATGTT CATAA AATGTTTGCA TCATCAATGC 60

ACGTTCTGAA CCAGAGCCTT CAACACAAAG ATAAATTTGT ACAGCAATAC CGCCTCTAAC 120

50

TCTTCGTTGC GATGATGGCA ATACCACTCA ATCCAACGCT CnG 163

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 765 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1213:

GTCACCTTTC ACGTCTTGAA GAAAAAGGCT ATATAAGAAG AGATCCAACG AAACCACGTG 60
 CTATAGAAAT TGTAAGTGAT CAAACAAATG ATAATATTAA TATGGAAGAA ACGATTCATG 120
 TGCCAGTTAT TGGTAAAGTC ACAGCAGGTG TTCCTATTAC CGCAGTAGAA AATATTGAAG 180
 AATATTTTCC ATTACCTGAA CACTTAACAT CGACACACAA TAGCGACATA TTCATATTAA 240
 ACGTCGTAGG CGACAGTATG ATTGAGGCTG GTATATTAGA CGGAGACAAA GTAATTGTnC 300
 GCAGTCAAAC CATAGCAGAA AATGGAGACA TTATTGTTGC TATGACTGAG GAAGATGAAG 360
 CAACTGTCAA ACGCTTCTAT AAAGAAAAAA ATCGTTATCG ATTACAACCT GAAAATAGTA 420
 CAATGGAGCC AATTTACCTA GACAATGTTG CTGTAATTGG GaaAGTAATT GGTtTGtACC 480
 GCGAAATGTA ATATTTTAAA CCGTTATATA TTATCGTAAT TGTTAAGCCC TCATTTTTTAT 540
 AAATTTTGGG nCTCTTGAAA AAGTTACGTT TTCAAGAGGT TTTATTTATT CTAATCTAAA 600
 TTCAGTTCAA ACAGAAATTG CGAATTGTTA GATAATCTCA TTCTTTA-TA TAAATnTTGA 660
 ATTACAGATT TCTGCAAAAT GTTACACAAA TTAAACTCG CCCCCGTAAT AATTTACAGG 720
 AGCGAGCCAT TATTATTTTT TTCnTTCTTT TTTATTAAGT ATATC 765

(2) INFORMATION FOR SEQ ID NO: 1214:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 495 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1214:

AGTTGCCATC AAAAATACAG AACCTAATnA ThACCTTTTTA TGATTTTTTTG TGAATTCAAA 60
 GTCACTTTCT TTTThAAGCT TTAAATTTCT CCCCATTTTTT TTAGCCCCCT ATAAGGATTG 120
 AATATCAATG CCTTCTTTCA TTAAATTTTC TCTAATTTTC GAAACAAATA ATAATGCATG 180
 TTCTCCATCA CCATGCACAC AAATTGTATC TGCTTGTAAC GTTACTTCCT TATTGTTTTT 240
 TGAAATAACT TTATyyycct TCACCATCTT TAAAACCTGC TTAAGTGCTT CGTCAGTATC 300

ATCAGCAAAC ACTTCAGAAG CTGTAATTAA TCCGACATTC TTTGCTTCTG AAATTAGATA 420
 TGAATTTGCT AATCCTACTA ACACTAGTGA TGGATCAAAG TCATAAACAG CTTGTGCTAT 480
 5 AACGTTTGCh ATTnC 495

(2) INFORMATION FOR SEQ ID NO: 1215:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 243 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1215:

AnTGTCcAGG TACCCGTGGT CCAAATTCG ATGTACGTAG ACCATTGnCT CGTATTTGTC 60
 20 AGATTGTGTT AATGATTGTC TCTGTATTAG GAAGTTTATT CTCCATTTTA ACAATTAGAA 120
 AAATAGATCC GTTAAAGGCG ATTGGGTAGG AGGTGTGCAA ATGTTGAAAT TGAAATGTAC 180
 AAGTCATTAA GTGGATCTAC ATGAGCGTAA GGTcATTGGA ATAGGTTTGC TGTGACTCGC 240
 25 CGA 243

(2) INFORMATION FOR SEQ ID NO: 1216:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 256 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1216:

GGGCTTCITT CTAAATAAC TAAGTGAATT ATCTATATAT GAAGAGGTCT CCAATACTTT 60
 40 TTATATAATT CGTGATTACA TTCTATACTT TTAATTGTAT TTTCTGcGTA GGAAAGTGGT 120
 TTTATTAATT CGATTAGTTC GAACTACAnT CAAGAAAAGG TAAAATTTGT GCATGAGCAA 180
 GAGAAACATA CCTGAACCAA AGGCACCCGA ACAGTACCTA GGCCTAAGCG GCTAATTATA 240
 45 AATGCTCAAG CGCAnG 256

(2) INFORMATION FOR SEQ ID NO: 1217:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1217:

5 TTTATTTTTT CTCTAAACG AAACATTGCT TCTTTTTGTG ATTGTTTTGT ATTCTTGTTT 60
 AACACAACAC TTACGCGCTT CCATTTATCT GTGTATGGAT CTTTATACTT CTCGTAGTAT 120
 CTGTATTTAG TTTGTTTATT TTTGTTTTTA AATTTTTCAA TCCACATGTT TATACCTCCT 180
 10 GAGAGAACGT ACGTTCTGTA AATTTGTAAA AAATAATAAG GGTAGGTGGG CTACCCAAAA 240
 TTTAGTACTA GGTACTAAAT ATGTTATAAT AAAATAAAAA GTAGGTGATA AGATGACTCA 300
 ATTTCTAGGG GCGCTTCTTC TTACAGGAGT TTTAGGTAC ATACCATATA AATATCTAAC 360
 15 AATGATAGGT TTAGTTAGTG AAAAAACAA GATTATCAAT 400

(2) INFORMATION FOR SEQ ID NO: 1218:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 400 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1218:

CCAAGTGTC ATTTTAGATG ATTTTTTTGT TGACCTATCG CCTTAACTGA AGAACTGAT 60
 30 AAATCATCCA TTTCAAAAAT AGGTCTAGAA AAATCTGTTT CGAAGGCTT AAACGATTCA 120
 TATCACGAAT ATTTTAAATC GTTATATCAT TTTCTGTTAA TAATACATCT ACTTGCTTTA 180
 CGGGATCTAA CGAAGTTGTT TTAGATAATT CTTTCATCCA TTTATTTAAA CCTTCAGCTA 240
 35 ACGATTCTAT ATTTTCAATA TCCATCGTCA TACCTGCAGC CATATGATGG CCGCCAAATT 300
 TAGCGATTAA CTCTTGATGT GCTGATAGTA TTCAAACAT CGACACTGA TCAATTGATC 360
 40 TGGCGGAGCT TTTGCATGAT TTTGCTCCCT ATCAATATTT 400

(2) INFORMATION FOR SEQ ID NO: 1219:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 458 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1219:

50 AAGATATTAG ACATAAAATC TAAAnACAGC AGTAAGATGA TTTATGATTA GAACTATCT 60
 55 TACTGCTGTT CACTTTTTAT AATACTTCTG AATGTCTTCA CTTATACTTC TAGTCACAGA 120

5 GCATTTCATTT CTTGTcTAGC AACGTTCTAC TCTAGCGGAA nTAActTAGCT ACCATCCTCG 240
 CTAAGAACCT TTCTTGACTT GTGACAATCG CTTGCTTCTT TCCTCTCCTT CGGcTCTCGC 300
 TTACTCATTT AGCTCTACTA AACTCGTTGC GCTCTTTTCT CGTTTCGTCA GATTCAAACG 360
 TTTTCACTTC GCCAAGCCAT TTTTCTTTGT GTTTACTTTT TATTTTGACG TTTTAGACAT 420
 10 AAAAAAGAG ACCTCACGGT CTCAACTTGG CCTGGGCA 458

(2) INFORMATION FOR SEQ ID NO: 1220:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 494 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1220:

CAAAATTAT TTCTTGTAAG GTACAATATA AGTTAAATGT TGTGTTAGG ATTGATTCTT 60
 25 TCATCTGAGC ATTTAATTGT AAGTCGCTAT ATTTAATTG TTCAGTATTT AmAAATAAAT 120
 TGAAAAATTC AACTGATAAT AAAGGTGTTT TAATCATACA AATACTTGAC GGTTCGTAAA 180
 TAGTGTTTAT TTGTAAGCAC CTCACATATT TGAATAGACA AATAGAGCCC TTTCAGGCTC 240
 30 TATTTGCTGT AATCTTAACA GCAGAAGCTA TTAAGAACTAC CAGTCTTAGC ACAACCAGGA 300
 GTACATAAAC TATGACTTGT AATACGTTCTG TCACCTGCTG AATCATTGTA GTnATTGTTT 360
 GCTTTAACTT GCACGTCTAA ATCAAGAACT TTTTCCATAG ATAAACCTC CTATTATTAG 420
 35 TTGAAGTTAA GGCCTACTTC AATTGTCATA TTATCCCTTT TCACACAACA AATCAATAAA 480
 TTATATAATT ATAT 494

(2) INFORMATION FOR SEQ ID NO: 1221:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 148 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1221:

50 ACTATTATTT TAATAGTGAC GACAGTTTTA ATTTTATTAA GTATTGCTTT AATTACCTTT 60
 ATGCAATACT ATTTACACA AGAnACCGAA AATGCCATAA GAGAAGATGC TAGACGTATA 120
 55 AGTTCACTGG TCGAACAATC ACATAATA 148

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 401 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1222:

TCGCTTTTAC ATCAAATAAA TTATCTTTAT ACAATTTCGT ATGAATTACT TCTTTATAAT 60
 TATAAAAATG AGGAATTTTT TGATTGTTTA ATTCACCTGC TAATTGATTA GCTTTTTGTT 120
 GGTCTTTAAT TGTAACGTCG TGTGGTGCAG TATATTTTAT TTCACCTGAT AAGGACGCTC 180
 TACTTATAGC AGCAAAGCAA AGAACTGAAA CAGTAATCGC TGAAATGATT GCCATGACCG 240
 TAAGTGAAAA AGCATTTTTC TTAATACGAT ACATAATAGA TGATGAGAAC ATGACATCAT 300
 TACACTTAT AACACCTTKT CTAAACtTCT TCACCATTTT AAAAActAGA GATACAGAGC 360
 TTTTAAAAAA TAAGnGTGCC CCAATCACAG TTGACAAnAA A 401

(2) INFORMATION FOR SEQ ID NO: 1223:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1223:

CTTGACCAAA TCCATTTTTT AACAAAAGTG GTCAAGTGAA ATCCCATCTT TCCAAATCAA 60
 AAAAGAGCTA AAGCAAAATG CTCTAGCCCT TGATATTACT GATTTCCTCAT CATTGTnTAG 120
 CGTATTAAATA TTGCTTCATG TACTGATCTC TTTCCCATTC AGAGACTTGA GTTCTG 176

(2) INFORMATION FOR SEQ ID NO: 1224:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1224:

AATACTTTTA CGATAATTAT AAAGGTTTCA TCAAATCTTA GGGCAAAATG TATAATATTT 60
 TTATGCAAAT TTAAAACAnT AACACTTATT TCAAGGTTCA ATATTTTGAG AATAAGGAGT 120

AATGCCAATC AAACTCGCTC GAAATAGGGG GAAACGAGGC nATCATTITT GACAATGAGT 240
 GTCAATAGGG AAGAATAACA ATTAGGAATG ACTCATAGGA GGA 283

(2) INFORMATION FOR SEQ ID NO: 1225:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 217 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1225:

AAAAGAAGAG AAGAAAGTTG AAGAACCTCA AGCACCTAAA GTTGATAACC AACAGAGGT 60
 TAAAACTACG GCTGGTAAAG CTGAAGAAAC AACACAACCA GTTGCACAAC CATTAGTAAA 120
 ATTCCACAGG CCATTACAGG AATTTAAGTC GGATACCACG TGAATnACGT CAGGATGTCh 180
 GCCGTTTACA TGACACGCTC TACATATTCA CCACTCG 217

(2) INFORMATION FOR SEQ ID NO: 1226:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1226:

TGGGTACTAT TTATTACGTA AAACTTTTCG TTAGCGATGC CGGCATTGCA AGAGCAAGGT 60
 TTACAAAAGC GGACTAGGTT TGCACTTTCT GCTGTTCCAT CGCATATGGA TTAGTAAGGT 120
 CTTTATGGGn ACTGTAAGTG ATCCGAGCAA TGTCGGATAA TCTAGGTCTT GGATAG 176

(2) INFORMATION FOR SEQ ID NO: 1227:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1227:

AATGAAAATA AAGAAGATGA TTTTGTGCA TATGGGTCTC CAGAACATAA TTATCAATTT 60
 GGTGGAAGTT TAATCAAAAG TAAAAATTA AGCACGTTAT TAAAACCAGT ACATCAAACC 120

GCGGCCGTTT TAATTTTGTG TTCAAGTTAA TCATACTTAC ATATAAAAAT GAAAAGTTGA 240
 ATTTTATAGTA CAATTTCCAC TGTATTTATC TGTTAACCAC TTTATACCCG CAATATTTTCG 300
 5 TTGATCTTAT ATAAACACTT TTGGCTGCTT TTCCTTGACC TAGATTAAAA TTATAGATTT 360
 TCCTTGTAAG AGTGTGTTTG TAGTATCATA AATATGTAAA 400

(2) INFORMATION FOR SEQ ID NO: 1228:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 150 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1228:

20 ATCTAGCATG CTTTGGAAA ACGATAAGAA TTTGGTGTGTT GAATTTTCATA ACCAACACCA 60
 GCAGTTTCAA CAACTACGTG TGTAGGATAT AAATGTGTGA CTTGACCTTT GACATACGCG 120
 25 TACATTATAG GCACATCCTT ACATACTCAN 150

(2) INFORMATION FOR SEQ ID NO: 1229:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 138 base pairs
 30 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1229:

35 TCGGTACAGT GCATTTGGTG CGTCATGGTT TACAAGGAA AAATAAGTGA TCACTATTTA 60
 TCAGTATGGG AAAAAGCAAC GACGTATCAA ATGTACCATG GCTTAGCATT ATTAATTATA 120
 40 GGTGTAATTA GTGGTACA 138

(2) INFORMATION FOR SEQ ID NO: 1230:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 182 base pairs
 45 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1230:

55 TTGGTATGTA CCTCCAATAG TAATTTCAAT AACTTTGTCT GTTGAACACT AAGAGCAATT 60

CCTTGATTAT nCTAAGAAAA GTAAAAGCAC ACGGAGTATC CTATCAGAAA ACCAGTATAC 180

CA 182

5

(2) INFORMATION FOR SEQ ID NO: 1231:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1231:

TCTCTTTAAA TTGTACACAA TTCTGGATAA CTATCCCAT TTCTGTGGAT AAACACCTTG 60

ATGTCTTATT ACTTATCCTC GTTTTTATAT CCCATATTTA TCAACGGkTA TCGCTTTTTT 120

20

TCATAAAAAA aCACTACCGT ACCyCTTATT TAAGAAATAC AGTAGTATTA TTTTCAATAT 180

ATTTTAGCTA GChTTTCTAC GTCGTACAGT CGCGATTAAA CCTATACCTG CTA 233

25

(2) INFORMATION FOR SEQ ID NO: 1232:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1232:

35

GTTTTTGGTA AAAGTACTAA TATTaCCAA GTATTGAGTC ATATGTTAAA ACAAGATGCG 60

AATGATATTG GCTTTGCTAA ATTACTACAA AATGAGAATA ATCGTATGAG TTATAACGAG 120

40

TTAATGAGTG AATGGCAATC ACATCAACGT GCATTTTTAG AAGAGTTGAG GCATGTTGAA 180

ATGTTAAAAG AAGAATCTAT TAGAGCATAT GATGTTTATA AAAATTGTGA GTCTTTCTCT 240

AAGATTGAAC AGGTTATTAA TAGTGAAAAA ACAAGTATTG AAGAACAGGT ATATCATTTA 300

45

GATAATGAAA CGTTACGAGA CAATAAAGAA ATAGAAGTTT GGGATAATCG ATTTAATTAn 360

ATTGnTAGCC CAATGnCCC TTAA 385

(2) INFORMATION FOR SEQ ID NO: 1233:

50

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1233:

CTGAAAAGAT TCAACTAATC TGAGTCAGGT TACAATCTCA TGATAACCGC TCGCACCTCG 60
 5 GAGAAAGACG CTCAAACAGT ACAGAGCTCG ACCGCAGGAT ATTTTTTGTG AAAATGAACG 120
 TATGAGCTGT TCAGAnGGCA CTTTATTGCG CAATAAAACG GTAAAAATCC TGAAGTGAAG 180

(2) INFORMATION FOR SEQ ID NO: 1234:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1234:

ACCATTTTAA TTTnChTTAT GACCTTGTAT GCCATTATCA aCAGTTATGA TTAATGTTAT 60
 20 GCyTTCATCA TGAGCATTTT TAAATGCTAG TTCGTTTGGT CCATATCTTC GGTAAAGCGG 120
 TTAGGAATAT GCCATCCTAC TTGTGCACCT AAAAGTTGTA ATGTTATCAC TAAAATTGTA 180
 25 GTTGCGGTAA CACCGTCGGC ATCGTAATCA CCATAAACTA GGATTTTCTC ATCATTGCT 240
 ATCGCTCTTT TAATTCCTTC AATAGTCTTA GTCATATCGC TCAATTGnAG TGCATCATGA 300
 30 TTGATATCTG TATCTGAAAT GATGGATTCT ATTGCTTGn CATCAATAAT CGATTTACTT 360
 nCTAATATTT TTTTACGAT TGGCGTTAAC TTTAATTTTG ATGTTAATTC ATCACTTATG 420
 TATTCAGCTG GTTTAGTTAA TTTCCACTTA TAC 453

(2) INFORMATION FOR SEQ ID NO: 1235:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1235:

CGTGATTAAG TTGGATACTA AAATTTGGGA TGTTAGAGAA TGTGATGAAG ATTTnCAGCA 60
 45 ATGTCCTGAA ATTAATGAAA TAAGAGACAT TGTTTTAAAG CGGTGGTTTG ATAGGTTTAC 120
 50 CAACTGAAGC AGTTTATGGA CTTGCAGCGA ATGCGGCAGA TGA 163

(2) INFORMATION FOR SEQ ID NO: 1236:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 base pairs

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1236:

GTTAACACAT CGCTTATAAC GTGTAAATAT AAACACTGTT TTATAAAAAC TTTTATACC 60
ATTACCGGCA CCGATAAATG TGTACTATGG AGCACAACAA TAACTGGAAT ACTTnGATTT 120
AATCCCGCTA TAACATTTCC TAAT 144

(2) INFORMATION FOR SEQ ID NO: 1237:

15

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 204 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1237:

ATAGTTTAAG GTGATATTCA AACAAATATTA GGTTTGATTT AACTATCAAG CATCTACTTT 60
GTCATGTTTA ATACCAGCAA ATAATAAAAT CCATGTCACG ATAATAAATG GCATGTTAAT 120
GCAGGGTAAG CCCAACCGGT CCAGCCATGG TGGTGGTACC CTAnTTGGCC nCCACCTACC 180
GGTTAATTAA CCACCCCCAA AAAA 204

(2) INFORMATION FOR SEQ ID NO: 1238:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 157 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1238:

CGGTCATTTT GGTtATTnAA CAACTATAACC TGAGGGGTGG TGGGTAAATC GTATTGCCGA 60
TTCAGTTTTG TTAATATGCT GACCACCTGC ACCAGAAGCT CTGAATGTAT CAACTGTGAA 120
TAGCATCCGG ATTGGTTTCA ATCTGCTATG GCATCAT 157

(2) INFORMATION FOR SEQ ID NO: 1239:

50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 164 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1239:

TTAAAGCATG TTCATAAATG TTTGCATCAT CATGCACGTT CTGGACCGAG CCTTCAnCAC 60
 AAGGTAATTT GTACAGCATA CCGGCTCTAA CTCTTCGTTG CGTATGCCTG CAATTTCTTA 120
 CCTCGATACT TAAGTCAATT TTCCGGGCAT AAGGGTGTTT ATTT 164

(2) INFORMATION FOR SEQ ID NO: 1240:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 239 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1240:

TTGAAAACT GAAAACGGTG AAGTTTTACA AACAACTAAA ACAGATAAAG ATGGTAAATA 60
 TCAATTTACT GGATTAGAAA ATGGAACCTA TAAAGTTGAA TTGAAACACC ATCAGTTACA 120
 CACCACACAA GTAGTTCAGA ACTATAAGTT GATCAATGTC TCACACAGTG CATAATnGAA 180
 CAACATGCCG GTCTCAACAC TCACTAGGCA GAGGATCATA ACGGTCGTnG TAAGCTCGG 239

(2) INFORMATION FOR SEQ ID NO: 1241:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1241:

TGATAAATAA TAGCGCCATC ATAAGTTGCA GGGTGTAAT CATAATTTCT ACCATGAGTC 60
 TTTGCCTTGA AAGACCGTGC ACCACTTTCT TTTAAATCGG TAAACGATAA TGnCAAGTTT 120
 AAAATAATTT GCGGTGTTGG CATTGAGTA TC 152

(2) INFORMATION FOR SEQ ID NO: 1242:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1242:

ATCTTTACCT TTCACTGTTC ACTTATGTAC CATAATACTT CTGACAGTTA CTAATTAACA 120

GCAACTCTAA CTCATTTATA TTTAACTA 148

(2) INFORMATION FOR SEQ ID NO: 1243:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 588 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1243:

GGTACCAATC GATCGTTACA AACGTTTCGT TAAAGACGAT AAAAAAGTAC CAACAGGCAA 60

TGAATATCGT GAATTAGTAT TAAAAGCAAT TCACATGATT ATGCTTGGTT TCTTGTATAA 120

ATATATTGTT GCTTACTTTA TTAACACATA TGCAATCATG CCGTTACAAT TAGACTTACA 180

TGGCTTTGTC AATTGTGGT TATATATGTA CGCATAACG TTATATTTAT TCTTTGACTT 240

TGCAGGTTAT AGTTTATTTG CGATAGCATT TAGTTATTTA TTCGGTATTA AAACACCACC 300

AAACTTCGAT AAACCTTTCA AAGCGAAAAA TATTAAAGAT TTCTGGAATA GATGGCATAT 360

GACATTATCA TTCTGGTTCA GAGATTGTAT TTACATGAGA TCTTTATTCT ACATGTCTCG 420

TAAAAAATTA TTGAAGAGTC AATTGCAAT GTCCAACGTG GCATTCTTAA TCAACTTCTT 480

CATAATGGGG ATTTGGGCAT GGAACGAAG TGTAATACCA TTGGTTTAAAG GGGTTAAACC 540

ATGGCCAGCC TTGGTTAAAG GGTAnGnCC AATAAGGACG GTGGGCGG 588

(2) INFORMATION FOR SEQ ID NO: 1244:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 376 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1244:

ATTGCTGGAG AAAGGCAAGG CTTTGCAGGC GGATTGAACT CGACATTCAC TAGTATGGGT 60

AATTTCATAG GTCCTTTAAT CGCAGGTGCG TTATTTGATG TACACATTGA AGCACCAATT 120

TATATGGCTA TAGGTGTTTC ATTAGCAGGT GTTGTTATTG TTTTAATTGA AAAGCAACAT 180

AGAGCAAAAT TGAAAGAACA AAATATGTAG CATAAGTATT TTGGTGTATA TTGATATAAA 240

GTAAAGCGTA ATATTATGAA TGATTAGCAT CGTTTTTCTT ATGAATTTTA TTAAGAAAAT 300

GTTATATGTh ACAAAA

376

(2) INFORMATION FOR SEQ ID NO: 1245:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 288 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1245:

CAATTTGTTT TAAAGCAAAT TCTTTACATT TCTCACGGAA TTCAGCTGTT GACATTTTCT 60
 TTCGGTCAAC ACCTTTTTTTC GTTAATGGCT TGTTC AATTG GGTAAACCAT GTGGTTATCC 120
 CCCAACCTGG GTACCGTATG GGGTGGCCAT TAGGAAACCC CTTGGCAATA GGTTTTTTAA 180
 TAAACCGGTA nCAAATTAAA AGGTCCTTTT TTA AAAAATT TTTGGTTTCC AAAGGGGCCA 240
 ATGGnTCCCC CAATTAATGG TTAAAGGTTT TAACCCCAAT TTTCCGGC 288

(2) INFORMATION FOR SEQ ID NO: 1246:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 161 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1246:

CATCATCAAT AGAACGGTAT TCGCAACGTC GTAGAAATCC CTAAACGAC GAGAACAAAG 60
 CAAGTCGACA TCAGTAAAC ATTAAAGATA TTATGGCTGA AGCAGAGAAC TTAATAACTT 120
 TTTATAAATC TATACnTCC AAAATGTAAA ACTACCAAAA T 161

(2) INFORMATION FOR SEQ ID NO: 1247:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 421 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1247:

TCACTACGTT GTTTAATTGT GGTAATTGCA GTTTTACAGG CTTGTCTAGT GAAGATTTAA 60
 AACTTAATTC AAGTGTTTTA GTCATGAATA TGTCTCTCTG ATTAAATTGA TAAAGATTTG 120

TTAATTTGGT CGTTTGATGC ATCAGGGTTA ATGTTAGCGA ATCGACGCTT AAATTCTGTT 240
 TGTTCGCCGT TAGCGTCTAC TTTAGTAAAT GaTAATACAA TAGTGATGTG GTTTATTTtA 300
 CTCATATTTT AAAACCTCCT TTCACACTAT ATATCGAAAC AAAATAATAA AATGGCTAAT 360
 TTTATTTTCT ATGTTTAAAA TCTATAAAAn AGGCAATAGA TATGTGTAAC TAAAATATAN 420
 G 421

(2) INFORMATION FOR SEQ ID NO: 1248:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 409 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1248:

ATAAGCTGCT GTTGATAATC CAAAGTGCTAA AGTTAAATAT AATATAGTGC CAAAAATTTT 60
 ATTCTTCTAT AAAAACACCT CTCATATTT AATCCATATC TCTTTTCGT TATTTATACA 120
 GAATTTTCTG AAAATAGTCA AGGTTAATTT ATTTGCTGCA AATTTCAACC ACTTTATGTT 180
 AAATGGACAT CAAAGTATAA TTGTATTATT ATTAAATTGT GCAGTGTATA AATAGACGAT 240
 ATGCAGGAAC ATGTCGCCTA TTGAGCCCGT TAAAAAGACG GTGACTAAAT GAGATTTTCT 300
 TTAACCATCA TTCGTTGTCA AAGTTTGGAA ATgATGGTTG TTTTtTATTG TTTAAATTAA 360
 TCATTGCGAC TTCTATATCA ATGAAAGTTA TCTTAATATA TGAACttCA 409

(2) INFORMATION FOR SEQ ID NO: 1249:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 217 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1249:

CTTTCAATTA CGAACAAATC ATTACTAGTT TATTTAATGA TAATGACTGn AATTTAGCAA 60
 AAAACGTAAC ATGATCATTT AGCTTGAATT GGTAAAACTG TTTATCCAAC ATTTCCAGAC 120
 CAAATCACAC CGGTAAAACG TCAAAATATA TTGTnTCTAT GCCCGTTTTT ACCAGTTATC 180
 TCATTGCAGG TATATTGATG AATTCACAAG TTGGTAT. 217

(2) INFORMATION FOR SEQ ID NO: 1250:

(A) LENGTH: 176 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1250:

10 CAGAACGTGC GGATCGTTTT TATGCAACCT AAACCAGGTA CGGATTTAGC GTGGTTAAGT 60
 GGCAGTGCAC TAAGTATATT CATTGATCCT GCATTTACAC GATGAAAGCA TTTATTGCAT 120
 GAGTGGGTAG ATGCATTTTG CTGGAATGAT TCACAAATGC ATTAGAGnAC ATTTAC 176

15

(2) INFORMATION FOR SEQ ID NO: 1251:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 212 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1251:

25

TTTCAACTTT CTAAATTATA ATATATCTTT TAAAATAAG CTAGAATTTT TATATAATAA 60
 ATGTTAATAA CGAAAAGGGA TGATGCATAG TGATCGCAAG CACGCCAGAG GCCGATTGTA 120
 30 TATTGCGAGT TAGTTTATAT GGTITGGGAT GATATGGATA GAATTGGTAA ACATCTCCTA 180
 AAGCATGGTA TAGTGCAATT nAAAAaCTGT GT 212

(2) INFORMATION FOR SEQ ID NO: 1252:

35

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 131 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1252:

45 CAGTAAGTCA ATTGAAATGG CTAATTTATT AAATCTATAA AAGCATAGAG CACTGTTGTG 60
 AGTTCATAAT CAAAGATTTA TAAATGTnTC AGACTGCCAA TATAACATTT AGGACCTAGA 120
 ACATTGATTA T 131

50

(2) INFORMATION FOR SEQ ID NO: 1253:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 462 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55